Theoretical Modeling of Hydrophobin Films at Air-Water Interfaces: Adsorption Kinetics and Mechanical Response

Dissertation

zur Erlangung des Doktors der Naturwissenchaften der Naturwissenschaftlich-Technischen Fakultät der Universität des Saarlandes.

von

Neda Safaridehkohneh

Saarbrücken
2024

Tag des Kolloquiums: 10.12.2024

Dekan: Prof. Dr.-Ing. Dirk Bähre

Berichterstatter: Prof. Dr. Ludger Santen, Prof. Dr. Albrecht Ott

Akad. Mitarbeiter: Dr. Adam Wysocki

Vorsitz: Prof. Dr. Jochen Hub

English Abstract

Hydrophobins are a family of small globular proteins expressed by filamentous fungi. Remarkable water-repellent properties enable them to cover interfaces and modulate surface wettability. Despite their potential applications across various fields, from the food industry to medical targets, the mechanisms underlying their linear adsorption, crystal formation, and film stabilization are not fully understood. This thesis investigates those critical aspects of thin films at liquid-air interfaces formed by hydrophobins in two parts. First, the adsorption kinetics of class II hydrophobins HFBI and HFBII, along with engineered variants, are analyzed through a combined approach of theoretical modelling and experiments. Experimental results reveal intriguing linear kinetics for wild-type proteins until saturation, contrasting with Langmuir kinetics for bulky variants and concentration-dependent kinetics for charge-mutated proteins. A stochastic model incorporating a subsurface layer and spin properties of proteins elucidates the underlying mechanism, highlighting a two-stage adsorption process and the role of diffusive motion and final adsorption rate in kinetics. The second part studies the mechanical response to external stress using a continuum space model incorporating thermal fluctuations and angle-dependent potentials. Simulation results reveal the crucial role of angle-dependent potentials in maintaining order and stabilizing films, contrasting with systems interacting only isotropically. Notably, the reordering process before rupture underscores the dominance of angle-dependent potentials in stabilizing structures. Moreover, the study emphasizes the challenge of achieving ordered parameters in off-lattice models due to persistent local defects. Through these investigations, this thesis contributes to a deeper understanding of hydrophobin behaviour and sheds light on potential applications in biodevice.

Zusammenfassung

Hydrophobine sind eine Familie kleiner globulärer Proteine, die von filamentösen Pilzen exprimiert werden. Dank ihrer bemerkenswerten wasserabweisenden Eigenschaften können sie Grenzflächen bedecken und die Oberflächenbenetzbarkeit modulieren. Trotz ihrer potenziellen Anwendungen in verschiedenen Bereichen, von der Lebensmittelindustrie bis hin zu medizinischen Zielen, sind die Mechanismen, die ihrer linearen Adsorption, Kristallbildung und Filmstabilisierung zugrunde liegen, nicht vollständig verstanden.

In dieser Arbeit werden diese kritischen Aspekte von dünnen Filmen an Flüssigkeits-Luft-Grenzflächen, die von Hydrophobinen gebildet werden, in zwei Teilen untersucht. Zunächst wird die Adsorptionskinetik von Hydrophobinen der Klasse II, HFBI und HFBII, sowie von konstruierten Varianten durch einen kombinierten Ansatz aus theoretischer Modellierung und Experimenten analysiert. Die experimentellen Ergebnisse zeigen eine faszinierende lineare Kinetik für Wildtyp-Proteine bis zur Sättigung, im Gegensatz zur Langmuir-Kinetik für voluminöse Varianten und zur konzentrationsabhängigen Kinetik für ladungsmutierte Proteine.

Ein stochastisches Modell, das eine unter der Oberfläche liegende Schicht und die Spineigenschaften der Proteine einbezieht, klärt den zugrundeliegenden Mechanismus auf, wobei ein zweistufiger Adsorptionsprozess und die Rolle der Diffusionsbewegung und der endgültigen Adsorptionsrate in der Kinetik hervorgehoben werden.

Im zweiten Teil wird die mechanische Reaktion auf äußere Belastung mit Hilfe eines Kontinuumsraummodells untersucht, das thermische Fluktuationen und winkelabhängige Potenziale einbezieht. Die Simulationsergebnisse zeigen die entscheidende Rolle der winkelabhängigen Potenziale bei der Aufrechterhaltung der Ordnung und der Stabilisierung der Filme, im Gegensatz zu Systemen, die nur isotrop wechselwirken. Insbesondere der Prozess der Neuordnung vor dem Bruch unterstreicht die Dominanz der winkelabhängigen Potenziale bei der Stabilisierung der Strukturen.

Darüber hinaus unterstreicht die Studie die Herausforderung, geordnete Parameter in Modellen außerhalb des Gitters zu erreichen, die auf anhaltende lokale Defekte zurückzuführen sind. Durch diese Untersuchungen trägt diese Arbeit zu einem tieferen Verständnis des Verhaltens von Hydrophobinen bei und wirft ein Licht auf potenzielle Anwendungen in Biodevices.

Publication

Dynamic Assembly of Class II Hydrophobins from T. reesei at the Air–Water Interface

Hendrik Hähl, Alessandra Griffo, Neda Safaridehkohneh, Jonas Heppe, Sebastian Backes, Michael Lienemann, Markus B. Linder, Ludger Santen, Päivi Laaksonen, and Karin Jacobs

Contents

Chapter 1

Introduction

Hydrophobins are a family of small globular proteins expressed by filamentous fungi. They have a relatively sizeable water-repellant surface area, which makes them able to cover interfaces and change the wettability of surfaces [22, 41]. Fungi employ their strong tendency to cover surfaces in biological processes, such as coating the surface and air cavities in fruiting bodies, which help the fungi survive and adapt to the environment [8, 48, 62]. Due to their elevated surface activity and non-toxic native origin, Class II hydrophobins HFBI and HFBII find various applications, from the food industry to medical targets [19, 25].

The inspiration for this project comes from the experiments performed in the Jacobs group at Saarland University, where the adsorption behaviour of class II hydrophobins HFBI and HFBII were investigated.
Class II hydrophobins HFBI and HFBII are compact proteins of $7.5\ kDa$ with a hydrophobic patch of about 12 percent of the surface area [21, 41, 57]. The adsorption from the bulk solution leads to a membrane comparable to a monolayer [20, 53, 65]. The interfacial film consists of observable two-dimensional structures, not fusing even in close affinity [20, 41]. Remarkably, the adsorption follows unusual linear kinetics rather than the expected Langmuir kinetics of proteins [20].

While there is massive interest in considering applications and mechanical properties of films formed by those hydrophobins and their mutants, the mechanism behind the linear adsorption, the reason for, and the contributing interactions that lead to crystal formation and stabilized films are not fully understood and not studied yet. This knowledge can result in controlling film formation and opening new windows for their applications in biodevices.

This thesis focuses on the formation and mechanical properties of monolayers resulting from the self-aggregation of class II hydrophobins HFBI and HFBII.

To investigate adsorption kinetics, this thesis introduces a stochastic model wherein proteins are viewed as particles possessing an internal degree of freedom, adsorbing to a two-dimensional lattice. The model suggests a "two-stage adsorption process", a concept that refers to the initial adsorption to the subsurface followed by a final adsorption to the interface, highlighting the significance of diffusive motion and a high final adsorption rate relative to other rates, resulting in linear adsorption kinetics. The Monte Carlo scheme and Metropolis-Hastings algorithm are employed for studying the adsorption kinetics, providing a comprehensive understanding of the process.

The second part of the project focuses on contributing interactions that stabilize the crystal structure in hydrophobins. Here, a continuum model approach is utilized. This approach is capable of analyzing thermal fluctuations in the film. The results conclude that anisotropic angular interaction is essential in stabilizing hydrophobin layers, accounting for their high elasticity.

The thesis is organized as follows: Chapter two will encompass background information from Physics and Biology. This chapter will explore essential terminology, fundamental models, and general background information crucial for comprehending the studied systems. Chapters three and four will focus on the fundamentals of phase transition and, specifically, the current understanding of two-dimensional phase transition. These chapters will set my work, presented in chapters five and six within the broader landscape of recent scientific research. The concluding chapter will summarize the accomplishments, conclude the proposed work, and offer insights into potential future research programs that integrate both investigative threads.

Chapter 2

Fundamental Concepts: Bridging Biology and Theory

This chapter provides essential terminology and background information, from biology to theory, that are all essential to capture the content of this thesis. It describes proteins and their structural compositions involving amino acids and proceeds to elaborate on the characteristics of hydrophobins, concluding with a brief overview of stochastic processes. The first section reviews the fundamental terminology of proteins. Initially, amino acids are introduced as organic constituents of proteins. Subsequently, the discussion extends to the four distinct levels of protein structure. The general structure of this part of the chapter is based on references [7, 30, 46]. Then, a comprehensive review of hydrophobins is presented, highlighting them as a group of small globular proteins with unique properties. Finally, the focus shifts to a more detailed exploration of hydrophobins of class II, the subject of study in this thesis.

The second section focuses on the theoretical concepts applied in this study. It briefly introduces the DLVO theory, a framework for describing the interactions among colloidal particles in a solution. The section then considers stochastic processes and explores selected models essential for a comprehensive understanding of this study. The approach in this section follows references [12, 31, 60].

2.1 Biological background

Proteins are one of the primary building blocks of biological systems. Proteins are macromolecules consisting of, on average, several hundred amino acids. They carry out diverse functions, all essential for a living object, from the simple formation of the system's structure to highly complex operations such as transporting objects, promoting intracellular chemical reactions, or DNA replication. The number and arrangement of the amino acids uniquely determine their physical and biochemical properties.

2.1.1 Amino acids
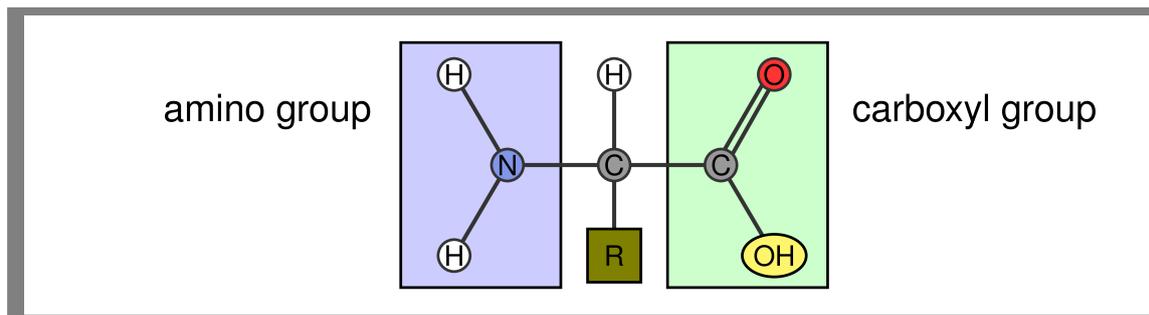
Figure 2.1: The general structure of amino acids: Amino acids consist of a central carbon called the α-carbon, attached to an amino group, a carboxyl group, and a side chain group denoted by R symmetrically. The R group determines the chemical nature of the amino acids, e.g., acidic, basic, polar, and non-polar.

Amino acids are organic molecules containing a central carbon called alpha carbon attached to an amino group and a carboxyl-acid group, as shown in figure 2.1.

Amino acids are differentiated by a group of atoms bonded to the side of the alpha carbon called the residual side-chain group or, equivalently, the R group. The side chain group determines the chemical nature of the amino acid, e.g., polar, non-polar, and charged. Although more than 400 amino acids are known, only 20 of them commonly form proteins. Rather than the complete name, it is common to use three or one-letter abbreviations of the amino acids. Those proteinogenic amino acids, their chemical properties, and both three and one-letter abbreviations are shown in figure 2.2.
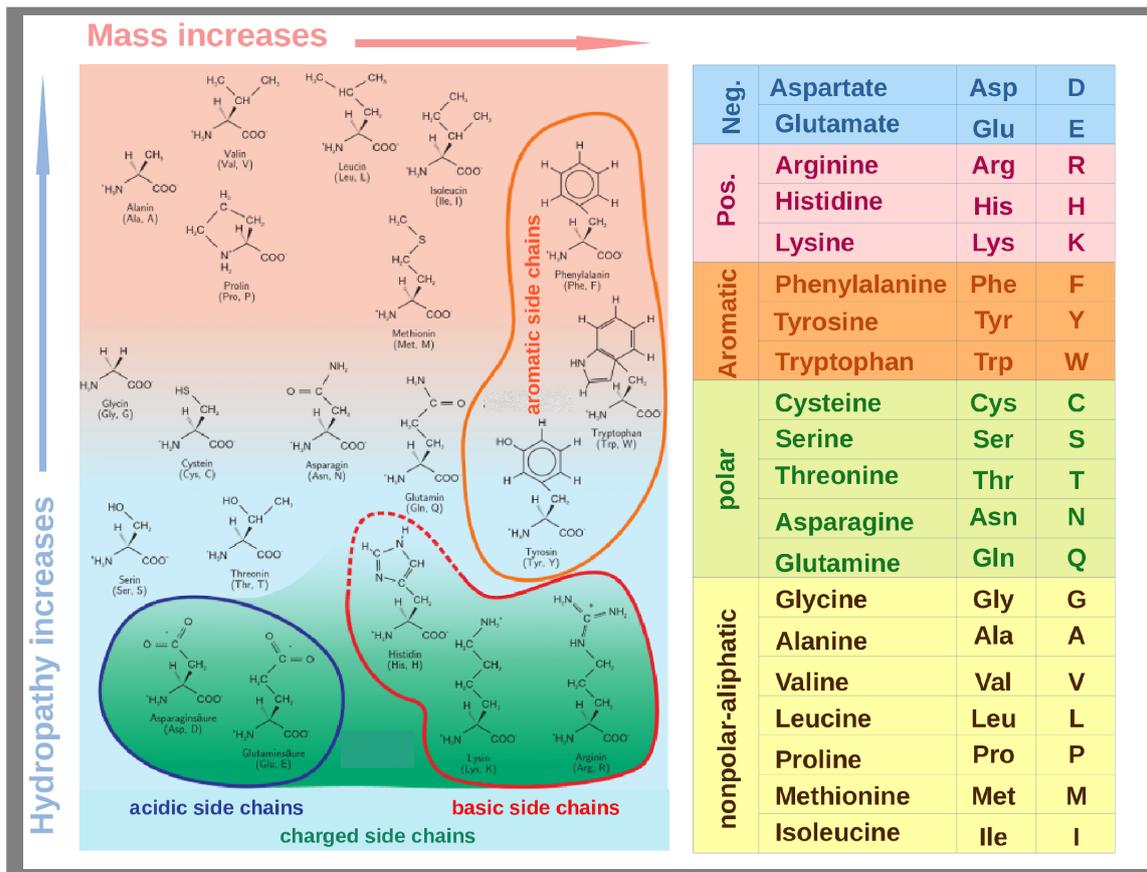
Figure 2.2: Left: Molecular structure of the common proteinogenic amino acids in detail, sorted by size and hydrophobicity. The figure is adapted from [18], originally based on [54]. **Right:** The complete name, three-letter, and one-letter abbreviations of twenty amino acids side chain groups are sorted based on the amino acids chemical nature adapted from [7]

Aliphatic amino acids, highlighted in yellow in the panel, are those of nonpolar and hydrophobic R groups. The hydrophobicity increases with the number of carbon atoms in the hydrocarbon chain. The aliphatic amino acids from more minor to higher hydrophobicity are glycine, alanine, valine, leucine, and isoleucine. Aliphatic amino acids are commonly buried inside the protein molecule; however, among them, glycine and alanine are ambivalent, meaning that they can be found outside of protein molecules, as in the case of hydrophobins.

The characteristic feature of hydrophobins comes from a relatively sizeable connected surface area formed by these hydrophobic, aliphatic residues called the hydrophobic patch on the surface [39]. As the rest of the outer region is hydrophilic, they have both hydrophobic and hydrophilic heads, making them strongly amphiphilic molecules. This amphiphilic nature imparts particular properties to hydrophobins, as

discussed shortly in Section 2.1.4.
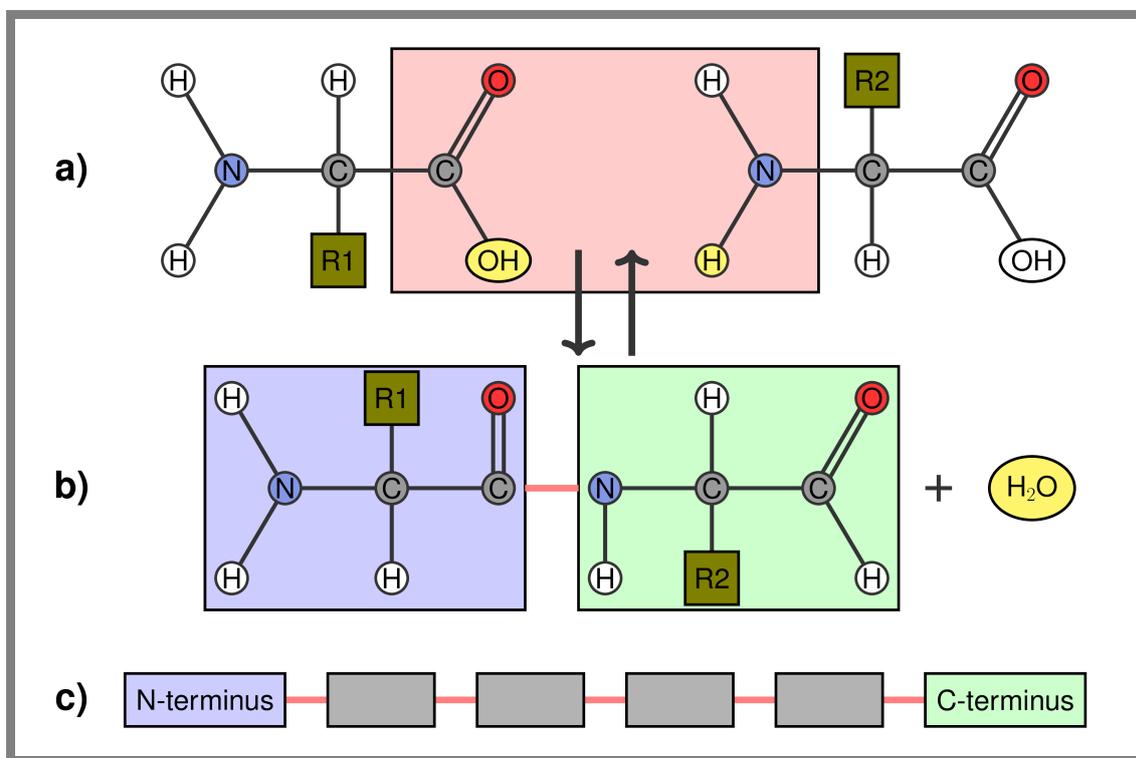
Figure 2.3: (a) A covalent peptide bond joins two amino acids to construct proteins. A peptide bond is formed when the carbon atom of the carboxyl group shares electrons with the nitrogen atom of the amino group of the second amino acid. The formation of a peptide bond is accompanied by the release of a water molecule. The pink rectangle pronounces the atoms participating in the peptide bond formation. **(b)** Two amino acids joined by a peptide bond are shown. A polypeptide chain always starts from the left by the amino terminus, known as the N-terminus, and ends by the carboxyl terminus, called the C-terminus. Blue and green rectangles mark the N and C-terminus, respectively. **(c)** A polypeptide backbone of six amino acids is shown.

A protein comprises a specific sequence of amino acids linked to a repeating structure called a polypeptide backbone. A peptide bond is a covalent bond that connects two amino acids by sharing electrons of nitrogen of the amino group to the carbon of the carboxyl group of the second one, as illustrated in figure 2.3. A particular backbone chain shows a specific directionality; since peptide bonds link the chain of amino acids, each has two ends, one of the amino groups called the N-terminus and the second of the carboxyl group called the C-terminus. Other interactions, e.g., hydrogen bonds and disulfide bridges, play a role in constructing the three-dimensional structure of a protein.

2.1.2 Protein structure
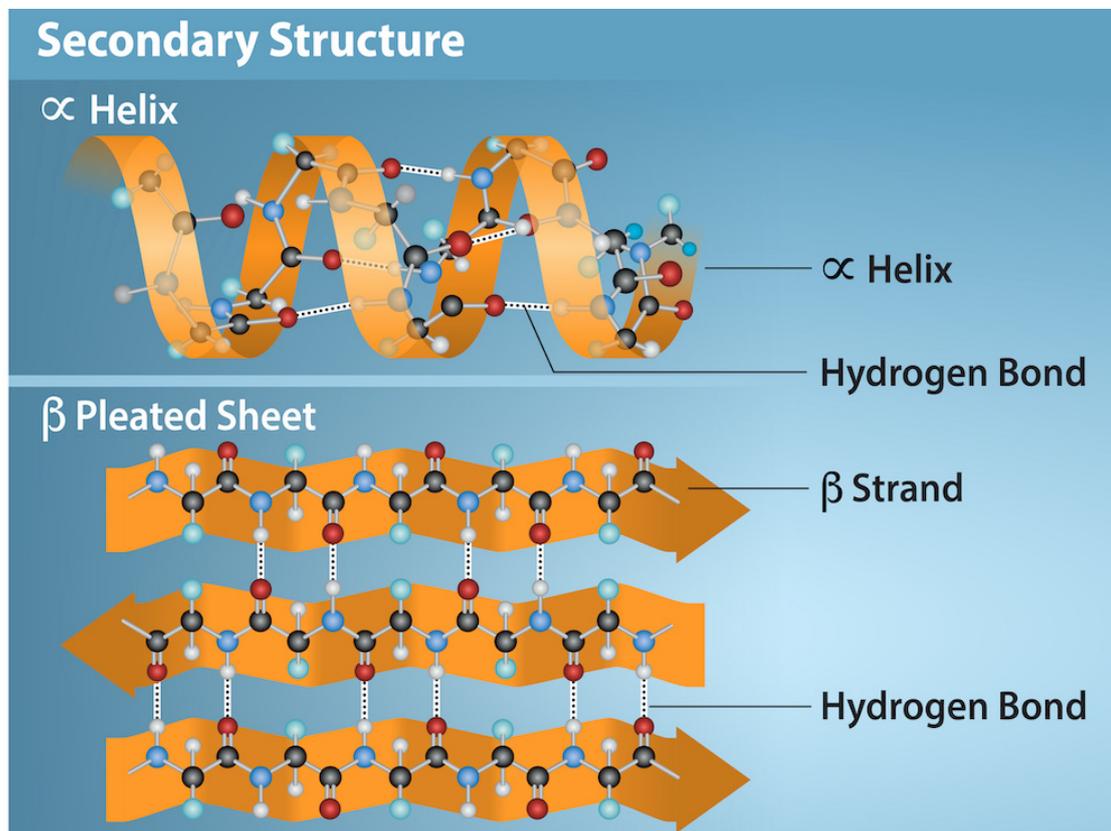
Figure 2.4: α-helix and β-sheet structure: The α-helix and β-sheet structures are common motifs defined by the specific hydrogen bonds between the amino hydrogen and carboxyl oxygen atoms in the polypeptide backbone. In this figure, carbon atoms are depicted in black, hydrogen in white, nitrogen in blue, and oxygen in red. The figure is created by Rao, A., Ryan, K., Fletcher, S., and Tag, A., Department of Biology, Texas A & M University and reprinted from [7] under the **Creative Commons Attribution License**.

The structural composition of proteins is comprehensively discussed through four distinct structures: primary, secondary, tertiary, and quaternary. The primary structure pertains to the actual sequence of amino acids connected by peptide bonds. Atomic interactions between amino acids induce a local folding of the polypeptide chain, resulting in a compact arrangement of the main chain known as the secondary structure.

Two prevalent substructures, depicted in Figure 2.4, are the α-helix and β-sheet. The stability of both substructures is achieved through hydrogen bond formation. In

the α-helix structure, hydrogen bonds link one amino acid to the other four down the chain, creating a bonding pattern that imparts a helical structure to the chain, with each turn typically involving 3.6 amino acids. On the other hand, the β-sheet structure involves two or more segments of the chain lining up next to each other, forming a sheet-like structure stabilized by hydrogen bonds.

The three-dimensional structure of the polypeptide chain, resulting from interactions between the R groups of amino acids, is termed the tertiary structure of the protein. This unique and stable conformation arises from the sequence alignment and is determined by energy considerations. Under favorable conditions, the unfolded chain consistently folds into this distinctive 3D structure.

The folded structure of a protein plays a pivotal role in the aggregation of multiple proteins into multimers, forming the quaternary structure. The 3D configuration of proteins influences factors such as cooperativity, surface activity, and the ability to form biofilms. Specific proteins, such as insulin, consist of multiple polypeptide chains known as subunits. These subunits are bound together by various interactions, including hydrogen bonds or weak forces like London dispersion forces, which contribute to constructing the protein's quaternary structure.

2.1.3 Protein functions in biological systems

The essential role proteins play in almost all functions of a living organism cannot be stressed enough. Transport proteins carry substances into cells, out of cells, or within cells. Motor proteins are essential to cellular transport, while enzymes catalyze metabolic reactions. Proteins also contribute to diverse regulatory processes, influencing gene expression, intercellular communication, and signaling mechanisms.

The presence of proteins on surfaces can lead to the formation of biofilms, characterized as populations or communities of organisms enclosed in a matrix and adhering to interfaces or each other. In multicellular organisms such as humans, animals, and plants, various microorganisms exist in biofilms or biofilm-like structures, playing essential roles in development, physiology, and immunity. However, some biofilms can have adverse effects on host health.

In summary, proteins are the fundamental building blocks of all organic systems, making them central to various natural sciences. Moreover, there is a growing interest in their applications in biotechnology, food processing, and medicine. Two hydrophobin proteins, HFBI and HFBII, produced by Trichoderma reesei, are particularly intriguing due to their potential industrial applications and suitability for creating

novel vesicles with functions such as drug delivery. The rest of this chapter provides a concise overview of hydrophobins, explicitly focusing on proteins in this thesis.

2.1.4 Hydrophobins: Proteins with the particular pattern in the primary structure
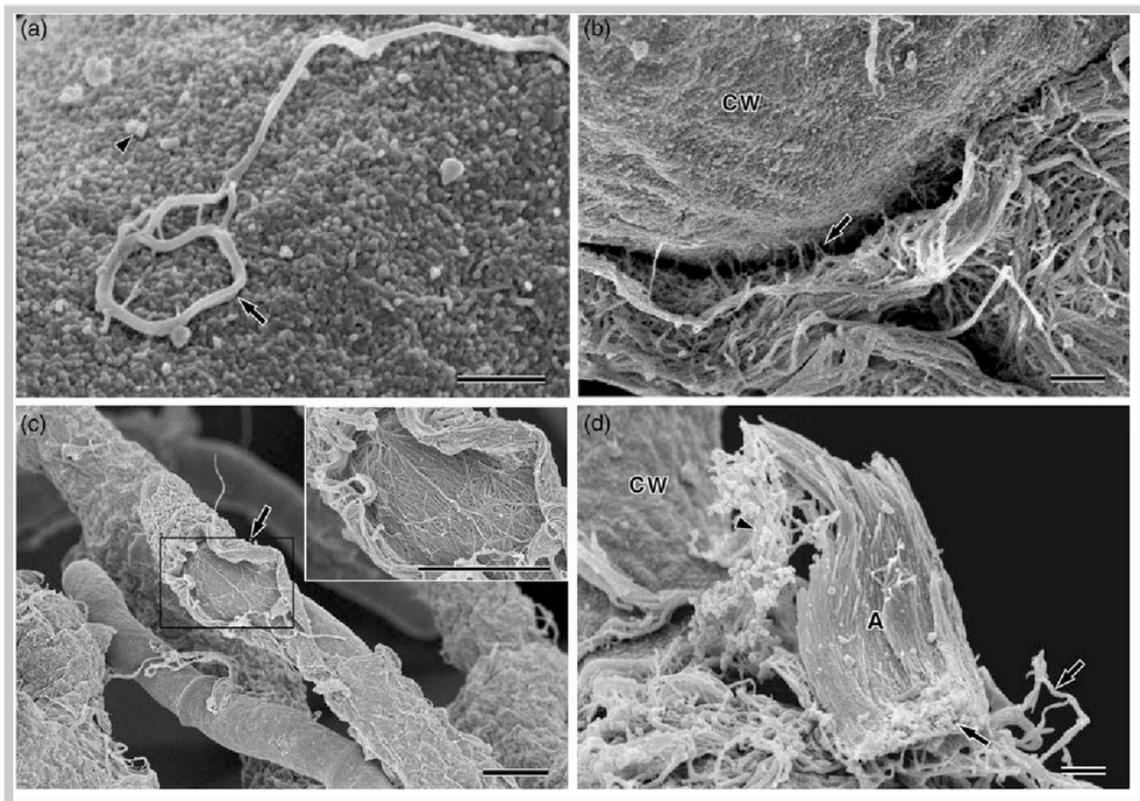
Figure 2.5: The generation of a fibrous material or fibrous layer is a distinctive characteristic of T. reesei. The surface structure of a T. reesei mutant grown on **(a)** glucose, and **(b-d)** Avicel for one **(a)** and three days **(b-d)**. **(a)** Initially, the tip of the fibrous material was emerging from the cell wall surface (arrowhead); subsequently, it enlarged to form fibrous material (arrow). **(b)** The fibrous material became entangled to form layers of threads. The interface between the cell wall and the reticulate fibrous layer is shown by the arrow. **(c)** The layers do not adhere tightly to the cell wall. **(d)** Avicel is entangled within fibrous material (arrowhead) or becomes trapped in the fibrous layers. The abbreviations **A** and **CW** are used for Avicel and cell wall, respectively. The figure is reprinted from [48] with permission from **Journal of Electron microscopy**, Copyright (2012).

This thesis concerns a family of small globular proteins called hydrophobins produced by filamentous fungi. Hydrophobins have several biological roles in fungal

growth in surviving and adapting them to the environment. The proteins regarded in this thesis are HFBI and HFBII, originating from Trichoderma reesei.

The filamentous fungus T. reesei is a fast-growing Ascomycete found widely in soil environments. T. reesei secretes high amounts of cellulolytic enzymes [52]. These microbial cellulases find significant industrial applications in converting cellulose into glucose [32], making this fungus the focus of intensive study. Figure 2.5 illustrates the transformation of the fibrous material into a fibrous layer on the hyphal surface of a hyper-secreting T. reesei [48].

Hydrophobins exhibit distinct hydrophilic and hydrophobic parts on their surfaces, making them prone to migrating to hydrophobic-hydrophilic interfaces, such as air-water or solid-water. By forming stable layers, they effectively alter the surface tension or wettability of these interfaces [1, 53, 65].

Hydrophobins are divided into classes I and II based on the sequence of amino acids in their primary structure [62], resulting in variations in solubility and aggregate patterns. Class I hydrophobins are characterized by the formation of rodlets, a distinctive feature absent in class II hydrophobins—the type to which the proteins discussed in this thesis belong. Adsorbed class II hydrophobins at air-water interfaces can form stable crystalline monolayers exhibiting high elasticity under specific experimental conditions. Moreover, these biofilms have the potential to create stable protein bilayers.

Hydrophobin bilayers have been successfully generated at various interfaces within microfluidic devices [23]. Microfluidic measurements indicate a bilayer adhesion energy on the order of several $k_B T$. While mutations do not impact the hydrophobic component, variations in the hydrophilic adhesion energy reveal that fewer charges result in higher adhesion energy, signifying repulsive Coulomb interactions between proteins [19]. Notably, hydrophobin adhesion energies tend to exceed those of other highly surface-active molecules, such as lipids [19].

The remarkable stability of hydrophobin bilayers positions them as ideal candidates for creating novel vesicles, particularly valuable in applications like drug delivery. In contrast to lipid vesicles, hydrophobin vesicles can be formed with either a hydrophobic or hydrophilic shell [19], rendering them versatile for use in diverse environments.

Moreover, hydrophobins HFBI and HFBII have demonstrated highly linear adsorption kinetics, as observed in [20]. This behavior is inconsistent with the typical Langmuir kinetics [1], suggesting the involvement of nontrivial mechanisms in the adsorp-

---

[1]Langmuir isotherm suggests an exponential adsorption rate, as will discuss in 2.2.6.

tion process. In-depth investigations into genetically modified variants of HFBI have revealed surprising results, with the adsorption kinetics for mutants aligning with the expected Langmuir type [20]. Therefore, a thorough examination of the structure and properties of these proteins becomes a crucial step in unraveling the precise mechanism behind the remarkable features of hydrophobins.

2.1.5 particular pattern in the primary structure

Hydrophobins are diverse in amino acid sequence. However, all hydrophobins share a distinguishing feature in their primary structure, i.e., the existence of eight cysteine residues that appeared in a conserved spacing pattern [62]. As shown in figure 2.6, the Cys-residues are regulated in two different arrangments, classifying hydrophobins into two subgroups, class I and II [62]. The disulfide bonds formed between Cys-residues result in stabilization and the spherical shape of hydrophobins [41].
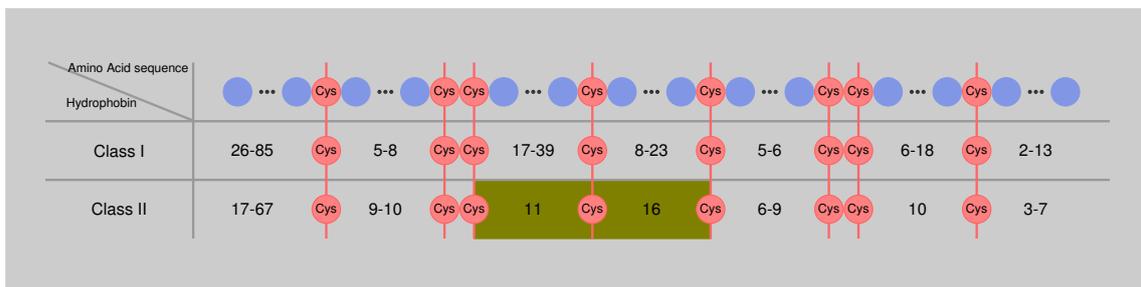
Figure 2.6: The sequence of amino acids in hydrophobins presents a characteristic feature. Eight Cys-residues exist in a particular design, classifying hydrophobins into two subgroups, class I and II. This arrangement reveals a more conversed way in class II proteins: the number of residues between Cys-residues 3 and 4 and Cys-residues 4 and 5 is entirely conserved, as shown in green in the panel. The numbers represent the variation of the number of amino acids in hydrophobins.

Class II hydrophobins, the focus of this thesis, exhibit a conserved sequence variation. Specifically, the lengths of polypeptide segments between Cys-residues 3 and 4 and between Cys-residues 4 and 5 in class II hydrophobins, highlighted in green in Figure 2.6, are entirely conserved. In contrast, class I hydrophobins display more diversity in these segment lengths.

This observed pattern may be linked to the various assemblies formed by hydrophobins, as illustrated in Figure 2.7. Experimental observations indicate that some class II hy-

drophobins, such as HFBI and HFBII, can form crystalline monolayer films at air-water interfaces [41, 57]. These soluble structures differ from the insoluble rodlet structures detected in class I hydrophobins [9].

While proteins of class II undergo no conformational changes at the interface, class I hydrophobins do. The rodlet structure represents the final stage of assembled aggregates, indicating that both class I and II hydrophobins form similar interface membranes. However, the membranes formed by class I hydrophobins become insoluble rodlet structures due to further conformational changes at the interface [15, 39].
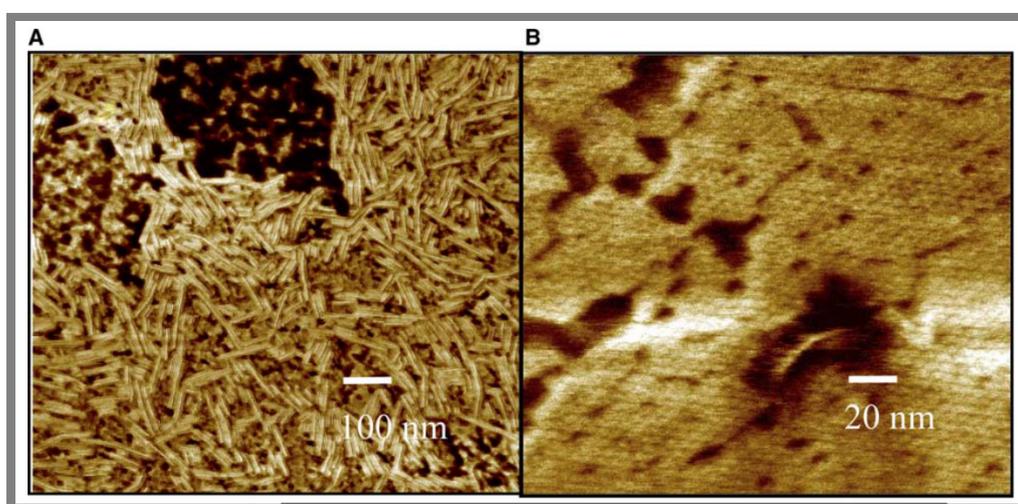
Figure 2.7: Atomic force microscopy images of hydrophobin film assembled at the air-water interface. **(a)** An AFM image of S. commune SC3 protein of class I hydrophobins shows rodlet structure. The sample was prepared by drying down a protein solution on a mica substrate. **(b)** An AFM image of T. reesei HFBI protein of class II presenting long-range ordered structure. The self-assembled film at the air-water interface was deposited on a mica substrate using the Langmuir-Blodgett technique. The figure is reprinted from [41] with permission from **FEMS Microbiology Reviews**, copyright (2005).

The following section describes the hydrophobins regarded in this thesis and some of their characteristic features.

2.1.6 Studied hydrophobins

The hydrophobins interested in this thesis comprise HFBI, HFBII, and several mutated variants of HFBI. HFBI and HFBII, wild-type proteins produced by T. reesei, exhibit a similar overall structure. The subsequent section offers a brief overview of the general aspects of their structure, focusing on features pertinent to the objectives of this thesis.

HFBII
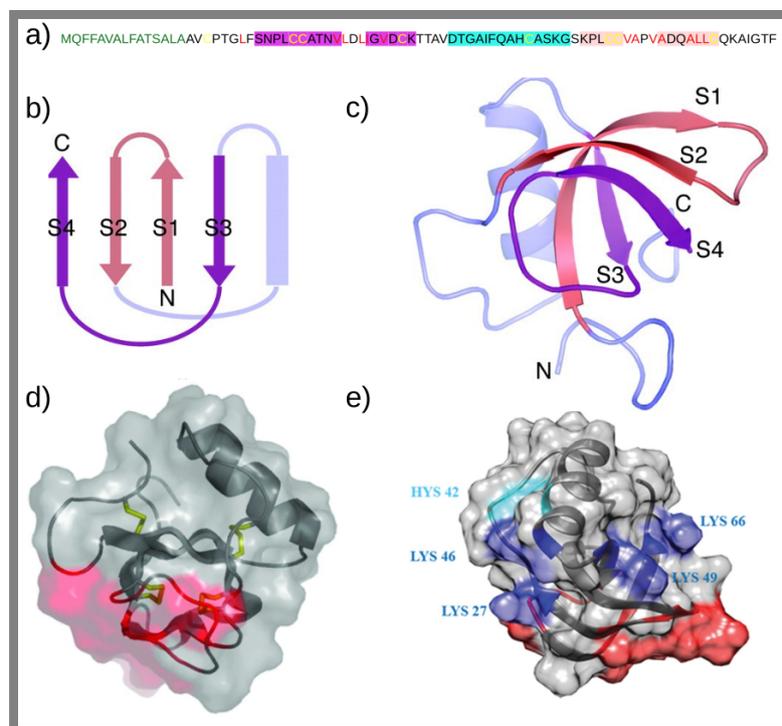
Figure 2.8: HFBII **(a)** The sequence of HFBII, with yellow color indicating the Cys residues and red for the hydrophobic patch residues. The residues highlighted in purple, cyan, and pink present the β-sheets number four and three, α-helix, and β-sheets number two and one, respectively. **(b)** Topology and structure of HFBII. The HFBII protein comprises four β-sheets and an α-helix. The β-sheets are shown in purple and pink arrows, and the blue rectangle indicates the α-helix structure. The N- and C-terminus are shown. **(c)** The three-dimensional structure of the HFBII. **(d)** The space-filling model of the HFBII structure, the disulfide bridges, and the hydrophobic patch are highlighted in yellow and red, respectively. **(e)** The structure of HFBII in which the positively charged residues are highlighted in blue (Lys) and cyan (His). The figures **(b-c)** are reprinted from [22] under the **CC BY Licence**, Elsevier. The figures **(a and d)** are reprinted from [21] with permission from **Biological Crystalography**, and the figure **(e)** is adapted from [14] under the **CC BY-NC-ND License**.

HFBII is a single-domain protein of 71 amino acids in the primary structure [21, 22]. The structure of the HFBII protein consists of four antiparallel β-sheets and an α-helix, as shown in the figure 2.8. The eight Cys-residues form four disulfide bridges symmetrically located in the same plane. These disulfide bridges make a network spanning the entire molecule and stabilize the protein's compact structure. The aliphatic side-chain groups, all located in the β- sheets except one positioned in the N-terminal loop, construct a relatively flat hydrophobic patch of the protein [22]. The 3D structure of the protein has a dimension of $24 \times 27 \times 30$ Å [22]. The total hydrophobic area of the protein is about $4 \ nm^2$, about 12 percentage of the total protein's area [14]. The remaining part of the surface is more hydrophilic. The protein has a molecular mass of $7.2 \ kDa$.

The HFBII includes five positive and four negative amino acids, Fig. 2.8-e. The negatively charged side chains are next to the hydrophobic patch. Below it is a half-ring composed of three positively charged amino acids, and the last two positive charges are on the hydrophilic side of the surface [14].

The HFBII is a rigid globular protein molecule in which the essential conformational changes at the air-water interfaces are not expected [39]. The thickness of a monolayer protein examined by AFM images is comparable to the size of the 3D structure determined by the X-ray structural analysis of crystal [53].

HFBI

The HFBI protein shares a similar structure to HFBII, consisting of four β-sheets, an α-helix, and the same disulfide bridges [22]. It differs from HFBII in the number of charged residues [22, 38].

The structure of HFBI has six charged residues exposed on its surface, as shown in figure 2.9. Four amino acids, Asp40, Asp43, Arg45, and Lys50, are positioned at the hydrophilic side of the protein opposite the hydrophobic patch. The remaining two charged residues, Lys32 and Asp30, are located in proximity to the hydrophobic patch. HFBI has a size of 70 amino acids and a cross-section of about $2 \ nm$ in diameter [22]. Both wild-type proteins show crystalline film structure, as discussed in 2.1.8.
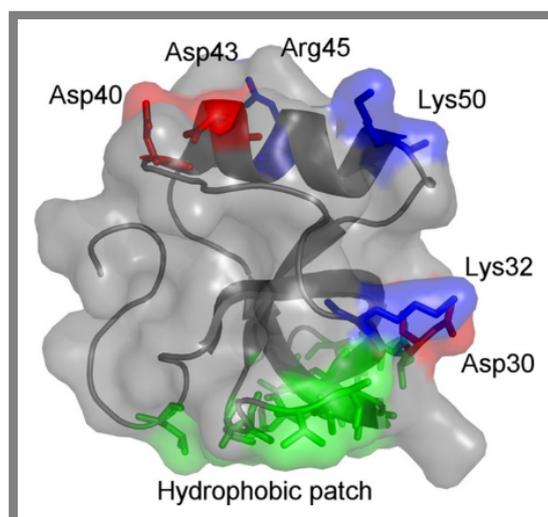
Figure 2.9: (a) Three-dimensional structure of T. reesei hydrophobin HFBI: The hydrophobic patch colored green comprises nearly all hydrophobic aliphatic side chains of protein. The remaining surface is hydrophilic—the protein self-assembles to the hydrophobic-hydrophilic interfaces due to its amphiphilic nature. The structure of HFBI consists of six charged amino acids exposed on the surface, labeled red and blue for negatively and positively charged amino acids, respectively. The figure is reprinted from [38] with permission from **American Chemical Society**, Biomacromolecules, Copyright (2015).

The crystal structure observations indicate lateral interactions between proteins at the interface. Those interactions are supposed to play a role in high buckling strength and exceptionally high shear moduli of hydrophobin membrane [65]. The mechanism and type of these interactions are unclear; hence, selected mutations probing the lateral interactions can shed light on it.

The role of lateral interactions can be investigated in two ways: (i) replacing the two charged amino acids proximal to the hydrophobic patch of HFBI with neutral amino acids, or (ii) glycosylation or adding cellulose domains to the HFBI. The larger size in the latter prevents proteins from getting close together, and gaining weight suppresses the protein's diffusive motion at the interface. In the following, I describe the structure of a few variants suitable for studying the role of lateral interactions.

HFBI-D30N-K32Q

HFBI-D30N-K32Q is a mutant of HFBI in which two charged amino acids, Asp30 and Lys32, close to the hydrophobic patch, are replaced by electrically neutral residues, Asn and Gln, respectively. This mutation does not affect the intermolecular interactions of hydrophobin, indicating that both proteins, HFBI-D30N-K32Q and HFBI,

maintain a similar conformation and orientation at the air-water interface [43], as confirmed by AFM examination of films from both proteins [38].

The AFM images of Langmuir-Blodgett films of HFBI-D30N-K32Q revealed a stable raft-like structure with confirmed crystallinity [38]. However, altering these residues influences the intramolecular interactions, as evidenced by measurements of the loss modulus. The loss modulus of HFBI-D30N-K32Q increased rapidly compared to HFBI [38].

FpHYD5
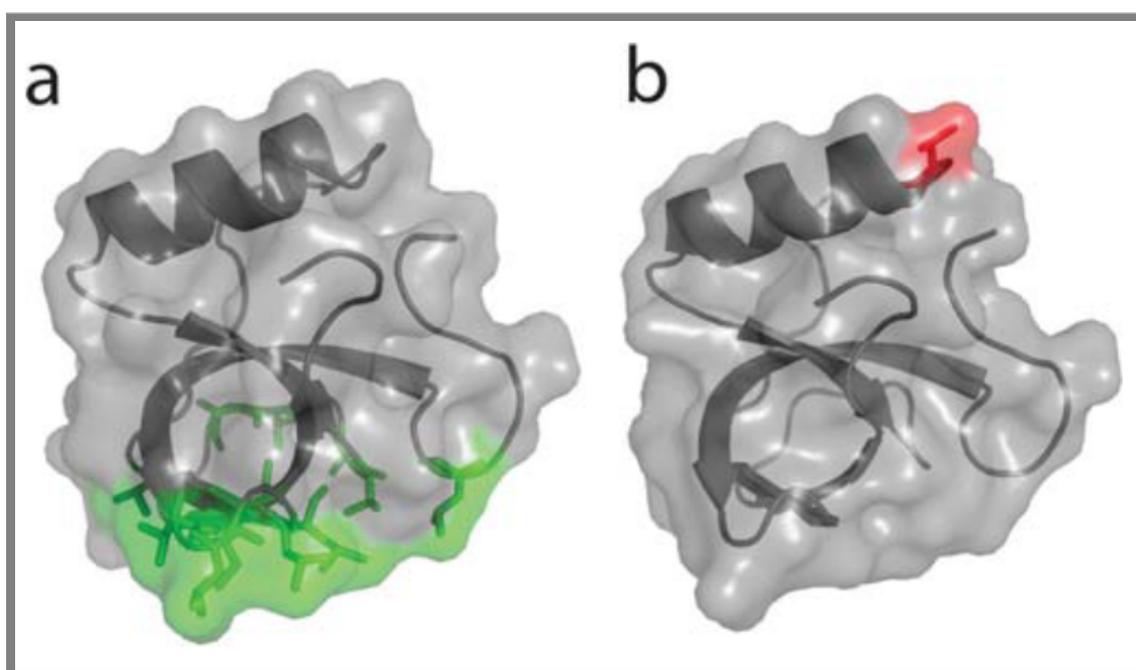
Figure 2.10: FpHYD5 is a glycosylated hydrophobin stemming from F. poae. The glycosylated site is in a surface-exposed opposite on the hydrophilic part of the surface. **(a)** The homology structure of HFBII with the hydrophobic patch colored green. **(b)** The homology structure of FpHYD5 compared to the HFBII. The glycosylated site is colored red. The figure is reprinted from [51] with permission from **Journal of Basic Microbiology**, copyright (2012).

FpHYD5 is a glycosylated hydrophobin stemming from Fusarium poae [51]. The glycosylation site is positioned on the hydrophilic side of the protein opposite the hydrophobic patch, as shown in figure 2.10. Its total molecular weight is $9.2\ kDa$, from which the isolated protein contributes $7.5\ kDa$, comparable to the HFBI molecular weight, and the leaving mass of $1.7\ kDa$ contributes from the glycan structure attached to protein [51].

HFBI-dCBD
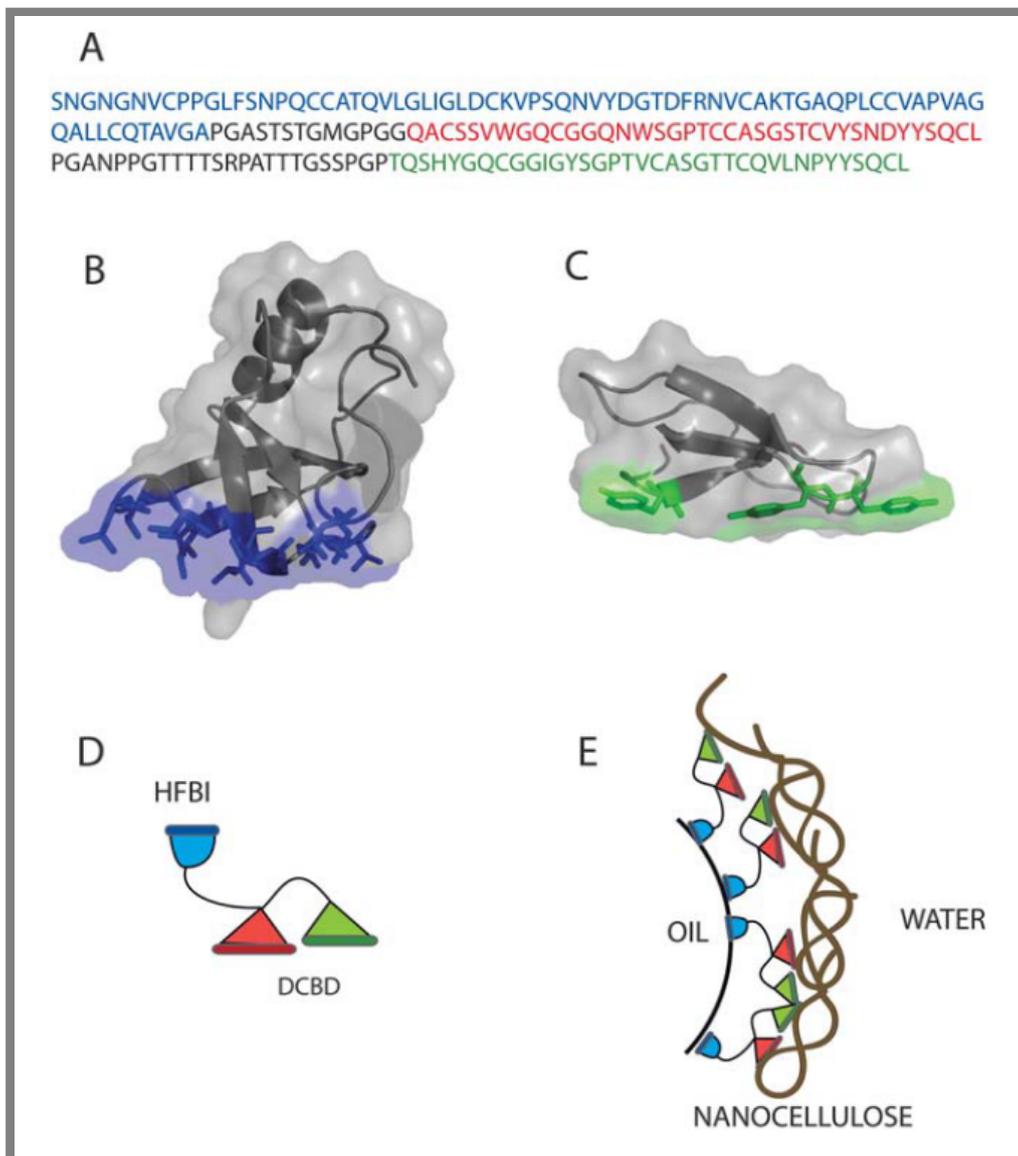
Figure 2.11: HFBI-dCBD **(a)** The amino acid sequence of HFBI-dCBD; The first blue script presents HFBI, followed by the cellulose-binding sequences of Cel6A-CBD in red and Cel7A-CBD in green. Two domains are separated by flexible linkers, with the sequence shown in black. **(b)** The three-dimensional structure of HFBI. The hydrophobic patch is shown in blue. **(c)** The structure of Cel6A-CBD. The aromatic side chains that bind the cellulose are marked in green. Both CBDs have a similar structure. **(d)** A cartoon representation of the HFBI-dCBD protein. The main protein, HFBI, is shown in blue, while the two cellulose domains and linkers are depicted in red, green, and black lines, respectively. **(e)** Schematic illustration of how the fusion protein can lead to an assembly at a hydrophobic-water interface. The figure is reprinted from [61] with permission from **The Royal Society of Chemistry**, copyright (2011).

HFBI-dCBD is a hydrophobin fusion protein that incorporates two cellulose-binding domains (CBD), Cel6A-CBD and Cel7A-CBD, linked to the C-terminus of the HFBI protein. Schematically depicted in Figure 2.11, these cellulose-binding sequences are combined to form a construct known as double cellulose-binding domains (dCBD). The CBDs and the linker collectively have a mass of 11 kDa, contributing to a total molecular weight of 18.5 kDa for HFBI-dCBD [40]. The flexibility of the entire fused protein architecture is ensured by the sufficiently long length of the linker [61].

Table 2.1.6 summarizes protein data and some of their properties. In the following sections, I will delve into the essential features of these proteins, providing insights for understanding and modeling the adsorption kinetics.

protein	weight [kDa]	electric charge [e]	remarks
HFBI	7.5	0.6	wild type
HFBII	7.2	1.8	wild type
FpHYD5	9.2	-1.6	glycosylated hydrophobin stemming from Fusarium graminearum
HFBI-dCBM	18.5	0.6	two cellulose binding domains connected via linker to HFBI
HFBI-D30N-K32Q	7.5	0.4	two charged amino acids were substituted with non-charged residues

Table 2.1: Weight, electric charge and type of the used proteins

2.1.7 Solution oligomerization

At low concentrations, hydrophobins dissolve into monomers that self-aggregate at the hydrophobic-hydrophilic interface, effectively shielding their hydrophobic patch. However, monomers aggregate in solution as concentrations increase, forming various oligomers. This oligomerization process serves as a mechanism to cover the hydrophobic patch of hydrophobins. Factors such as the size of the hydrophobic patch relative to the hydrophilic part, concentration, and even mutations can influence the shape and size of the resulting oligomers.

The critical micellar concentration (CMC) is the specific concentration at which aggregates begin to form. The CMC depends on experimental conditions, including temperature, pressure, and the concentration of other surface-active materials in the solution. Beyond the CMC concentration, HFBI and HFBII proteins form dimers and likely tetramers at higher concentrations [56, 58, 67], as illustrated in Figure 2.12.
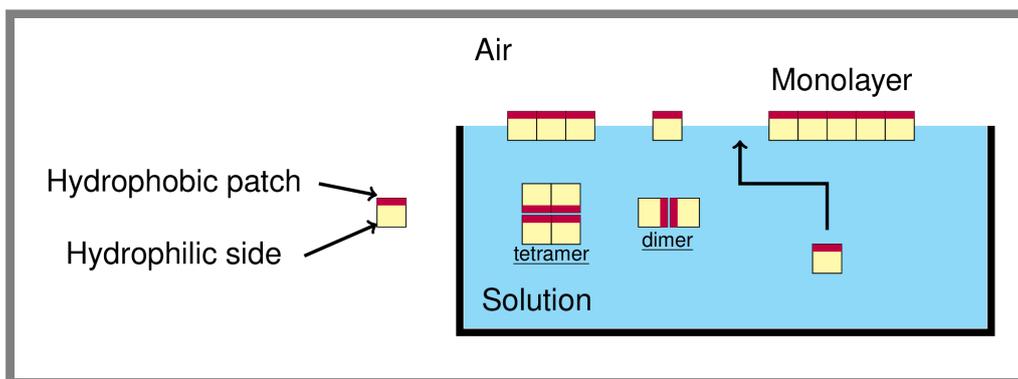
Figure 2.12: A possible model for forming different oligomers in solution: at concentrations less than CMC, proteins dissolve into monomers, which self-assemble to the interface to shield the hydrophobic patch. At higher concentrations, dimers and tetramers are likely. The figure is reprinted from **Hähl *et al.* [20]**, with the permission of **American Chemical Society**

The charged residue mutations affect oligomerization ability at a certain concentration, even though the explicit behavior of changes is unclear. However, oligomerization and adhesion to surfaces are correlated [37]. For instance, the mutant HFBI-D30N-K32Q shows higher stability of multimers and less binding capacity to surfaces than the WT protein HFBI.

2.1.8 Structure, crystal formation, film formation
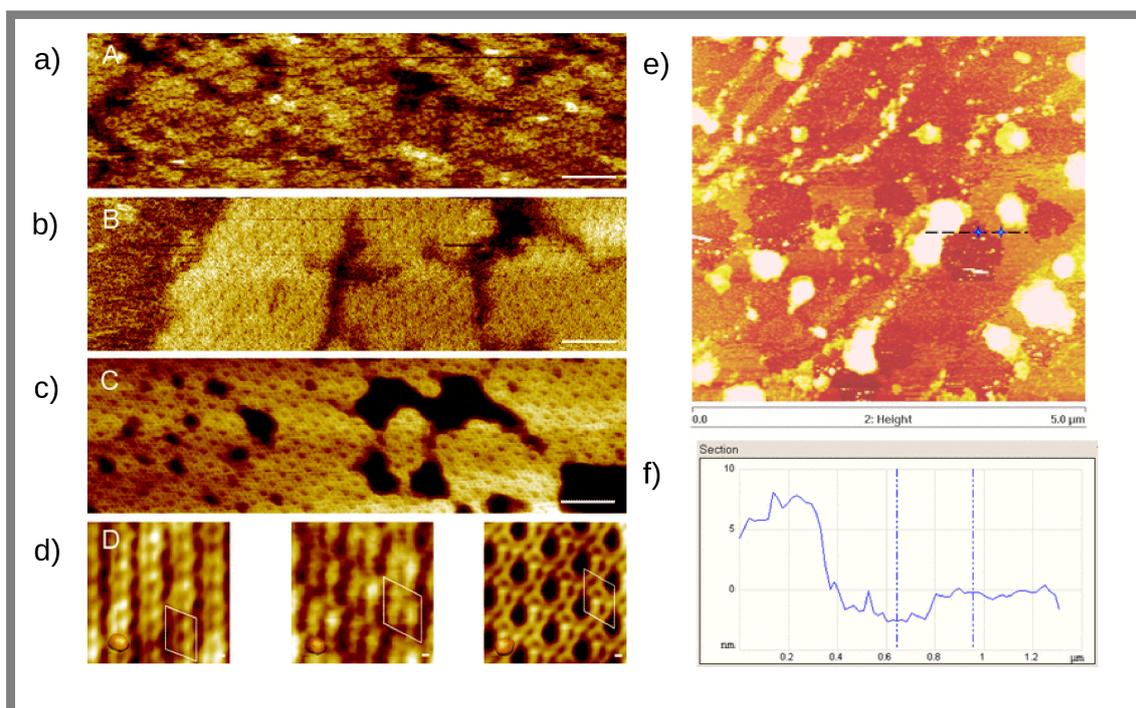
Figure 2.13: AFM examination of the HFBI films: a) The examined film is formed using the pendant drop and transferred to oriented pyrolytic graphite, as described in [57]. **b-c)** Both examined films were formed by spreading protein on a subphase using a Langmuir through, as explained in [57]; the film in **(b)** is transferred to oriented pyrolytic graphite, while in **(c)**, Langmuir Blodgett method is used to place the film on a piece of mica. **d)** The low-resolution surface structure of the HFBI monomer was produced using the UCSF Chimera package. **e)** AFM images of the HFBII layer at surface pressure of $\Pi = 25 \frac{mN}{m}$; The sample has been prepared based on the protocol discussed in [53]. The surface is highly inhomogeneous. The bright spots represent thicker layers of three to five-layer proteins. The darker spots show the voids. **f)** Thickness of the HFBII layer vs. the distance along the section of the dashed line in the figure **(e)**. The zero on the vertical axis is chosen arbitrarily. The height difference between each of the two points is independent of the position of the zero. The figures **(a-d)** are reprinted from [57] with permission from **American Chemical Society**, (copyright 2007). The figures **(e-f) are reprinted from** [53] with permission of the **American Chemical Society**, copyright (2013).

Hydrophobins adsorb at hydrophobic-hydrophilic interfaces and form films. The capability to adhere to different surfaces is an essential function of hydrophobins. However, the developed membranes show diverse characteristic dependence on the subdivisions of hydrophobins. The class I hydrophobins adhere to the surface very firmly and form insoluble film, while the class II members include more dissociable films. In

the formation of a film, the saturated time is defined when the highest value of adsorbed mass, saturated adsorbed amount, is achieved at the interface. The adsorption process is irreversible for hydrophobins [1, 53]; hence, the adsorption process stops at saturation, and the saturated amount stays constant later. The saturated adsorbed amount is independent of the adsorption process, and its surface density is comparable to the formation of a monolayer [1, 20, 53].

Different techniques, such as adsorption from a bulk solution or adsorption by spreading protein solution on the interface form a film. Those films consist of voids and domains of proteins [53, 57]. The class II hydrophobins interested in this thesis form mesoscopic interfacial layers despite the formation technique [53, 57, 65].

The adsorption from the bulk solution leads to a monolayer film, as shown in figure 2.13 a [20, 57]. The films result in the spreading of protein, dependence on the experimental conditions such as protein solution and compression, are inhomogeneous layers of voids, monolayers, multilayers of an odd number, and even spots of aggregation of proteins as shown in figure 2.13-e [57]. The bright spots in Fig. 2.13-e are likely the protein aggregates deposited in the spreading drop solution of protein.
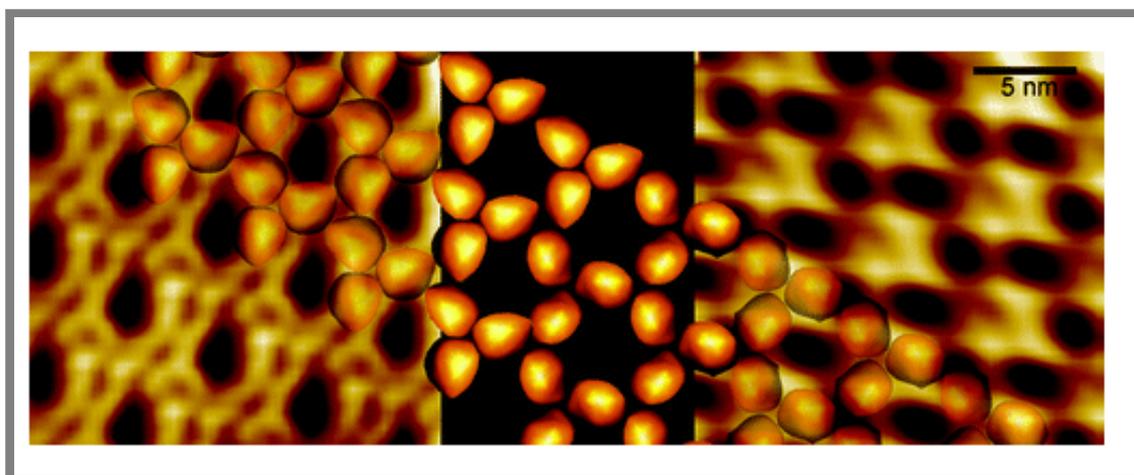
Figure 2.14: The left panel shows HFBI film sampled based on Langmuir Blodgett's technique. The right panel presents the HFBI-CysC-biotin film prepared using the Langmuir Schaefer method. The hexagonal structure is visible in both films. The middle panel suggests a possible arrangement that matches both films. The figure is reprinted from [57] with permission from **American Chemical Society**, copyright (2007).

The interfacial film of HFBI shows a repeating pattern close to the hexagonal lattice [57, 65]. Figure 2.14 illustrates a model suggesting two arrangements of protein leading to a hexagonal structure: (i) a large ring consisting of six proteins or (ii) an arrangement of units of three proteins, trimers, in each intersection of hexagons [57].

As previously noted, observations of the crystalline structure indicate lateral interactions between proteins at the interface. The role of these lateral interactions is examined by proposing various interaction mechanisms in chapters 5 and 6 of this thesis.

2.1.9 summary

This section has highlighted several remarkable features of hydrophobins, capturing substantial interest in various industrial applications. However, to harness the properties of hydrophobins most effectively, a comprehensive understanding of their functionality and the ability to control their interactions are crucial. These questions are explored in two chapters of this thesis.

In Chapter 5, I present experimental results and address open questions concerning adsorption kinetics. Subsequently, I introduce a model that describes the mechanisms underlying the adsorption process. Finally, in Chapter 6, I explore the interactions that contribute to the high elasticity observed in hydrophobin biofilms. Before delving into these chapters, briefly review the theoretical foundations that will be utilized.

2.2 Theoretical Background

This thesis aims to elucidate the intricate dynamics governing hydrophobin adsorption behaviour at air-water interfaces. Hydrophobins tend to self-assemble at air-water interfaces and form multimer complexes or monolayers based on the experimental conditions. Notably, their self-assembly shields hydrophobic regions, facilitating their interaction with the environment. As I will mention in chapter 5, experimental observations have revealed a distinctive feature: the linear kinetics characterizing hydrophobins' adsorption process until saturation, making them distinguishable from other proteins, which usually follow Langmuir adsorption kinetics. This thesis focuses on developing a simplified model system that explains the underlying mechanisms governing this behaviour. This model, constructed with a focus on fundamental processes, allows exploring their contributions to the observed linear kinetics. The model initially adopts a lattice gas framework in two dimensions, omitting intricate protein-protein and protein-surface interactions for computational tractability. The experimental results hint at the importance of long-range electrostatic interactions; hence, the model further incorporates the DLVO theory to describe particle interactions within a bilayer system. The Monte Carlo approach based on the Metropolis algorithm offers a practical way in which a few deterministic variables describe the system, and noise includes the randomness in the system. The following provides a review of concepts setting the theoretical basis of this thesis.

2.2.1 DLVO theory

Derjaguin, Landau, Verwey, and Overbeek developed a theory to explain the stability of colloidal systems. According to the theory, two forces impact the stability of colloidal systems: (i) the repulsion of electrostatic double-layer interactions and (ii) the van der Waals attraction.

The electrostatic double-layer interaction is calculated as pair interactions in an infinite electrolyte solution using the Poisson equation combined with the charge density distribution characterized in the Boltzmann distribution. The resulting equation is known as the Poisson-Boltzmann equation. The Poisson-Boltzmann equation, for a solution composed of N ions of the valency of z_i, can be formulated as

$$\nabla^2 \psi = -\frac{e}{\epsilon} \sum_{i=1}^{N} z_i n_i e^{-\frac{z_i e \psi}{k_B T}},$$

where ψ is the electrostatic potential, e is the elementary charge, n_i denotes the concentration of ions, and ϵ is the absolute permittivity of the solution. The Poisson-Boltzmann equation is a non-linear equation without any analytical solution. However, in the limit of $\frac{z_i e \psi}{k_B T} \ll 1$, i.e., Debye-Hückel approximation, it converts to the Helmholtz equation

$$\nabla^2 \psi = \kappa^2 \psi,$$

where $\kappa^{-1} = \sqrt{\frac{2e^2 I}{\epsilon k_B T}}$ is the Debye screening length and $I = \frac{1}{2}\sum_{i=1}^{N} z_i^2 n_i$ is the ionic strength of the solution. This equation can be solved analytically for various geometries and boundary conditions. A practical example is the geometry of two infinite parallel flat-charged plates immersed in an electrolyte solution. Derjaguin used this geometry to calculate the interaction between the two spheres as a sum of corresponding interactions between infinitesimal surface elements having planar geometry.

The interaction energy of two atoms is described by London interaction in the microscopic scale, in which the energy decays according to r^{-6}, where r denotes the radial distance between two atoms. The microscopic approach and the energy additivity principle result in an analytic formalism of interaction between complicated geometries of the interacting bodies, known as the van der Waals interaction on the macroscopic scale. Hence, the van der Waals interaction of two macroscopic bodies derived by Hamaker reads

$$U_{vdW} \propto \int\limits_{v_1} \int\limits_{v_2} \frac{\rho(r_1)\rho(r_2)\,dv_1\,dv_2}{|r_1 - r_2|^6}.$$

For instance, computing the integral for the two parallel semi-infinite plates at a distance r results in $\frac{A}{r^2}$, where A is known as the Hamaker constant. The Hamaker constant depends on the London constant and the charge of interacting bodies.

The DLVO theory combines the van der Waals interaction energy and the double-layer interaction energy to describe the interaction between two particles in a liquid. In most cases, the Hamaker constant is positive; hence, it has an attraction contribution with a negative sign in the total energy. On the other hand, the double-layer interaction can be either positive or negative depending on the charge of particles, surface potential, and the distance between them. Additionally, the range of the interaction can be affected remarkably by changing the ionic strength of the solution.

Thus, the superposition can result in a complicated energy profile determining if the colloidal system coagulates.

A typical energy profile for a system of charged particles is shown in figure 2.15. The energy minimum corresponds to the adsorption of the colloidal particles. The characteristic feature is the appearance of a maximum energy barrier after the primary minimum. The energy barrier is followed by a minimum called the secondary minimum for a system of low solution concentration and large particles. The secondary minimum is more shallow than the primary minimum; however, it is expected that particles will accumulate around its position, which can impact the adsorption.
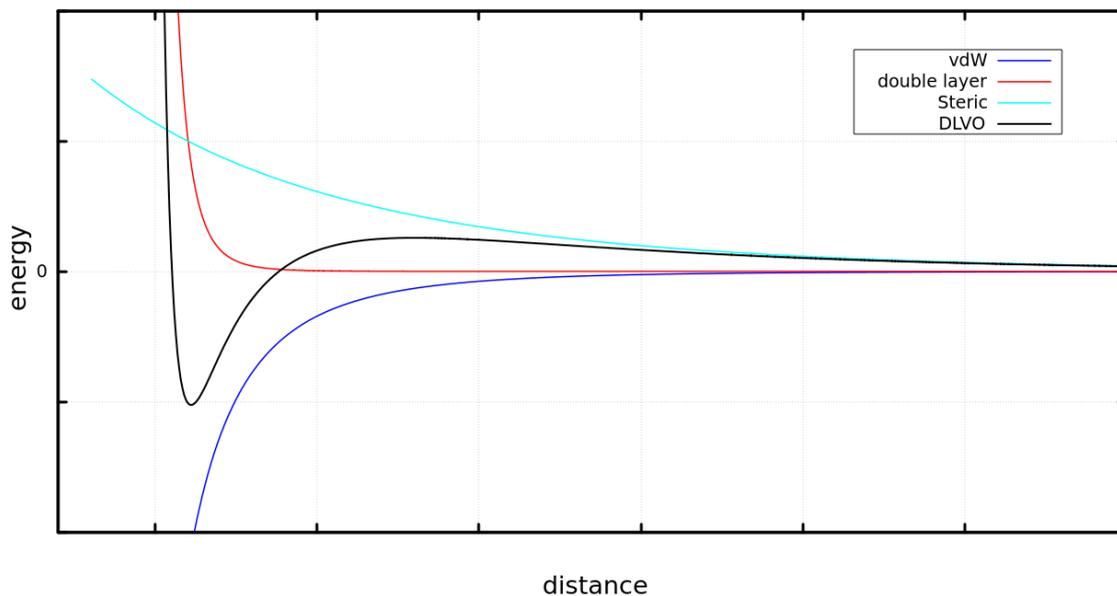
Figure 2.15: Gibbs free energy of pure van der Waals attraction, double layer repulsion, Steric repulsion, and DLVO.

2.2.2 Stochastic Processes

A stochastic variable, or equivalently a random number, is an object X defined by (i) a set of possible values x_i, and (ii) a probability distribution over this set, i.e., p_i. As the stochastic variable is defined, any other one can be derived from it, e.g., using a mathematical function such as $Y_X(t) = f(X,t)$. $Y_X(t)$ is generally a random variable; however, it is called a stochastic process if the variable t stands for time. The probability density that the random variable takes value x_i at time t_i is $p_i(x_i, t_i)$. A set of joint probabilities, defined as $p_k(x_1, t_1; x_2, t_2; \ldots; x_k, t_k)$, describes the system completely. Under certain conditional probability, called the transition probability

$p_{1|1}(x_1, t_1 | x_2, t_2)$, the stochastic variable evolves from (x_1, t_1) to (x_2, t_2). Thus, a complete description of the system is obtained in terms of the joint conditional probabilities and the joint probability density using the following equation

$$p_{k+l}(x_1, t_1; \ldots; x_{k+l}, t_{k+l}) =$$
$$p_k(x_1, t_1; \ldots; x_k, t_k) p_{k|l}(x_{k+1}, t_{k+1}; \ldots; x_{k+l}, t_{k+l} | x_1, t_1; \ldots; x_k, t_k), \quad (1.1)$$

where the joint probability density in the left-hand side, i.e., $p_{k+l}(x_1, t_1; \ldots; x_{k+l}, t_{k+l})$, determines the probability that the stochastic variable takes values $(x_{k+1}, \ldots, x_{k+l})$ at $(t_{k+1}, \ldots, t_{k+1})$ given that it had values $(x_1, x_2, \ldots, x_k)$ at $(t_1, t_2, \ldots, t_k)$. The above equation is valid independently of the ordering of the time $(t_1, t_2, \ldots, t_{k+l})$. In the case of the right sequence of time, i.e., $t_1 \leq t_2 \leq \cdots \leq t_{k+l}$, one obtains the time evolution of the system, meaning that having enough knowledge of the past is sufficient to predict the future of the process. In the most important class of the processes, i.e., Markov processes, the future outcome depends only on the current state. This property simplifies the above equation considerably. Markov processes are used to model systems involving stochastic decisions over time.

2.2.3 Markov Processes

A Markov process is a stochastic process defined by the Markov property, which states the transition probability of a given stochastic process is determined by the last current state; hence, the knowledge of the previous random variables does not affect the transition probability. For any set of successive times, i.e., $t_1 \leq t_2 \leq \cdots \leq t_k$, the Markov property is written as

$$p_{1|k}(x_k, t_k | x_{k-1}, t_{k-1}; \ldots; x_1, t_1) = p_{1|1}(x_k, t_k | x_{k-1}, t_{k-1}) \quad t_1 \leq t_2 \leq \cdots \leq t_k.$$

Thus, the entire hierarchy of the Markov process can be obtained from the initial condition, $p_1(x_1, t_1)$, and the transition probability, $p_{1|1}(x_{i+1}, t_{i+1} | x_i, t_i)$, by following a straightforward algorithm, e.g., as follows

$$p_3(x_3, t_3; x_2, t_2; x_1, t_1) = p_1(x_1, t_1) p_{1|1}(x_2, t_2 | x_1, t_1) p_{1|1}(x_3, t_3 | x_2, t_2) \quad (1.2),$$

which makes Markov processes manageable and applicable but not history-dependent. Equation (1.2) leads to the Chapman-Kolmogorov equation, i.e.,

$$p_3(x_3, t_3|x_1, t_1) = p_1(x_1, t_1) \int dx_2 \; p_{1|1}(x_2, t_2|x_1, t_1) p_{1|1}(x_3, t_3|x_2, t_2),$$

where the times satisfy the order. i.e., $t_1 \leq t_2 \leq t_3$.

Here, it is useful to introduce the transition rate, which is the probability transition per unit time

$$\omega_t(x_2|x_1) = \lim_{\Delta t \to 0} P_{1|1}(x_2, t + \Delta t|x_1, t).$$

A special case of Markov processes is homogeneous processes in which the transition probabilities are time-independent. A particular case of homogeneous Markov processes is stationary, which describes equilibrium fluctuations. A stochastic process reaches stationarity if it has the same statics at all times. For a stationary Markov process, the probability transition, $p_{1|1}(x_2, t_2|x_1, t_1)$, does not depend on the times but only on the time interval, i.e., $\Delta t = t_2 - t_1$. Hereon, for convenience, I consider the stationary Markovian processes, and I use a new notation for the probability transition;

$$P_\tau(x_2|x_1) := p_{1|1}(x_2, t_2|x_1, t_1),$$

where τ denotes the time interval, i.e., $\tau = t_2 - t_1$. From the transition rate and the Chapman-Kolmogorov equations, the differential form of the Chapman-Kolmogorov equation can be derived called the master equation. In the limit of $\tau' \to 0$, where $\tau' = t_3 - t_2$, the master equation reads

$$\frac{\partial P_\tau(x_3|x_1)}{\partial t} = \int \{\omega_\tau(x_3|x_2) P_\tau(x_2|x_1) - \omega_\tau(x_2|x_3) P_\tau(x_3|x_1)\} dx_2.$$

The master equation is interpreted as a conditional differential equation that connects conditional probabilities; suppose the conditional probability of being in state x_3 at time t_3, if the system has been in state x_1 at time t_1. The master equation considers all conditional probabilities of transitioning from any arbitrary state x_2 into x_3 and leaving from x_3 to any random state x_2.

From the master equation, the evolution of the single time probability, i.e., $p_i(x_i, t_i)$, can be derived. In the special case of stationary Markov process in which $\frac{\partial p(x,t)}{\partial t} = 0$,

it becomes

$$\omega(x'|x)p(x) = \omega(x|x')p(x'),$$

which is known as the detailed balance equation. Here, $P(x)$ is the probability density. The detailed balance condition states there is no probability flux between the two states in equilibrium. The sum of all transitions into any state x must be balanced by the sum of all transitions from x'.

Theoretically, the solutions of the master equation determine the time evolution of a stochastic process entirely. However, the master equation is rarely analytically solvable. Numerical methods, approximations, and sampling are techniques used to describe a stochastic process. In sampling techniques, drawing random numbers from a given distribution generates a realization of the desired probability density. The following section serves the sampling methods used in this thesis. To this end, I introduce an exact method to produce a Markov chain in continuous time, i.e., the stochastic simulation algorithm, which works based on inverse transform sampling. Then, the Metropolis-Hastings algorithm is introduced, which constructs a sequence of Markov processes from a given distribution known up to a normalization factor.

2.2.4 Inverse transform sampling and tower sampling

Let x be a probability random variable with a known cumulative distribution function $F(x)$, and let U be a random variable uniformly distributed in $[0, 1)$. We are interested in random variables x distributed according to $F(x)$. The inverse transform sampling states that the random variable $F^{-1}(Y)$ has distribution function $F(x)$, where $F^{-1}(Y)$ is the inverse of F defined as [10]

$$F^{-1}(Y) := \{x \in \mathbb{R} | \ F(x) = Y, \ 0 \leq Y < 1\}.$$

The statement covers discrete and continuous distribution functions. In the case of a continuous distribution, the method is reduced to finding the inverse function, which sometimes needs to be solved numerically. As an example, consider the sampling of an exponentially distributed random variable, $p(x)$,

$$p(x) \propto e^{-\omega x}, \quad x \geq 0.$$

The cumulative probability function, $F(x)$ is

$$F(x) = \omega \int_0^x e^{-\omega x'} dx' = 1 - e^{-\omega x}.$$

$F(x)$ is a continuous analytical function. Hence, the inverse function of $F(x)$,

$$F^{-1}(Y) = -\frac{1}{\omega} \log(1 - Y),$$

would sample from the desired distribution. Since $(1 - Y)$ and Y are both uniformly distributed in $[0, 1)$, it is convenient to use a simplified form of $F^{-1}(Y)$, i.e.,

$$F^{-1}(Y) = -\frac{1}{\omega} \log Y. \tag{2.1}$$

In a discrete distribution function, the algorithm is also called tower sampling algorithm, which allows generating desired random variables [31].

Consider N random variables each has the probability of p_i, where $i = 1, 2, \ldots, N$. Then, the cumulative distribution function of $F(x)$ is defined as

$$F(x) = \sum_{i=1}^{N} p_i$$

The problem is sampling integer values m such that the random variable $F^{-1}(P)$ fulfills

$$\sum_{i=1}^{m-1} p_i \leq F^{-1}(P) < \sum_{i=1}^{m} p_i.$$

Based on the implementation of the algorithm, tower sampling is only practical when the probabilities can be listed.

2.2.5 Stochastic simulation algorithm

The stochastic simulation algorithm generates a trajectory of states according to the correct probability distribution; hence, the probability of the generated samples is precisely the probability that would come out of the solution of the Master equation. Sampling is based on the fact that the density probability of entering the state X at

time t given that the state is X_0 at the current time of t_0 is an exponential function [17], i.e.,

$$P(X, t | X_0, t_0) = \omega(X | X_0) \exp\left(-\omega(X | X_0)(t - t_0)\right).$$

Three methods use the probability mentioned above to generate a single trajectory: (i) the Direct method, (ii) the First reaction method, and (iii) The next reaction method. These methods are described briefly in the following.

Direct method

In the direct method, all transition rates are updated, and a reaction is chosen based on the statistical weight of the reaction, i.e., $\frac{\omega_i}{\sum_i \omega_i}$. In the next step, the waiting time for the reaction is calculated according to eq. 2.1 [17].

First reaction method

The first reaction method calculates the waiting time for each reaction. Then, the reaction of the minimum waiting time is chosen [17].

The next reaction method

The next reaction method is an alternative to the first reaction method with the difference that the calculated times are stored and allowed to be reused. The algorithm samples a new waiting time for a reaction only if the transition rate changes. The next reaction method allows reusing generated random numbers, which is generally risky but, in the case of an exponential variable, is legitimate [16].

Metropolis-Hastings algorithm

The Metropolis-Hastings algorithm can generate a discrete Markov chain of states from a complex probability distribution $p(x)$ with an unknown normalization constant. An example of such distribution is the Boltzmann distribution, $\frac{\exp(-\beta H(x))}{K}$, where $H(x)$ denotes the energy of the state x and the normalization constant, K, is $\sum_x e^{-\beta H(x)}$. The resulting chain fulfills the detailed balance conditions. In the limit of sufficiently long enough steps of sampling, it approaches the stationary state, which samples from the desired distribution $p(x)$. The Metropolis-Hastings algorithm consists of

two steps: (i) given the current state of x', a new state x is proposed using an arbitrary transition probability function of $f(x|x')$, and (ii) the proposed state is accepted by the probability of $min(1, \frac{p(x)f(x|x')}{p(x')f(x'|x)})$.

Since the acceptance probability uses the relative probabilities of $p(x)$, the normalization constant cancels out. For a symmetric transition probability, i.e., $f(x|x') = f(x'|x)$, such as the Boltzmann distribution, the acceptance probability is simplified to $min(1, e^{-\beta(H(x)-H(x'))})$.

2.2.6 Langmuir Isotherm model

Langmuir adsorption model describes the adsorption of a gas at a solid surface in equilibrium. The substrate consists of identical distinct sites, each of which can be occupied by a single adsorbent. The model considers non-interacting gas molecules on the gas phase or the surface. Langmuir classified six different mechanisms of adsorption [33], from which I present the most straightforward case in one dimension. All the possible processes between the gas and the surface are (i) incident at the surface, (ii) reflection from the surface, (iii) adsorption, and (iv) desorption. By considering three fundamental assumptions, a simple relation between equilibrium site occupancy, θ, and the pressure of the gas phase can be derived: (i) The rate of incident particles is proportional to the gas pressure, P. The adsorption rate, r_{ads}, depends on, up to a constant, K_{ads}, the incident rate and the probability of incidence at a free site, i.e.,

$$r_{ads} \propto P\left(1-\theta\right) \quad \Rightarrow \quad r_{ads} = PK_{ads}\left(1-\theta\right).$$

(ii) The desorption rate, r_{des}, is proportional to the density of the occupied sites

$$r_{des} \propto \theta \quad \Rightarrow \quad r_{des} = K_{des}\theta.$$

(iii) in equilibrium, the adsorption rate is equivalent to the desorption rate, hence

$$r_{ads} = r_{des} \quad \Rightarrow \quad \theta = \frac{K_{eq}P}{1+K_{eq}P},$$

where K_{eq} is called the Langmuir constant, independent of pressure, and has a tem-

perature dependency. The Langmuir constant is determined experimentally.

The adsorption kinetics is given by

$$\frac{d\theta}{dt} = r_{ads} - r_{des},$$

where, r_{ads} and r_{des} are the adsorption/desorption rates, respectively. The desorption rate can be proportional to the occupied sites, $r_{des} = K_{des}\theta$, or a constant, $r_{des} = K_{des}$, depending on the particular method of desorption. For a system of a constant desorption rate,

$$\frac{d\theta}{dt} = PK_{ads}\left(1 - \theta\right) - K_{des}$$
$$\frac{d\theta}{dt} = -\left(PK_{ads}\right)\left(\theta + \left[\frac{K_{des} - PK_{ads}}{PK_{ads}}\right]\right),$$

which results in

$$\theta(t) = \theta_m\left(1 - e^{-\frac{t}{\tau}}\right),$$

where $\theta_m = \frac{PK_{ads} - K_{des}}{PK_{ads}}$ and $\tau = \left(PK_{ads}\right)^{-1}$. Hence, the Langmuir adsorption model results in an exponential adsorption rate; see figure 2.16.

The self-aggregation of proteins is anticipated to adhere to the Langmuir isotherm. Surprisingly, hydrophobins investigated in this thesis exhibit linear adsorption kinetics. Remarkably, the introduction of mutations converts the kinetics to the Langmuir type. This thesis delves into the underlying mechanism by introducing a stochastic model, as discussed in Chapter 5.

$$\theta(t) = \theta_M (1 - \exp(-t/\tau))$$
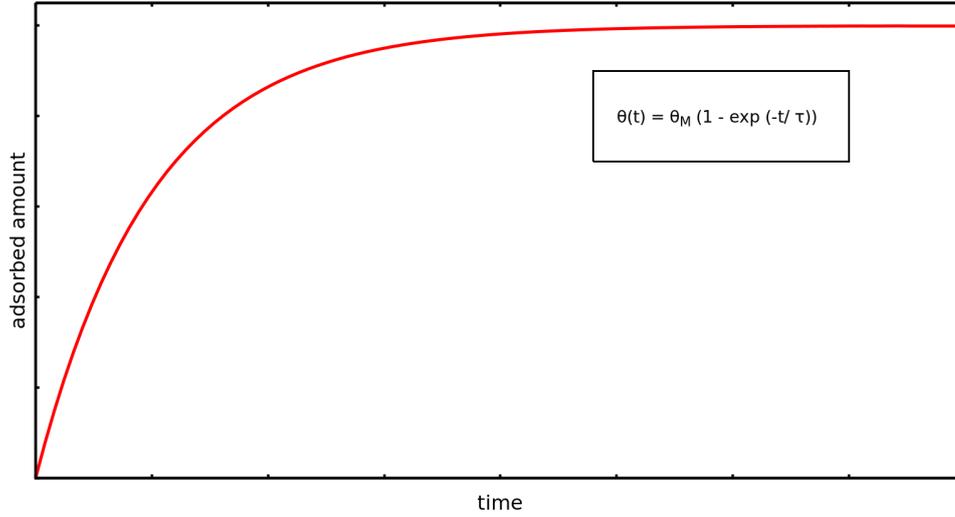
Figure 2.16: Langmuir adsorption kinetics.

2.2.7 Potts Model

The Potts model is a generalization of the Ising model, which describes a system of interacting spins on a lattice [11]. In this context, each lattice site possesses a spin property pointing to one of q equally spaced directions specified by the angles,

$$\theta_{i,n} = \frac{2\pi n}{q},$$

where i identifies a site on the lattice and n is an integer number, $n = 0, 1, \ldots, q$. While, in most systems, only nearest-neighbor interactions are considered, one can also conceive of a multi-site interacting system as interacting with an external field. For a Potts model on a lattice with N sites, the Hamiltonian, H, is generally given by [64]

$$-\beta H = \sum_{\langle i \rangle} L\delta_{Kr}(\sigma_i, \sigma_0) + K \sum_{\langle i,j \rangle} \delta_{Kr}(\sigma_i, \sigma_j) + M \sum_{\langle i,j,k \rangle} \delta_{Kr}(\sigma_i, \sigma_j, \sigma_k) + \ldots,$$

where $\beta = \frac{1}{k_B T}$, L denotes the external field, σ_i specifies the spin state at the ith site, and

$$\delta_{Kr}(\sigma_i, \ldots, \sigma_k) = \begin{cases} 1 & \sigma_i = \cdots = \sigma_k \\ 0 & otherwise, \end{cases}$$

and $\langle \dots \rangle$ specifies the range of interacting spins with $\langle i, j \rangle$ used for nearest neighbours.

The one-state Potts model addresses the percolation problem in two-dimensional systems, while the two-state Potts model corresponds to the Ising system. The limit of $q \to \infty$ reduces to the XY model, which characterizes a two-dimensional system with continuous phase transitions, as detailed in Chapter 3.

In this thesis, the Potts model is employed to simulate the formation of two-dimensional crystals in the film formation process of hydrophobins.

Chapter 3

Fundamentals of Phase Transitions

This chapter is an overview of phase transition. In the first section, two model systems are introduced. The second section presents the terminology of phase transition based on thermodynamics. The chapter continues with the fundamentals related to the phase transition from the statistical point of view. The following chapter is based on the presentations of Reichl [50], Huang [24], and Chaikin and Lubensky [6].

3.1 Model systems

In this section, I discuss the significance of two lattice models: the Ising and XY models. These models, with each lattice site carrying a classical quantity called spin, play a crucial role in our understanding of phase transitions. Within the realm of phase transition, both the Ising and XY models exhibit intriguing and non-trivial behavior. In the Ising model, for instance, we observe the spontaneous growth of the mean magnetization without an external field, a phenomenon that represents the transition. Similarly, in the XY model, the transition is driven by the unbinding of topological defects, a complex and fascinating process. I introduce these models briefly. Throughout this chapter, I will often refer to the Ising model as a primary example. Its concepts and principles serve as a solid foundation for understanding several other models and concepts in the field of phase transitions. The idea behind the XY transition, i.e., unbinding topological defects, is fundamental to understanding two-dimensional melting.

3.1.1 Ising model

The simplest lattice model applied extensively to different ordering phenomena is the Ising model. The Ising model is a particular case of the more general Potts model, in

which each lattice site can only take two distinct values, see 2.2.7. The Ising Hamiltonian is given by

$$H = -\sum_{(i,j)} J_{i,j} S_i S_j - h \sum_i S_i,$$

and $S_i = \pm 1$ is the spin value on the lattice site i. In the first sum, (i, j) denotes the nearest-neighbor pairs. Hence, the Ising model considers only nearest-neighbor interactions. $J_{i,j}$ is the interaction energy between nearest neighbor. For simplicity, the same interaction between all pair sites is considered, i.e., $J_{i,j} = J \quad \forall (i, j)$. The second term represents the interaction of the sites with an external field with the strength h.

The mean magnetization, $\langle M \rangle = \sum_i S_i$, quantifies ordering in a spin system, where the sum is taken over all sites.

The one-dimensional Ising system does not show any phase transition, but the two-dimensional system does. The two-dimensional Ising system exhibits the first-order transition in the presence of an external magnetic field. It undergoes a continuous phase transition at the Curie temperature T_C and below that in the absence of an external magnetic field.

3.1.2 XY model

The second spin system, which is used as an example in this chapter and will be discussed in detail in the next chapter, is the XY model. The two-dimensional XY model is a classical system of unit vectors S_i on each lattice site. Alternatively, these vectors are called spin. In contrast to the Ising system, spins can take any continuous value in the XY model. They are confined to rotate in the plane of the lattice. The Hamiltonian of the system, in the absence of an external magnetic field, is

$$H = -\sum_{(i,j)} J_{i,j} S_i \cdot S_j = \sum_{(i,j)} J_{i,j} \cos(\phi_i - \phi_j),$$

where $J_{i,j} > 0$ is the interaction energy between spins. Without losing generality, the same interaction energy between spins is considered, i.e., $J_{i,j} = J \quad \forall (i, j)$. The angle ϕ_i is the angle that i^{th} spin makes with arbitrarily fixed axes. The summation over (i, j) ensures the nearest neighbor interactions. Thus, the Hamiltonian is invariant

under the rotation of all spins, i.e., it possesses complete rotational symmetric. For simplicity, the lattice is considered a simple square lattice with the lattice constant a. The configuration in which all spins are aligned parallel is the ground state of the system.

The XY model does not show the phase transition associated with growing mean magnetization. In contrast, it undergoes a type of phase transition called the topological phase transition. Topological phase transition is the main subject of chapter 4.

3.2 Thermodynamic perspective

Thermodynamics studies many interacting particles in thermal equilibrium. The investigated systems are isolated, closed, or open, depending on the interaction with the surroundings[1]. In the thermal equilibrium state, all macroscopic properties are time-independent and defined by a few parameters called state variables.

State variables can be extensive or intensive. An extensive variable's value depends on the system's size, while an intensive variable is size-independent. The volume, V, and the number of particles, N, are examples of extensive variables, while temperature, T, and pressure, P, are intensive variables.

A collection of state variables with particular macroscopic properties, such as color, density, etc., defines a phase. For instance, consider the behavior of water as you change the temperature. Water can exist in three phases: solid, liquid, and gas. At low temperatures, it is solid and may form a crystal structure. Molecules are arranged highly packed in a regular pattern. From the microscopic point of view, molecules fluctuate around particular lattice sites. Increasing the temperature may make the transition to the liquid or gas phases possible. In the liquid phase, molecules are still highly packed but without any pattern. The molecules can move freely. In the gas phase, molecules move freely. Hence, the density is less than that of both liquid and solid states; thus, the gas phase has the most compressibility.

---

[1]An isolated system has no interaction with surroundings. A closed system is in contact with a thermal bath, and an open system is in mechanical equilibrium with its surroundings. So, the reservoir and the system can exchange energy and particles

Since only a few state variables are independent, an equation of state determines the state. For example, the ideal gas law, $\frac{PV}{T} = const.$, describes the state of a gas consisting of non-interacting localized particles. The constant in the equation depends on the number of particles.

The equation of state can be determined experimentally. In an experiment on a thermodynamic system, one changes a parameter and examines how another parameter responds to this change. The quantity, which measures how the system responds to the change, defines the response function. A response function is, in general, a function of state variables. Heat capacity and compressibility are two common examples of response functions.

Heat is the adsorbed energy that raises the temperature of a system while no work is done. The heat capacity, in general, defines the amount of heat needed to raise the system's temperature by a given amount. Usually, it is measured by fixing all independent variables except temperature. So, $C_X = \left(\frac{\Delta Q}{\Delta T}\right)_X$, where X denotes the particular fixed variable. The value of C_X depends on several variables. It depends on the nature of the substance, the amount of material, and the way of absorbing heat. Thus, it is an extensive variable. The heat capacity measured for a specific amount of matter is called specific heat and is an intensive variable.

The compressibility, κ_X, is a mechanical response function that, for instance, in a PVT system, defines how the volume changes as the pressure varies. For a fixed number of particles at a fixed temperature, it is called isothermal compressibility.

It is an experimental fact that both response functions introduced are positive quantities. However, I will show that being positive originates from the statistical behavior of the matter.

The internal energy of a system, U, is generally a function of state variables. If external forces work on the system, a thermal bath exchanges heat with the system, or a reservoir exchanges particles with it, the internal energy of the system changes. Then, the amount of the change in the internal energy becomes

$$\Delta U = \Delta Q + \Delta W + \mu \Delta N,$$

where ΔQ denotes the heat, ΔW the work done on the system, and ΔN the num-

ber of exchanged particles. μ is called the chemical potential and defines the energy added to the system by adding a particle.

In general, a pair of an extensive variable X and an intensive variable Y are conjugate variables[2] if a small increment in X at fixed Y results in small change in the internal energy, $dU = \pm Y dX$. Then, in general, $\Delta U = \Delta Q \pm Y \Delta X$.

Other thermodynamic functions related to internal energy are called potentials. A thermodynamic potential, e.g., Gibbs free energy, G, contains all thermodynamic information of the system. Thermodynamic potentials are Legendre transforms of internal energy. Thus, a change in them is independent of the path of change. The Helmholtz free energy, enthalpy, and grand potential function are other examples. The equilibrium state minimizes thermodynamic potentials. All thermodynamic information can be deduced from potentials, e.g., $P = -\left(\frac{\partial G}{\partial V}\right)_T$. One can find other equations in thermodynamic course books.

The Gibbs potential plays an essential role in the Ehrenfest classification of the phase transitions. The Gibbs free energy corresponds to the work performed by a system consisting of a fixed number of particles at pressure P in thermal equilibrium. The Gibbs potential, $G(P, T, N)$[3], is a function of pressure, temperature, and number of particles as

$$G = U - TS + PV = \sum_j \mu_j N_j.$$

Here, S is called the entropy of the system. Entropy measures the randomness at the molecular level of the system. μ_j and N_j are the chemical potentials and the number of particles for species j, respectively.

The Helmholtz free energy measures the work obtained from a closed system at a fixed volume and temperature. It is defined as

$$F = U - TS.$$

Enthalpy defines the potential of a closed thermally isolated system in mechanical

---

[2]Some of the conjugate pairs are $(P, V), (T, S)$ and, (μ, N).
[3]in a PVT system

contact with the surroundings. The entropy is constant for such systems and is given as

$$H = U + PV.$$

Finally, the grand potential is a practical function of an open system in a reservoir's thermal and mechanical equilibrium. It is a function of temperature, volume, and the chemical function defined as

$$A(T, V, \mu) = U - TS - \mu N = F - \mu N.$$

A thermodynamic system can exist in several phases, each exhibiting different macroscopic behavior. Varying variable states, such as external magnetic fields or temperature, may cause a transition between phases. The transition occurs at a particular temperature depending on the other state variables. The transition temperature usually is called the critical temperature, T_c.

Each realized particular state must minimize the corresponding potential for a given set of independent variables. It might happen that for a specific set of values, two or more phases satisfy the condition of minimizing the free energy, which causes the appearance of consistency. The Coexisting phases might exchange particles. Since the coexistence phases are thermally and mechanically in equilibrium, the Gibbs free energy must change continuously among these phases. However, there is no restriction on the Gibbs derivatives. How these derivatives behave at the transition point is the basis of the Ehrenfest classification of phase transitions.

In the Ehrenfest classification, the phase transition is classified based on the continuity derivatives of the Gibbs free energy as a function of other thermodynamic variables.

In a first-order transition, one discontinuity exists in the Gibbs potential's first derivative. Thus, for instance, the entropy $S = -\left(\frac{\partial G}{\partial T}\right)|_{P,N_j}$ changes abruptly at the transi-

tion point as

$$\Delta S = S^{II} - S^{I} = -\left(\frac{\partial G}{\partial T}\right)^{II}_{P,N_j} + \left(\frac{\partial G}{\partial T}\right)^{I}_{P,N_j} \neq 0.$$

Here, superscripts II and I denote phases on both sides of the transition. The discontinuity in the entropy leads to an infinite peak in the heat capacity[4]. A first-order transition always involves latent heat, which is defined by the changes in the enthalpy, $\Delta H = T\Delta S$ [5].

Similarly, a second-order transition is associated with a second-order discontinuity. Entropy varies continuously. Since the first derivative changes continuously, the heat capacity shows a finite peak. The behavior of the Gibbs potential and its derivatives during first-order and second-order transitions are shown schematically in figure 3.1.
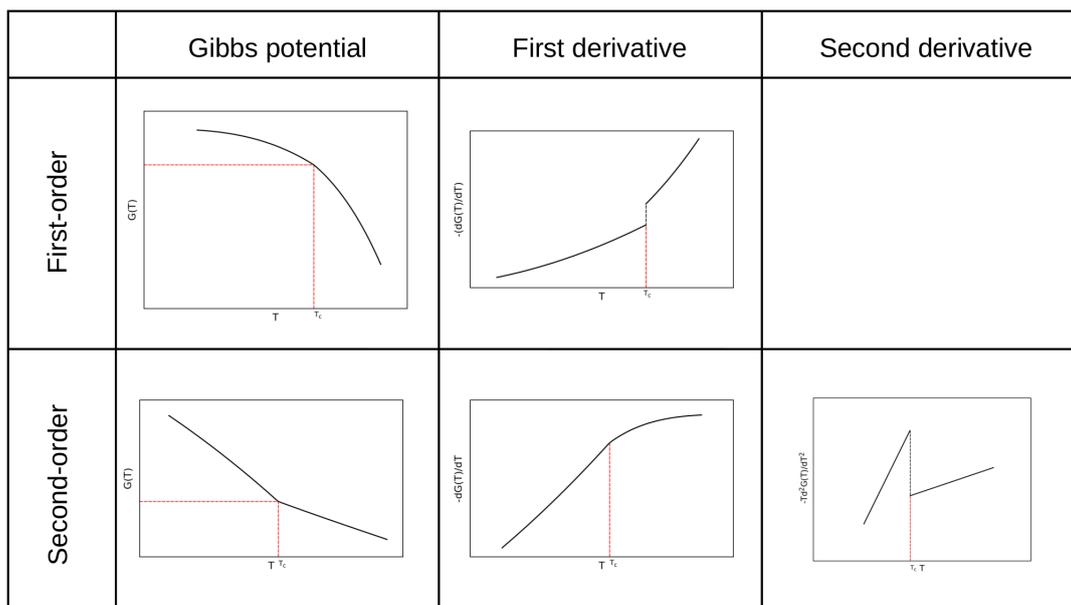
Figure 3.1: Phase transition based on the mathematical behavior of Gibbs free energy. The Gibbs potential is a continuous function of state variables, for instance, T, but its derivatives may display discontinuity. If the first derivative shows discontinuity, the transition is called the first-order transition. A second-order transition is associated with a discontinuity in the second-order derivative.

---

[4]For instance, $C_V = \left(\frac{dU}{dT}\right)_V = T\left(\frac{dS}{dT}\right)_V$ is not defined at the transition point.
[5]For a first-order transition, entropy does not change continuously, then enthalpy, $H = G - TS$, has not the same values during the transition.

3.3 Statistical mechanics perspective

Thermodynamics studies a system from an empirical viewpoint, while microscopic interactions of particles are not considered. However, statistical mechanics establish the connection. In statistical mechanics, the macroscopic state of the system is deduced from the dynamics of its constituent particles.

From the microscopic point, the coordinates and their conjugate momenta of individual particles, (q_i, p_i), uniquely determine the state of the system.

A $6N$-dimensional space formed by (q_i, p_i) is called the phase space. Each point in the phase space represents the state of the entire N-particle system. The Hamiltonian, H, and the initial state at t_0, (q_{i_0}, p_{i_0}), determine the system for entire time t. Hence, each point represents an instantaneous state, and a system moves through these points. However, a macroscopic observable is not an instantaneous value but an average measured in a time scale bigger than the collision times. A large number of these instantaneous microstates correspond to a macroscopic condition observed in thermodynamics, and one can not distinguish these microscopic points from the macroscopic point of view. In principle, one can follow the dynamic of the system for a given initial condition. However, for a system consisting of many interacting particles, instead of following the instantaneous states, a statistical description is essential.

Instead of considering the time evolution of a microscopic state, one can imagine the collection of time-independent microstates, all representing the same macroscopic state. The collection is called an ensemble, and the macroscopic condition defines the type of the ensemble. The distribution of points in the phase space represents the ensemble geometrically. Since the number of these systems is finite, the density of points in the phase space, ρ, is time-independent, $\frac{d\rho}{dt} = 0$.

In the phase space scheme, the system in equilibrium moves through the microstates. The time spent in a macroscopic state is proportional to the number of these states or the occupied volume in the phase space. In the ensemble scheme, all members can be attained equipossibly over time. This hypothesis is called ergodicity. Then, for an ergodic system, it is sufficient to look at the trajectory of an ensemble instead of look-

ing at the entire system over time[6]. Then, the mean value of a physical observable over time equals the average value obtained over the time-independent-ensemble collection

$$\lim_{T \to \infty} \left(\frac{1}{T} \int_0^T O(t)dt \right) = \lim_{M \to \infty} \left(\frac{1}{M} \sum_{i=1}^M p_i O_i \right),$$

where $O(t)$ is, in general, a time-dependent observable. M refers to the number of ensembles. P_i and O_i represent the probability of finding the system in state i and the mean value of observable in this state, respectively.

The volume occupied by an ensemble in the phase space is called the partition function. The partition function includes inter-particle interactions. On the other hand, as I state in the following, it connects to Helmholtz's free energy and entropy. Then, in principle, all thermodynamic information is obtained by having the partition function. Therefore, the partition function of a system plays an essential role in statistical mechanics.

Three kinds of ensembles correspond to different thermodynamical systems: (i) A microcanonical ensemble represents an isolated system with the energy between $[E, E + \Delta E]$. The entropy for such an ensemble referring to a system of N particles in volume V is defined by

$$S(E, V) = k_B \log \Gamma(E),$$

where $\Gamma(E)$ defines the partition function, and k_B is the Boltzmann constant. Solving this equation for E gives the internal energy of the system, $U(S, V) \equiv E(S, V)$. Having the internal energy is sufficient to deduce all thermodynamic information.

(ii) A canonical ensemble corresponds to a closed system in equilibrium with a thermal bath. While the temperature is constant, the energy fluctuates. The probability of finding a microstate i with the given energy E_i is proportional to the Boltzmann factor, $e^{-\frac{E_i}{k_B T}}$. Then, without losing generality, for non-degenerate discrete microstates,

---

[6]not all physical systems are ergodic

the partition function of a canonical ensemble is

$$Z_{Can}(N, V, T) = \sum_i e^{-\frac{E_i}{k_B T}},$$

where the sum is taken over all members of the ensemble. Thus, the probability of having a state i with the energy E_i is

$$p_i = Z_{Can}^{-1} e^{-\frac{E_i}{k_B T}}.$$

The canonical partition function is related to the Helmholtz free energy via

$$F(N, V, T) = -k_B T \log Z_{Can}(N, V, T).$$

Other thermodynamic information can be also obtained from the Helmholtz free energy. Additionally to the thermodynamic, the fluctuations of an observable can be obtained from the partition function. For instance, consider the global fluctuations in the energy,

$$< (\Delta E)^2 > = < E^2 > - < E >^2 .$$

Each term is related to the derivative of the partition function as

$$< E > = U = \sum_i E_i p_i = -\frac{\partial \ln Z_N}{\partial \beta},$$

$$< E^2 > = \sum_i (E_i)^2 p_i = \frac{\partial^2 Z_N}{(\partial \beta)^2}.$$

Then,

$$< (\Delta E)^2 > = \frac{\partial^2 \ln Z_N}{(\partial \beta)^2} = -\frac{\partial U}{\partial \beta} = k_B T^2 C_V,$$

where $\beta = \frac{1}{k_B T}$. Thus, the heat capacity, C_V, as an experimentally observable is positive since it is related to the energy fluctuations in the system.

In general, for a conjugate pair (X, Y), where Y is an intensive fixed variable and X

is a fluctuating extensive variable,

$$< X > = \pm \frac{\partial \ln Z_{Can}}{\partial(\beta Y)},$$

$$< (\Delta X)^2 > = < (X - < X >)^2 > = \frac{\partial < X >}{\partial(\beta Y)} = \frac{\partial^2 \ln Z_{Can}}{\partial(\beta Y)^2}.$$

(iii) In addition to microcanonical and canonical ensembles, there is another ensemble known as the grand canonical ensemble. The grand canonical ensemble refers to an open system in equilibrium with a reservoir. Due to this equilibration, both energy and the number of particles are allowed to vary. The probability of a given state with N_i particles at the energy E_i is

$$p_i = (Z_{Gr}(z, V, T))^{-1} \sum_{N_i=0}^{N} e^{-\beta(E_i - \mu N_i)},$$

where $z = e^{\mu}$ is fugacity, and the grand canonical partition function is defined as

$$Z_{Gr}(z, V, T) \equiv \sum_{N=0}^{N} z^N Z_{Can}^N(N, V, T).$$

Then, the grand potential function is defined as

$$A(T, V, z) = -k_B T \ln Z_{Gr}(z, V, T).$$

I showed that distinct ensembles correspond to different macroscopic systems. However, one can show that in the thermodynamical limit, all diverse ensembles give equivalent results.

In statistical mechanics, the thermodynamic limit refers to a large number of particles at a large volume proportional to the number of particles. In contrast, the particle density, $n = \lim_{(N,V) \to \infty} \frac{N}{V}$, is held fixed. At the thermodynamic limit, thermal fluctuations of quantities are negligible. For instance, I showed that the thermal fluctuations

in the internal energy are related to the heat capacity, then,$< \Delta E >\propto \sqrt{N}$. Hence,

$$\lim_{(N,V)\to\infty} \frac{< \Delta E >}{E} \approx \lim \frac{\sqrt{N}}{N} = \lim_{N\to\infty} \frac{1}{\sqrt{N}} \approx 0.$$

Therefore, in the thermodynamic limit, fluctuations in the total energy are negligible. Thus, for instance, in a gas system, the internal energy can be obtained by other thermodynamic variables such as temperature and pressure. So, any statistical system must be considered in the thermodynamic limit to serve as a macroscopic condition.

3.3.1 phase transition from statistical perspective

As mentioned before, phase transitions are associated with singularities in thermodynamic functions. From a statistical mechanics viewpoint, the partition function is sufficient to deduce all thermodynamic potentials. The partition function and its derivatives are analytical functions as long as the system is finite. However, in the thermodynamic limit, $\lim_{N,V\to\infty} \frac{N}{V} \to n$, singularities might appear in its derivatives. Yang and Lee studied the phase transition in the scheme of statistical mechanics. They indicated that for a hard-core monatomic gas with finite-range radial interactions, the equation of states follows [66][34]

$$\frac{P}{k_B T} = \lim_{N,V\to\infty} \frac{1}{V} \ln Z_{Gr}(z, V, T),$$
$$\frac{1}{V} = \lim_{N,V\to\infty} \frac{1}{V} z \frac{\partial}{\partial z} \ln Z_{Gr}(z, V, T).$$

It has been proved that the first equation is always an analytical monotonically increasing function. Also, the second equation is always a non-decreasing function. A phase transition occurs depending on the behavior of derivatives of $P(z)$ around the roots of the partition function.

The grand partition function does not have any real positive root for a finite system. However, some of these may approach the real axis in the thermodynamic limit. One can imagine a distribution of the roots in a complex space. A single phase corresponds to a region where no root coincides with the real axis. As a root approaches the real axis at point z_0, it splits the space into two subspaces, each representing a single phase. As mentioned above, $P(z)$ is always an analytical function, and $\frac{1}{V}$ is a

non-decreasing function. If $\frac{\partial P(z)}{\partial z}$ varies discontinuously at z_0, $\frac{1}{V}$ must have an upward jump to match on both sides of the root, which is the hallmark of a first-order transition. On the other hand, a continuous phase transition corresponds to a discontinuity in the higher-order derivatives of $P(z)$.

3.3.2 Symmetry and statistical observables

Phase transitions are often associated with symmetry breaking. For example, the average density in a crystal phase shows translational symmetry concerning the lattice unit vectors, so a solid in the crystal phase has discrete translational symmetry. It is also orientationally discrete symmetric. In a liquid, the average density is independent of the position. So, a liquid phase has a complete orientational and translational symmetry, i.e., melting transition breaks the discrete to complete symmetry in the liquid phase. As another example, applying a magnetic field in a spin system breaks the rotational symmetry. In the paramagnetic phase[7], the system is rotationally symmetric. The magnetic field induces spontaneous magnetization, defining a unique space direction. Here, complete rotational symmetry changes to discrete symmetry.

After this qualitative description of symmetry breaking, I explain it specifically by an instructive example regarding the spin system[8]. In general, the Hamiltonian determines the dynamical properties of any thermodynamic system. A group of operators under which the Hamiltonian remains invariant defines a symmetry group of the system. For instance, the Ising Hamiltonian is invariant under continuous rotations. This symmetry is broken during a transition.

Lowering temperature transforms the system to the ground state. However, sometimes, the ground state of the system is not unique. A physical system can be in one of these ground states, which has no symmetry similar to the Hamiltonian. Then, the transition spontaneously breaks the symmetry. In the Ising model, a configuration in which all spins are aligned in one direction defines the ground state. However, the energy of this system is invariant if all spins orient in the opposite direction. Therefore, although the ground state obeys the Hamiltonian, the symmetry of the ground state is different. Most of the time, one phase is less ordered than the other. For instance, at the high temperature of an Ising system, the paramagnetic phase in the

---

[7]In a spin system, a random distribution describes the alignment of spins.
[8]Ising system is an excellent example to clarify the concept of symmetry breaking. I will explain this model in detail in the next section.

absence of an external field is less ordered. The symmetry group of this phase is similar to the Hamiltonian. So, the paramagnetic phase of the Ising system has a complete orientational symmetry.

There are a set of thermodynamic functions, called order parameters, which are affected by these operators. These functions take zero at the high symmetric phase and nonvanishing values at the other. In general, an order parameter is the global average of an operator, $< O >$, where O is a function of dynamical variables and reflects the symmetry of the Hamiltonian. Hence, it takes zero in the disordered phase and a non-zero value in the ordered phase. Two or more phases may coexist on a coexistence line in a phase transition. The order parameter provides a proper measurement to distinguish them. For instance, the mean magnetization, $m = \frac{<M>}{N}$, is the order parameter of the Ising system. At the high-temperature randomly distributed paramagnetic phase, it vanishes. Still, it takes a finite value as an external magnetic field is applied[9]. This new phase, i.e., the ferromagnetic phase, is orientationally symmetric about the unique axis defined by the external field. Here, the higher-ordered phase has less symmetry. The ferromagnetic phase breaks the complete orientational symmetry of the Hamiltonian.

The correlation function relates to the order parameter and contains spectacular information about phase transition. To clarify, consider the alignment of spins in the Ising system without an external magnetic field. For a particular spin at time t, $S(x,t)$, there is a tendency for its neighbors to be aligned parallel to it. This tendency propagates over the entire system. The probability of finding another spin, $S(x + r, t)$, aligned parallel to $S(x, t)$ decreases as r increases. The two-point-spin correlation function measures this probability. This qualitative description suggests a correlation length, ξ, such that for $r \ll \xi$, the spin-spin correlation function is near one, and for $r \gg \xi$, it vanishes.

In general, for an order parameter M, one can define the density of the order parameter, $m(r)$, such that

$$M =< \int d^3 r m(r) >,$$

---

[9]Depending on the strength of the applied field, some spins tend to align in the direction of the external field. Then, the distribution of spins is not random anymore. This phase is called the ferromagnetic phase.

where the $<...>$ denotes the ensemble average. The correlation function measures how the density of the order parameter is spatially correlated, and it is defined as

$$C_{m,m}(r,r') =< m(r)m(r') > - < m(r) >< m(r') > .$$

In the next chapter, I will show that the mathematical behavior of the correlation function characterizes different phases in a two-dimensional melting transition.

Chapter 4

Two-dimensional melting

In the previous chapter, I discussed that a phase transition is usually associated with symmetry breaking. Different factors like thermal fluctuations, specific boundary conditions, or external fields can break a continuous symmetry. This symmetry breaking leads to the appearance of some discontinuities in the system. These discontinuities, called topological defects, result in a famous phase transition generalized by topological defects. Topological defects have different names depending on the broken symmetry and the studied system. However, they are characterized by two factors: a core and a far-field property. The core region, e.g., a point, is where the order is destroyed. Far from the core region, an elastic property changes slowly. A topological defect can be stable. Stable topological defects play an essential role in phase transitions.

In the previous chapter, order parameters and correlation functions are introduced. As I presented, in a two-dimensional spin system, the two-point correlation function in the ordered phase is constant within the limit of far separations. In contrast, it decays exponentially in the disordered phase. So, the former state has a long-range order, and the latter shows a short-range order. The correlation function may also behave as a power-law function with a temperature-dependent exponent, $C_{m,m}(r, r') \sim |r - r'|^{\eta(T)}$. In this case, the system has a quasi-long-range order. The transition from quasi-long-range order to the disordered phase is driven by topological defects. These transitions are known as topological phase transitions.

This chapter presents the topological phase transition in two-dimensional systems. I address this problem by introducing topological defects in two-dimensional systems, XY models, and two-dimensional crystals.

Before going to the main subject of this chapter, i.e., topological phase transitions, it is worth studying an analogous statistical model system, a two-dimensional lattice of charged particles. The method of deriving a phase transition in this model is generalized to the XY systems to understand how topological defects derive the phase transition.

4.1 A model system

As a model system, consider a two-dimensional square lattice occupied by a dilute gas of charged particles, as introduced by Kosterlitz and Thouless [29]. Each particle has a charge of $\pm q$, so the system is electrically neutral overall. Particles interact via a logarithmic potential [1] with a low-cutoff radius r_0 to avoid artificial divergences at small-scale separations. The cutoff radius, r_0, is of the order of the lattice constant. The Hamiltonian is

$$H(r_1, r_2, \ldots, r_n) = \begin{cases} \frac{1}{2}\sum_{i \neq j}\left(-2q_i q_j \ln\left(\left|\frac{r_i - r_j}{r_0}\right|\right)\right) + 2\mu & r > r_0 \\ 0 & r < r_0, \end{cases}$$

where q_i and r_i denote the charge and the position of the ith particle, respectively, and μ is the chemical potential to create a pair of equally opposite-charge at a distance r_0.

Although authors considered only a high value of chemical potential to construct a tractable theory, they emphasized that the argument is qualitatively applicable to a system with any value of μ. The result is generally unchanged if the low-cutoff radius increases to r_0' such that $e^{-2\beta\mu(r_0')} \ll 1$. A configuration of bounded particles as dipole pairs well separated from others minimizes the Hamiltonian such that the interaction between different dipoles is negligible. For the outlined system, the calculation of the mean square distance, $< r^2 >$, suggests a transition from a dielectric phase with bounded charges to a plasma phase with free charges, where the average, $< \cdots >$, is taken with the Boltzmann factor, $exp\left(-\beta H\right)$. Since the significant contributions to the Hamiltonian come from the long-range part of interactions, the continuum limit is applicable. Thus, the mean square distance reads

---

[1]The solution of Poisson equation for a Coulomb gas at a square lattice reveals the logarithmic interaction, as I present in section 4.2.4.

$$< r^2 >= \frac{\int_{r_0}^{\infty} r^3 dr \quad exp\left(-2\beta q^2 \ln\left(r/r_0\right)\right)}{\int_{r_0}^{\infty} r dr \quad exp\left(-2\beta q^2 \ln\left(r/r_0\right)\right)} = r_0^2 \frac{\beta q^2 - 1}{\beta q^2 - 2},$$

where $\beta = \frac{1}{k_B T}$. From this, one can observe that the mean square distance diverges in the limit $\beta q^2 \to 2$. Thus, the polarizability of the system, defined as $p(r) = \frac{1}{2}\beta q^2 r^2$, diverges, too.

Introducing a dielectric constant may describe the effective interaction between pairs. Iterative calculations in the framework of mean-field theory confirm the transition. Those calculations improve the value of the critical temperature by a correction. The proof is beyond the scope of this thesis. However, it reveals the mechanism underlying this transition. Above the critical temperature, the largest bounded pairs dissociate, the polarizability diverges, and the dielectric constant becomes infinite. The density of dissociated pairs tends to zero as the temperature approaches T_C from above, and hence, the system becomes dielectric. Thus, unbinding-pair charges drive a phase transition from a dielectric to a plasma phase at a critical temperature.

The following shows that the XY system is statistically similar to the two-dimensional Coulomb gas. Hence, the results of this model system are also applicable to the XY systems.

4.2 The KT-transition in two-dimensional XY system

4.2.1 The XY model

As mentioned in 3.1.2, the two-dimensional XY model is a classical lattice model of an interacting spin system. Here, for simplicity, the lattice is considered as a simple square lattice with the lattice constant a. Spins interact only with nearest neighbors; hence, the Hamiltonian is

$$H = -\sum_{i,j} J S_i \cdot S_j = -\sum_{i,j} J \cos(\theta_i - \theta_j),$$

where θ_i is the angle of the ith spin relative to an arbitrary direction, and j runs only over the nearest neighbors of spin i. A uniform alignment of spins defines the configuration that minimizes the Hamiltonian. Two ground states of the XY Hamiltonian are shown in figure 4.1. The configuration of the lattice can be, alternatively, given by a set of angles $\theta(\vec{r})$, where $\vec{r}$ denotes the site defined by $\vec{r} = \sum_i n_i \vec{a_i}$, for an integer n_i and lattice vectors $\vec{a_i}$. Thus, the Hamiltonian is

$$H = -J \sum_{\vec{r},\vec{r'}} \cos(\theta_r - \theta_{r'}),$$

such that $|\vec{r} - \vec{r'}| = a$.
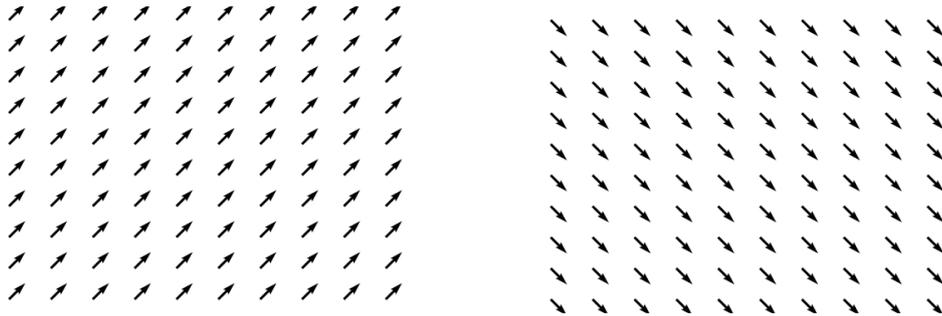
Figure 4.1: A two-dimensional XY system is a classical lattice model of an interacting spin system. Spin is a classical vector in the plane of the lattice. The ground state of the XY model is the configuration in which all spins point out the same direction. Two configurations represent examples of the ground state for the XY Hamiltonian.

As I present in the following, the spin-spin correlation function shows a quasi-long-range order as the temperature reaches a critical value from above. Hence, a transition from a disordered phase to a quasi-long-range order is expected in this system.

4.2.2 Spin-spin correlation function

Since, in general, having the partition function is sufficient to obtain the value of an observable, I first express the energy and then the partition function of the XY system in two temperature limits. After that, I will discuss the behavior of the spin-spin correlation function within these limits.

At low temperatures, only a slowly varying configuration of spins contributes significantly to the partition function, the configuration in which the angles defining neighboring spin directions are nearly equal. Thus, the expansion of Hamiltonian about its local minimum to the quadratic terms gives [3]

$$
\begin{aligned}
H - E_0 &\approx -J \sum_{r,r'} \left(-\frac{1}{2} \left(\theta_r - \theta_{r'} \right)^2 \right) \\
&= -J \sum_{r,r'} \theta_r \, \triangle_{r,r'} \, \theta_{r'},
\end{aligned}
$$

where $\triangle_{r,r'}$ is a matrix operator representing the first difference operator and is an analogue to the Laplacian operation. It is defined as

$$
\triangle_{r,r'} = \begin{cases} -2 & r = r' \\ 1 & |r - r'| = a \\ 0 & otherwise. \end{cases}
$$

The long-range effects dominate critical properties of the system [28], which are expressed sufficiently accurately in the continuum limit. Hence, the set of angels, $\{\theta_r\}$, translates to a continuous function, $\theta(r)$. The Hamiltonian, $H - E_0 \approx \frac{J}{2} \int d^2r |\nabla\theta|^2$, can be solved analytically [3], with the resulting distribution being Gaussian. For a Gaussian distributed variable, in two dimension, $< \left(\theta(r) - \theta(r') \right)^2 >$ increases logarithmically with distance [3]. Therefore, using the Gaussian integral properties gives

$$
< S_i \cdot S_j >= \mathrm{Re} \left(e^{-\frac{<\left(\theta_i - \theta_j\right)^2>}{2}} \right) \approx e^{-\ln\left(\frac{|r_i - r_j|}{r_0}\right)/4\pi K(T)} = \left(\frac{|r_i - r_j|}{r_0} \right)^{-\frac{1}{4\pi K(T)}},
$$

where $K(T) = \beta J$, and $\frac{r_0}{a} = \frac{e^{-\gamma}}{2\sqrt{2}}$. Therefore, at low temperatures, the spin-spin correlation decays algebraically, so the system does show quasi-long-range order.

At high temperatures, the partition function of the system can be expanded in terms of $-\beta H$,

$$Z = \int\limits_0^{2\pi} \Pi_i \frac{d\theta_i}{2\pi} e^{-\beta H} = \int\limits_0^{2\pi} \Pi_i \frac{d\theta_i}{2\pi} \left(\sum_p \frac{\left(K(T) \sum_{i,j} \cos(\theta_i - \theta_j) \right)^p}{p!} \right)$$

where $K(T) = \beta J$ and, $-\beta H = \beta \sum_{i,j} J \cos(\theta_i - \theta_j) = K(T) \sum_{i,j} \cos(\theta_i - \theta_j)$.

Regarding the high-temperature limit, the spin-spin correlation function is

$$< S_0 \cdot S_x > = < \cos(\theta_0 - \theta_x) >$$

$$= \mathrm{Re} \left(\frac{1}{Z} \int\limits_0^{2\pi} \Pi_i \frac{d\theta_i}{2\pi} e^{i(\theta_0 - \theta_x)} \right) \left(\sum_p \frac{\left(K(T) \sum_{i,j} \cos(\theta_i - \theta_j) \right)^p}{p!} \right).$$

The exponent $(\theta_0 - \theta_x)$, in general, can be considered as a path connecting sites r_0 and r_n via neighboring sites. By this consideration, the exponent can be rewritten as

$$(\theta_0 - \theta_x) = (\theta_0 - \theta_1) + (\theta_1 - \theta_2) + \cdots + (\theta_{n-1} - \theta_n),$$

where each term represents the revolution of spin along the bond connecting two neighboring sites, as shown in figure 4.2. Thus,

$$< S_0 \cdot S_x > =$$

$$= \mathrm{Re} \left(\frac{1}{Z} \int\limits_0^{2\pi} \Pi_i \frac{d\theta_i}{2\pi} e^{i((\theta_0 - \theta_1) + (\theta_1 - \theta_2) + \cdots + (\theta_{x-1} - \theta_x))} \right) \times$$

$$\left(1 + K(T) \sum_{i,j} e^{i(\theta_i - \theta_j)} + \frac{(K(T))^2}{2} \sum_{i,j} \sum_{i',j'} e^{i\left((\theta_i - \theta_j) + (\theta_{i'} - \theta_{j'}) \right)} + \cdots \right).$$

Since $\int d\theta = 2\pi$, and $\int d\theta \cos \theta = 0$, therefore, the non-zero terms are the terms where the exponent adds to zero. Hence, the first none-zero term is a term of the
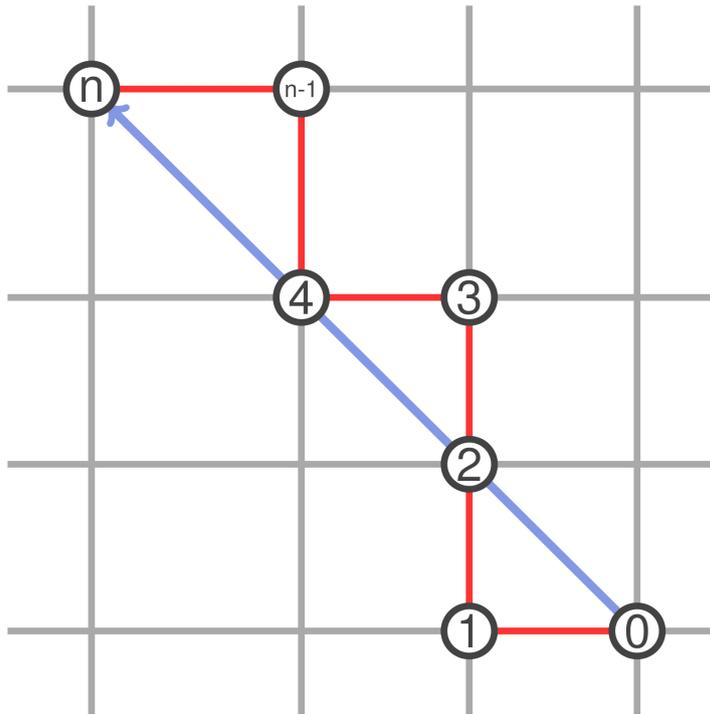
Figure 4.2: The red lines show the path connecting sites r_0 and r_n via neighboring sites. The sum of spin revolution along the bond-connecting path, i.e., $\theta_0 - \theta_x$, appears in the spin-spin correlation function. The blue line represents the distance between the two sites. At large distances, the length of the blue line, r, is in the order of n.

order n, $(K(T))^n$. Therefore,

$$< S_0 \cdot S_x > \sim (K(T))^n .$$

The simple sketch in figure 4.2 shows that, at large distances, where the physical behavior of the system is nearly homogenous, $|r| \sim n$. Thus,

$$< S_0 \cdot S_x > \sim (K(T))^{|r|} = e^{(r \ln K(T))}.$$

At the limit of high temperatures, $\ln K(T)$ is negative. So, we conclude that the spin-spin correlation function, $e^{-\frac{r}{\xi}}$, decays exponentially.

The mean magnetization, $m = \frac{<S>}{N}$, is the order parameter of the system, where N

is the number of sites. The mean magnetization takes zero value at all temperatures [44]. This statement is evident by direct calculation; see, for instance, Appendix A. Thus, the model does not show a usual phase transition, i.e., a disordered phase to an ordered phase with a finite magnitude of the mean magnetization. However, the system undergoes another type of transition, a continuous phase transition, in which the correlation function changes from exponential decay to an algebraic relation [28]. This transition, i.e., the transition from the disordered phase at high temperatures to a quasi-long-range ordered at a critical temperature, is called the Kosterlitz-Thouless transition. Vortices, the topological defects in the XY system, drive this phase transition, as I explain in the following. To this end, I will introduce the vortex as the topological defect in the XY system. Then, I represent the energy associated with a single vortex and the interaction exchange energy of vortices. The analogy of the interaction energy of vortices and a model system of interacting charged particles introduced in 4.1 completes the study and addresses the KT transition. To present this section, I follow [6, 28, 29].

4.2.3 Topological defect in the XY system: Vortex

Vortices are topological defects in a system, such as the XY model. The angular variable $\theta(x)$ defines the spin in this system, where $x = (r, \phi)$ denotes the polar coordinate of sites. The order parameter of the system is the mean magnetization, or simply, $< S(x) >=< S \left(\cos \theta(x)\hat{i} + \sin \theta(x)\hat{j} \right) >$. One can find a subspace d_s where the order parameter shows a singularity. To clarify, consider a spin configuration defined by function $\theta(x) = \phi + \frac{\pi}{2}$, as shown in figure 4.3-a. This configuration is associated with a singularity of dimension zero at the origin: At the origin, the direction of the spin or the derivative of θ, $\nabla \theta = \frac{1}{r}\hat{e}_\phi$, is infinite. Hence, there exists a singularity at the origin. The singularity is called the core of the defect. Since vortices in 2D are point defects, $d_s = 0$.

A vortex is characterized by its strength. I would give the mathematical definition of the strength of a vortex. It is defined by the total angle changes θ along an arbitrary path enclosing the singularity. For instance, in figure 4.3, two vortices of strength ± 1 are sketched. The strength does not uniquely define the field. Different functions can address different spin configurations of the same strength, e.g., $\theta(x) = \phi + \theta_0$ for any arbitrary constant θ_0 represents a vortex of strength $+1$. The strength of a vortex is related to a discrete charged particle and, hence, is called a topological charge. The

word vorticity is alternatively used instead of strength.
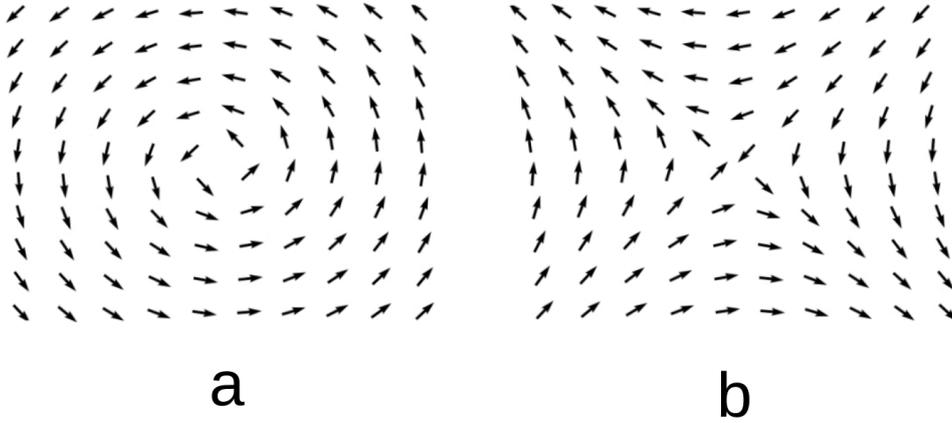
Figure 4.3: Two spin vortices on a simple square lattice are shown. Panel **(a)** presents a vortex with positive vorticity and, panel **(b)** shows a vortex with negative vorticity. Vortices are topological defects that drive transition in some two-dimensional systems, e.g., the XY model.

4.2.4 Energy of vortices

At low temperatures, the expansion of Hamiltonian about its local minimum to the quadratic terms, which leads to the approximation called the spin-wave approximation [4], gives

$$H - E_0 \approx \frac{J}{2} \sum_{i,j} (\theta_i - \theta_j)^2 = \frac{J}{2} \sum_{r} (\triangle\theta(r))^2,$$

where $\triangle$ is the first deference operator, the sum is taken over all neighboring sites. In the continuum limit, the energy associated with the formation of a single vortex is

$$\sum_{r} (\triangle\theta(r))^2 \to \int |\nabla\theta(x)|^2 d^2r \quad \Rightarrow \quad E_v \approx \frac{J}{2} \int_{a}^{L} \left(\frac{1}{r}\right)^2 (2\pi r dr) = J\pi \ln\left(\frac{L}{a}\right).$$

Here, L is the linear size of the system. Due to the logarithmic behavior of the energy concerning the size of the system at low temperatures, in the thermodynamic limit, a single vortex is not favorable.

A classic argument predicts that vortices can exist abundantly at a critical temperature, as I will present in the following. In a square lattice of size L, there exist $(\frac{L}{a})^2$ sites. Then, the entropic contribution of such a vortex yields

$$S = k_B \ln\left(\frac{L}{a}\right)^2,$$

where k_B is the Boltzmann constant. Both together lead to the free energy

$$F_v = U - TS = J\pi \ln\left(\frac{L}{a}\right) - 2k_B T \ln\left(\frac{L}{a}\right) = (J\pi - 2k_B T)\ln\left(\frac{L}{a}\right).$$

At low temperatures, the entropic term does not play a role. Hence, the free energy takes positive values, and the minimum of U in the above calculation minimizes the free energy. At high temperatures, the entropic term is dominant, and the sign of the free energy changes at a critical temperature, $T_C = \frac{J\pi}{2k_B}$. At temperatures $T > T_C$, the free energy gets negative values, meaning that free vortices are favorable and abundant. However, as I present later, the logarithmic divergence with the size disappears for even number vortices with total zero vorticity. Hence, they can be formed thermally and exist as bounded vortices at low temperatures. Thus, the KT transition is topologically a transition from a forbidden-single-vortices phase to a phase where vortices are free.

Consider the vortex shown in figure 4.4. A function $\theta(x)$ describing this configuration is not single-valued. For each round on a closed path, enclosing the core, $\theta(x)$ changes by 2π, or simply,

$$\oint d\theta(x) = 2\pi.$$

The shortest closed path is a circle of radius $\sqrt{2}a$ whose center lies on a dual lattice, a lattice whose sites x^* located at the center of the original lattice, as shown in figure 4.4. This visualization maps the distribution of vortices to an electrostatic problem. Hence, a configuration of vortices now is, in analogy, a distribution of discrete charged particles located at the dual lattice sites, such that $\rho(x^*) = \sum_i q_i \delta(x^* - x_i^*)$ describes the charge distribution. Thus, $\theta(x)$ is a potential produced by that distribution in 2D. Hence, $\theta(x)$ obeys Poisson's equation
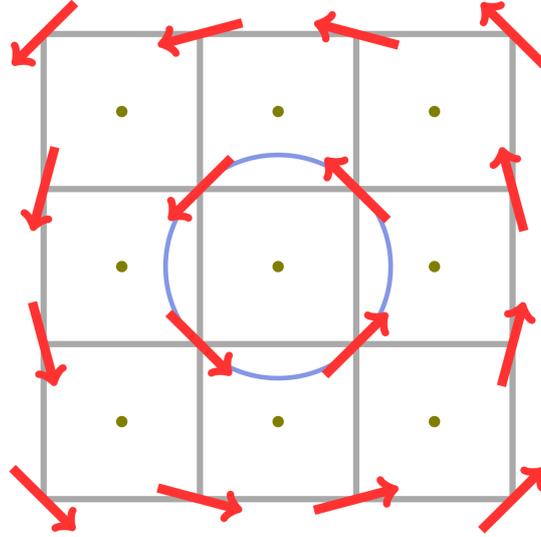
Figure 4.4: The sketch shows a positive vortex of strength $+1$ on a square lattice. The vortex system is mapped to a discrete charged system by considering a dual lattice. The green dots located at the center of the studied lattice present the sites of the dual lattice. The dual lattice is square, too.

$$\nabla^2 \theta(x^*) = 2\pi\rho(x^*).$$

The solution of this equation reads

$$\theta(x) = 2\pi \int d^2x' \rho(x')g(x - x') + O\left(\ln\left(\frac{R}{x_0}\right) \int d^2x \rho(x)\right),$$

where $g(x) = \frac{1}{2\pi} \ln\left(\frac{x}{r_0}\right)$ is the Green's function of a two-dimensional square lattice defined so that $g(0) = 0$ [26]. In the second term, since $\ln\left(\frac{R}{r_0}\right)$ diverges with the size of the system; therefore, a physical solution requires that $\int d^2x \rho(x) = 0$, meaning that the analogous system, similar to the model system in 4.1, is electrically neutral. Thus

$$\theta(x^*) = 2\pi \sum_i q_i g(x^* - x_i^*) = \sum_i q_i \ln\left(\frac{x^* - x_i^*}{r_0}\right).$$

Now, Let's go back to the original Hamiltonian expanded about its local minimum,

$$H - E_0 = \frac{J}{2} \sum_x \left(\triangle \theta(x) \right)^2 .$$

Consider both spin-wave excitation, i.e., deviation from local minimum, and vortex configuration. Let $\theta(x) = \bar{\theta}(x) + \psi(x)$, where $\bar{\theta}(x)$ defines the direction of a spin in its local minimum, and $\psi(x)$ the deviation of this configuration such that $\oint d\psi(x) = 0$. Hence, the Hamiltonian reads

$$
\begin{aligned}
H - E_0 &= \frac{J}{2} \sum_x \left(\triangle \bar{\theta}(x) + \triangle \psi(x) \right)^2 \\
&= \frac{J}{2} \sum_x \left(\triangle \bar{\theta}(x) \right)^2 + \frac{J}{2} \sum_x \left(\triangle \psi(x) \right)^2 + J \sum_x \left(\triangle \bar{\theta}(x) \cdot \triangle \psi(x) \right),
\end{aligned}
$$

where, by using $\oint d\psi(x) = 0$, the cross term vanishes. Using the charged system analogy in the continuum limit gives:

$$H - E_0 = \frac{J}{2} \int d^2 r \left(\psi(r) \right)^2 - \pi J \sum_{i \neq j}^{2n} q_i q_j \ln \left(\frac{r_i - r_j}{a} \right) + \mu \sum_i q_i^2,$$

where the details of this calculation are provided in Appendix B. Since the spin-wave excitation does not interact with the vortices, the problem is equivalent to the two-dimensional Coulomb gas introduced in section 4.1 with the interaction energy of

$$
E_v =
\begin{cases}
-\pi J \sum_{i \neq j} q_i q_j \ln \left| \frac{r_i - r_j}{a} \right| + \mu \sum_i q_i^2 & |r_i - r_j| > a \\
0, & otherwise.
\end{cases}
$$

Hence, the KT transition is a transition from bounded vortices to a state in which they are free.

In summary, the KT transition observed in the XY systems is a continuous transition driven by topological defects. While the mean magnetization is zero at any temperature, the order of the system characterized by the spin-spin correlation function switches from quasi-long-range at low temperatures to disorder at high temperatures.

4.3 Two-stage melting transition: KTHNY theory

A two-dimensional solid has much in common with the XY model. According to the Mermin-Wagner theorem, long-range order is impossible in a 2D crystal. Nevertheless, it does show long-range topological order associated with the response to shear stress [29]. Kosterlitz and Thouless proposed the KT transition for a 2D melting transition. The continuum elasticity theory describes the response of the solid to shear stress. A solid is rigid, meaning it resists shear stress. Thermal fluctuations at low temperatures lead to the formation of pairs of bounded dislocations, which are topological defects in a crystal and reduce rigidity. At a higher temperature, T_M, isolated dislocations are possible. Free dislocations cause the system to be unstable against shear stress. Thus, the system loses the long-range topological order. However, the two-dimensional transition is more complex. The state of the matter just above T_M is not isotropic [47]. It is still orientationally quasi-long-range order, evident by the orientational order parameter measurement. Hence, another step is essential to complete the melting transition. Furthermore, the substrate on which a 2D solid is adsorbed can influence the melting transition.

The following section presents the two-step melting transition according to a theory that Nelson, Halperin, and Young elaborated regarding the KT transition, known as the KTHNY theory. A two-dimensional crystal is both positionally and orientationally ordered. To present the KTHNY theory, I first introduce order parameters concerning underlying orders in the system. Then, I will introduce the topological defects in a 2D solid. Finally, I summarize the basics of the KTHNY theory according to [47] on a smooth substrate.

4.3.1 Fundamentals of solid states

A perfect crystal is a solid consisting of identical copies of a unit arranged in a highly ordered microscopic structure. The repeated pattern is simply a distribution of mass and charge, and it may be composed of, e.g., a single or a group of atoms, ions, or molecules. Since this periodic configuration affects the fundamental physical features of the matter, I survey a few of the most important geometrical properties of the lattice configuration, particularly in two dimensions[2].

---

[2]In this section, I follow the solid-state textbook of Ashcraft [2].
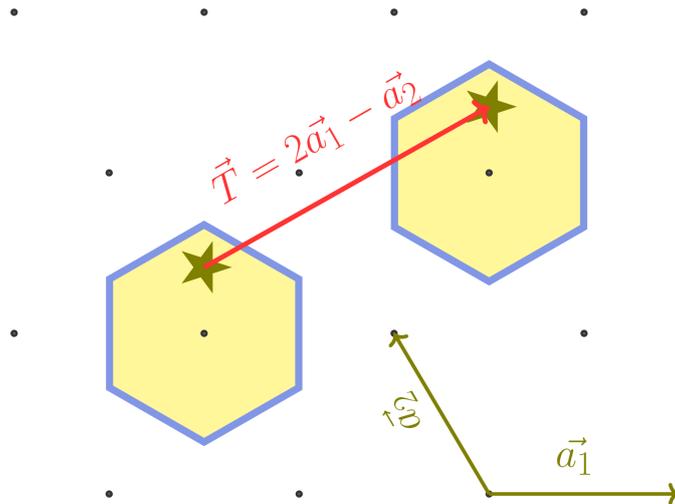
Figure 4.5: The triangular lattice provides an example of a two-dimensional Bravais lattice. The black dots indicate lattice sites. The set of vectors $\vec{a_i}$ is an exemplary basis set of this lattice. The yellow hexagons represent the unit cells. A translational vector, $\vec{T}$, connects two equivalent points in different unit cells.

The periodic array specifying a crystal is called the Bravais lattice. A Bravais lattice is an array of points defined by a vector $\vec{R} = \sum_{i=1,2} n_i \vec{a_i}$, where n_i are integer numbers and $\vec{a_i}$ are a set of basis vectors concerning lattice arrangement. The set of basis vectors is not unique. Figure 4.5 shows the Bravais lattice and a set of the basis of a two-dimensional triangular lattice. A primitive unit cell is the smallest area that builds the entire lattice without any overlap or leaving a void. A translational vector, $\vec{T}$, connects two equivalent points in different unit cells. Since each point in a Bravais lattice follows the periodic structure, the vector $\vec{T}$ is a summation of the basis set, $\vec{T} = \sum_{j=1,2} n_j \vec{a_j}$, where n_j are integer numbers. The existence of the translational vector in the Bravais lattice results from broken translational symmetry in crystals.

For a given three-dimensional Bravais lattice, a set of parallel planes constructs a family set such that each of them builds a two-dimensional Bravais lattice. Each original lattice site belongs to one of these planes. A unique way to classify a different set of families comes from introducing a second Bravais lattice correspondence to a given lattice in the momentum space called a reciprocal lattice.

While the Bravais lattice represents the periodic arrangement in the physical space, the Fourier transform of a Bravais lattice, called the reciprocal lattice, refers to the

periodicity in the momentum space. It plays a fundamental role in studies related to a periodic structure. A particular reciprocal lattice refers to a specific Bravais lattice. For a given Bravais lattice defined by a set of vectors $\vec{R}$, there exists a set of wave vectors $\vec{K}$ whose plane wave $e^{i\vec{K}\cdot\vec{r}}$ has the same periodicity as the Bravais lattice. Such a set of $\vec{K}$ defines the reciprocal lattice and satisfies

$$e^{i\vec{K}\cdot r} = e^{i\vec{K}\cdot\left(\vec{r}+\vec{R}\right)},$$

for any $\vec{r}$ in the real space. Hence, $e^{i\vec{K}\cdot\vec{R}} = 1$. The reciprocal lattice is itself a Bravais lattice. For the proof see, for instance, [2]. Hence, for a given Bravais lattice defined by the set of basis vectors $\vec{a_i}$, the reciprocal lattice is defined by the set of wave vectors $\vec{K}$ such that $\vec{K}$ is a sum of a basis set, $\vec{b_k}$, i.e., $\vec{K} = \sum_k m_k \vec{b_k}$, where,

$$\vec{b_k} = \frac{\pi}{|v|}\epsilon_{ijk}\left(\vec{a_i}\times\vec{a_j}\right).$$

Here, ϵ_{ijk} is the Levi-Civita symbol, and $|v| = \vec{a_1}\cdot(\vec{a_2}\times\vec{a_3})$ is the volume of the primitive cell. The following theorem represents the relation between a set of family planes in the physical space and reciprocal lattice:

For any set of family planes separated by distance d, there exist reciprocal lattice vectors perpendicular to them with the shortest length of $\frac{2\pi}{d}$. Conversely, for any reciprocal lattice vector $\vec{K}$, there is a set of family planes separated by distance $\frac{2\pi}{d}$, where d is the length of the shortest reciprocal vector.

In the field of condensed matter, scattering experiments provide crucial information. The conditions, called the Laue conditions, connect the scattering experiments and the microscopic structure. The Laue conditions state that the scattering of a lattice causes constructive interference if the change in the wave vector is equivalent to one of the reciprocal lattice vectors. In the following, I survey the fundamentals of scattering experiments and how such experiments describe the microscopic structure of matter.

4.3.2 Order parameters in a two-dimensional crystal

Scattering problem and the structure factor

Let's consider a discrete configuration of individual sites in two dimensions, as shown in figure 4.6. Each site is represented by the vector of position, $\vec{x}_n$, where n is the number of sites. We are interested in the elastic scattering of the entire system, in which the initial and final energy of the scattered particle is identical. Each site describes a scattering center of a weak potential $U(\vec{x}_n)$ for scattering particles. The scattering particles are indicated by an arrow in figure 4.6. First, I will address the issue for one site and then generalize it to the entire system. For simplicity, I do not use the index n, meaning that the phrase $U(\vec{x})$ represents the scattering potential either of a site or the whole system. Generalizing to the entire system is simply possible by substituting $U(\vec{x})$ by the total potential $U(\vec{x})$, where $U(\vec{x}) = \sum_n U(\vec{x}_n)$.
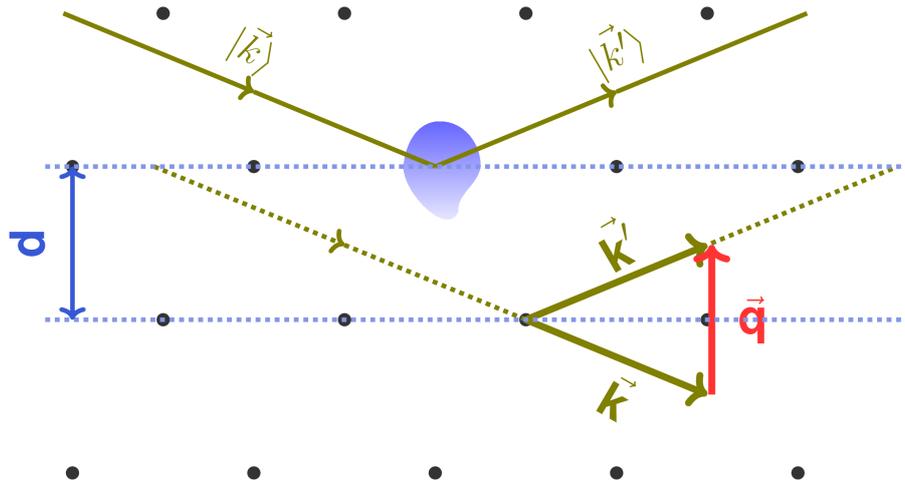
Figure 4.6: Scattering problem in two dimensions; A discrete configuration of individual sites in two dimensions is shown. Each site, shown as a black dot, describes a scattering center of weak potential. One scattering center is shown with a general shape schematically. Typical scattering geometry is shown by the incident, final, and scattering wave vectors $\vec{k}$, $\vec{k}'$, and $\vec{q} = \vec{k}' - \vec{k}$. Scattering from two parallel planes separated by a distance d is depicted.

Both incident and scattered particles are described by plane wave states, $|\vec{k}\rangle$ and $|\vec{k}'\rangle$, each of them has the momenta $\hbar k$ and $\hbar k'$, respectively. For a sufficiently weak potential, Fermi's golden rule gives the transition rates between the initial and final

state in the framework of perturbation theory. According to Fermi's golden rule, the transition rate between states $|\vec{k}\rangle$ and $|\vec{k'}\rangle$ is proportional to the square of the matrix elements of the perturbation field. In a d-dimensional space, the matrix elements are:

$$M_{\vec{k},\vec{k'}} = \langle \vec{k}|U|\vec{k'}\rangle = \int d^d x\, e^{-i\vec{k}\cdot\vec{x}} U(\vec{x}) e^{i\vec{k'}\cdot\vec{x}},$$

where $U(\vec{x})$ is the perturbation potential in the coordinate representation of the scattering particle, $\langle \vec{x}|\vec{k}\rangle = e^{i\vec{k}\cdot\vec{x}}$ presents the state of the scattered particle, and $\vec{k}$ is a vector. For elastic scattering, in which the initial and final energy of the scattered particle is identical, the transition rate, $\Gamma_{\vec{k}\to\vec{k'}}$, is:

$$\Gamma_{\vec{k}\to\vec{k'}} = \frac{2\pi}{\hbar}|M_{\vec{k},\vec{k'}}|^2.$$

The generalization of the problem to the scattering of a potential consisting of individual atoms yields:

$$\langle \vec{k}|U(\vec{x})|\vec{k'}\rangle = \sum_n \int d^d x\, e^{-i\vec{k}\cdot\vec{x}} U_n(\vec{x}-\vec{x_n}) e^{i\vec{k'}\cdot\vec{x}},$$

where $U(\vec{x}) = \sum_n U_n(\vec{x}-\vec{x_n})$ is the perturbation potential arising from individual N sites each of them placed at $\vec{x_n}$.

It is convenient to represent the matrix elements corresponding to the scattering direction from each atom. Hence, by defining $\vec{R_n} = \vec{x}-\vec{x_n}$ and $\vec{q} = \vec{k}-\vec{k'}$, the matrix elements are:

$$\langle \vec{k}|U(\vec{x})|\vec{k'}\rangle = \sum_n \int d^d R_n e^{-i\vec{k}\cdot\left(\vec{x_n}+\vec{R_n}\right)} U_n(\vec{R_n}) e^{i\vec{k'}\cdot\left(\vec{x_n}+\vec{R_n}\right)}$$

$$= \sum_n \int d^d R_n e^{-i\vec{q}\cdot\vec{R_n}} U_n(\vec{R_n}) e^{-i\vec{q}\cdot\vec{x_n}}$$

$$= \sum_n U_n(\vec{q}) e^{-i\vec{q}\cdot\vec{x_n}}.$$

Where $U_n(\vec{q})$ is Fourier transform of the atomic potential. Therefore, the transition

rate is proportional to

$$|M_{\vec{k},k\vec{k}'}|^2 = \sum_{n,n'} U_n(\vec{q})U_{n'}^*(\vec{q})e^{-i\vec{q}\cdot\vec{x_n}}e^{i\vec{q}\cdot\vec{x_{n'}}}.$$

For a statistical system composed of identical atoms, we get $U_n(\vec{q}) = U_{n'}(\vec{q})$. By the assumption that the system is ergodic, the transition rate for such a system is equivalent to the ensemble average over all possible configurations. Hence, it becomes

$$\Gamma_{\vec{k}\to\vec{k}'} = \frac{2\pi}{\hbar}|U_n(\vec{q})|^2\Big\langle\sum_{n,n'} e^{-i\vec{q}\cdot(\vec{x_n}-\vec{x_{n'}})}\Big\rangle$$

$$= \frac{2\pi}{\hbar}|U_n(\vec{q})|^2 I(\vec{q}),$$

where the $\langle\ldots\rangle$ states for the ensemble average. In the last step, the summation part of this equation is replaced by the term $I(\vec{q})$, which depends only on the position of individual perturbating centers. This parameter is called the structure function, $I(\vec{q})$. For a randomly distributed system, the only non-vanishing terms are those with $n = n'$. The structure function is an extensive variable of the system, i.e., it increases by the size of the system. The intensive variable derived from the structure function is called the structure factor, $S(\vec{q})$, and it is simply the structure function divided by either N or the volume of the system, V,

$$S(\vec{q}) \equiv \frac{1}{N}\Big\langle\sum_{n,n'} e^{-i\vec{q}\cdot(\vec{x_n}-\vec{x_{n'}})}\Big\rangle.$$

Hence, the structure factor, measured experimentally, contains important details of the relative position of individual parts of the system. Specifically, it is the Fourier transform of the density-density correlation function, as I will discuss below.

density-density correlation function

The number density operator, $n(\vec{x})$, defines the number of particles per unit volume at position $\vec{x}$. For instance, the number density operator corresponding to the configuration of figure 4.6 is

$$n(\vec{x}) \equiv \sum_n \delta(\vec{x} - \vec{x_n}),$$

where $\delta(\vec{x})$ denotes the Dirac delta function. The ensemble average of the number density operator defines the average density of the system, $\langle n(\vec{x}) \rangle$. In a homogeneous system of complete symmetry, the average density does not depend on $\vec{x}$, meaning that the system is not correlated. While dependency on $\vec{x}$, either the amount or direction, is the signature of the symmetry breaking and expresses a correlation between particles.

The density-density correlation function, $C_{nn}(\vec{x_1}, \vec{x_2})$, measures this correlation, and it is defined as the average ensemble of the density operator

$$C_{nn}(\vec{x_1}, \vec{x_2}) = \langle n(\vec{x_1}) n(\vec{x_2}) \rangle = \langle \sum_{n_1, n_2} \delta(\vec{x} - \vec{x_1}) \delta(\vec{x} - \vec{x_2}) \rangle.$$

The density-density correlation function is related to the previously introduced structure function, $I(q)$, via Fourier transformation

$$n(\vec{q}) = \int d^d x \quad n(\vec{x}) e^{-i\vec{q} \cdot \vec{x}} = \sum_n \int d^d x \quad \delta(\vec{x} - \vec{x_n}) e^{-i\vec{q} \cdot \vec{x}}$$

$$= \sum_n e^{-i\vec{q} \cdot \vec{x_n}}.$$

Therefore, according to the definition of $I(\vec{q})$, the structure function is simply the Fourier transform of the density-density correlation function,

$$I(\vec{q}) = \langle \sum_{n,n'} e^{-i\vec{q} \cdot (\vec{x_n} - \vec{x_{n'}})} \rangle = \langle \sum_n e^{-i\vec{q} \cdot \vec{x_n}} \sum_{n'} e^{i\vec{q} \cdot \vec{x_{n'}}} \rangle$$

$$= \langle n(\vec{q}) n(-\vec{q}) \rangle.$$

Order parameters: Translational order parameter, Bond-orientational order parameter
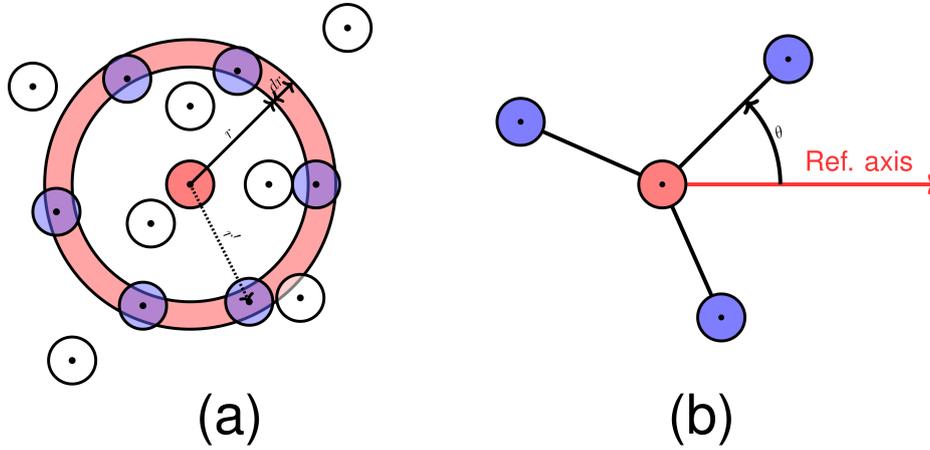
Figure 4.7: (a) Panel **(a)** shows how the pair correlation function, $g(r)$, is determined. $g(r)$ measures the probability of finding particles at a distance r of a given particle as a reference. In panel **(a)**, a permitted configuration of particles is shown in which a small circle depicts each particle; the pink particle presents the reference. A shell of dr at a distance of r of the reference particle is assumed. Each particle, shown in blue, whose center is located in the shell, is counted. $g(r)$ is determined by binning the pair radial distances into a histogram and normalizing them concerning an uncorrelated system distribution. **(b)** The bond-orientational order parameter is defined as $\psi_{nj} = \sum_{k}^{n_l} \frac{1}{n_l} e^{i(n\theta_{jk})}$, where θ is the angle between the physical bond and a fixed reference axis and n represents the expected symmetry in the lattice and equals to six for a triangular lattice interested here. In this context, a line connecting the center of two particles defines a physical bond. Panel **(b)** shows the reference axis and the angle between it and a bond.

An ideal crystal is a system in which translational symmetry holds discretely. A periodic function, $f(\vec{x}) = f(\vec{x} + \vec{T})$, presents the crystal in the real physical space, where $\vec{x}$ is a point and $\vec{T}$ is a translational vector. Based on the symmetry of the condensed matter, the translational order parameter is defined as

$$\psi_{Tj} = \exp\left(i\vec{K} \cdot \vec{r_j}\right)$$

proposes information regarding the structure of the system, where $\vec{r_j} = (x_j, y_j)$ states the position, and $\vec{K}$ states the primary vector of the reciprocal lattice. Thermal fluctuations deform the perfect lattice such that the position of individual atoms satisfies $\vec{r_j} = \vec{R} + \vec{u}(\vec{R})$, where $\vec{u}(\vec{R})$ defines the displacement from the perfect site $\vec{R}$. The global translational order parameter is defined as the average of all local trans-

lational order parameters, i.e.,

$$\psi_T = |\frac{1}{N} \sum_{j=1}^{N} \psi_{Tj}|,$$

provides information regarding positional order in a statistical system. Here, N is the number of sites.

The density-density correlation function, $\langle \rho(\vec{r})\rho(\vec{0})\rangle$, provides additional information, which is related to the structure factor via Fourier transform, and can be probed directly by diffraction experiments via

$$S(q) \equiv \langle \rho(\vec{q})\rho(-\vec{q})\rangle \approx \sum_{R} e^{i\vec{q}\cdot\vec{R}} \langle e^{i\vec{q}\cdot\left(u(\vec{R})-u(\vec{0})\right)}\rangle$$

The other function related to the density-density correlation function is the pair distribution function, $g(\vec{x_1}, \vec{x_2})$, which determines the probability of finding a particle at $\vec{x_2}$ if a given particle placed at $\vec{x_1}$, and is defined as

$$\langle n(\vec{x_1})\rangle g(\vec{x_1}, \vec{x_2})\langle n(\vec{x_2})\rangle \equiv \langle \sum_{n \neq n'} \delta\left(\vec{x_1} - \vec{x_n}\right)\delta\left(\vec{x_2} - \vec{x_{n'}}\right)\rangle.$$

A direct method of determining $g(\vec{x})$ in a permitted configuration is to choose a particle as the origin of the configuration. Then, consider a small volume at the separation $\vec{x}$ of the origin. $g(\vec{x})$ is equal to the number of particles in the volume; see figure 4.7-a. In a homogeneous system, where the density is independent of direction, the pair distribution function depends only on the radial distance, i.e., $g(\vec{x}) = g(r)$, where r denotes the radial separation of the origin, $r = |\vec{x}|$.

A solid in a crystal phase is, in addition, orientationally long-range ordered, which means that the orientation of bonds connecting nearest neighbors sites is not only correlated but also maintains a consistent pattern over large distances. To illustrate, consider two individual cells with a shared side in Voronoi tesselation, which are considered neighboring sites. This concept will be further explained in section 4.3.3. A measure related to this correlation is the bond-orientational order parameter, which is defined based on the expected symmetry in the studied system. In the case of a triangular lattice, which is the focus of this thesis, each site is orientationally sixfold

symmetric. Thus, the bond-orientational order is defined as [47]

$$\psi_{6j} = \sum_{k=1}^{n_l} \frac{1}{n_l} e^{6i\theta_{jk}},$$

where the sum runs over nearest neighbors of site j, θ_{jk} represents the orientation of the bond connecting two neighboring sites j, k relative to a fixed axis, and n_l is the number of nearest neighbors; see figure 4.7. The global bond-orientational order parameter, $|\psi_6| = |\frac{1}{N}\sum_{j=1}^{N}\psi_{6j}|$, quantifies the sixfold orientational symmetry in the system. The sum is taken over the number of particles, N. A system in a crystal phase exhibits a long-range orientational order, which means $|\psi_6| \approx 1$. For a liquid phase, on the other hand, bonds are oriented randomly, i.e., there is no preferred direction for bonds, which leads to $|\psi_6| \approx 0$.

The bond-orientational correlation function measures the spatial orientational ordering. In a triangular lattice, it is defined by:

$$g_6(|\vec{r} - \vec{r'}|) = \langle \psi_6(\vec{r})\psi_6^*(\vec{r'}) \rangle,$$

where $\psi_6(\vec{r}) = \frac{1}{N}\sum_{j=1}^{N}\psi_{6j}\delta(\vec{r} - \vec{r_j})$ is the local orientational order parameter of the particle at the position $\vec{r}$. This function is a crucial parameter in characterizing different phases in two-dimensional melting. According to the KTHNY scenario for melting, which I will discuss in detail in section 4.3.4, $g_6(r)$ approaches a constant in the crystal phase and decays algebraically and exponentially in the hexatic and liquid phases.

4.3.3 Topological defects in two-dimensional crystal

(a)   (b)
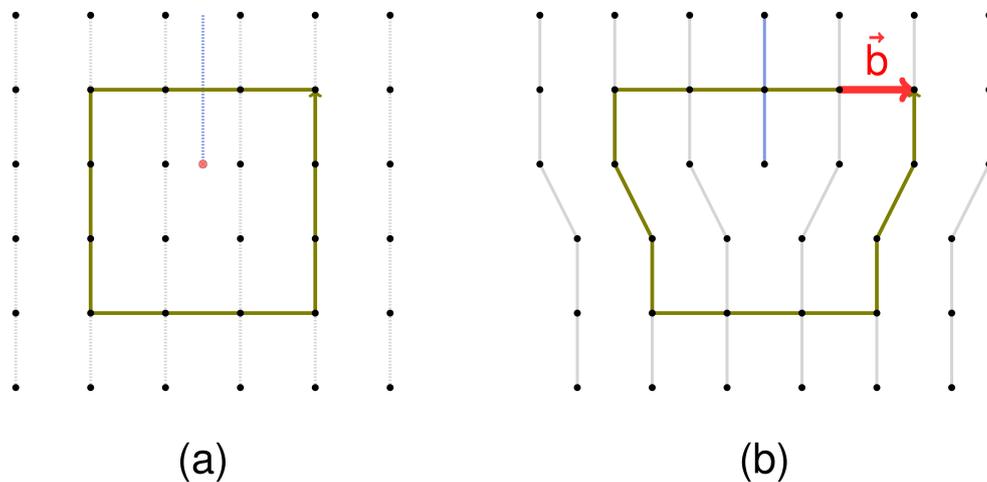
Figure 4.8: The figure shows an elementary dislocated solid and the related Burgers vector on a square lattice. A dislocated solid is formed if, for instance, a half-line is inserted in the perfect lattice. A Burgers vector characterizes the strength and direction of the dislocation. To find the Burgers vector, one needs to draw a closed-circuit surrounding the defected core called the Burgers circuit. The circuit begins at an arbitrary point, and it passes along the nearest neighbors in the counterclockwise direction until it reaches the starting point. The same procedure would take in the dislocated solid: The circuit would begin at the same position and move similarly through the nearest neighbors. The Burgers vector is the amount by which the path enclosing the defect would be closed compared to the perfect lattice. (a) A configuration of individual atoms of an ideal crystal on a simple square lattice and the Burgers circuit is shown. The point labeled in red represents the core of dislocation. The dashed line shows the position of the inserted half-line. The green path illustrates the Burgers circuit. It is composed of three steps to the left, down, right, and finally, up to the starting point. (b) The dislocated crystal and the corresponding Burgers vector are sketched. By the same procedure, the Burgers circuit fails to close by the vector $\vec{b}$.

An ideal two-dimensional crystal shows both translational and orientational order. A site in the lattice that does not possess these orders defines a defect. Thus, two kinds of topological defects are possible: dislocations and disclinations. Dislocations are defects in a system where the translational order has been lost. They are the first topological defects occurring in a crystal. Thermal fluctuations in an ideal crystal lead to dislocation formation, as I present shortly later for a triangular lattice. The nature of dislocation defects, e.g., on a square lattice, can be imagined by inserting an extra

row of atoms into the perfect lattice such as sketched schematically in figured 4.8-b.

Burgers vectors characterize dislocation defects. A Burgers vector describes the strength and the orientation of the dislocation. A practical way to determine the Burgers vector is to draw a loop counterclockwise around the defect points before and after defect formation, as shown in figure 4.8. In the defected lattice, the Burgers vector is the vector required to close the loop. In the case of inserting one half-line, the Burgers vector is equal to one of the lattice vectors in magnitude and is perpendicular to the imaginary inserted half-line. The simple square lattice shown in figure 4.8 provides a convenient example. Forthcoming, we consider the triangular lattice, which better explains what is coming in this thesis.

(a)   (b)
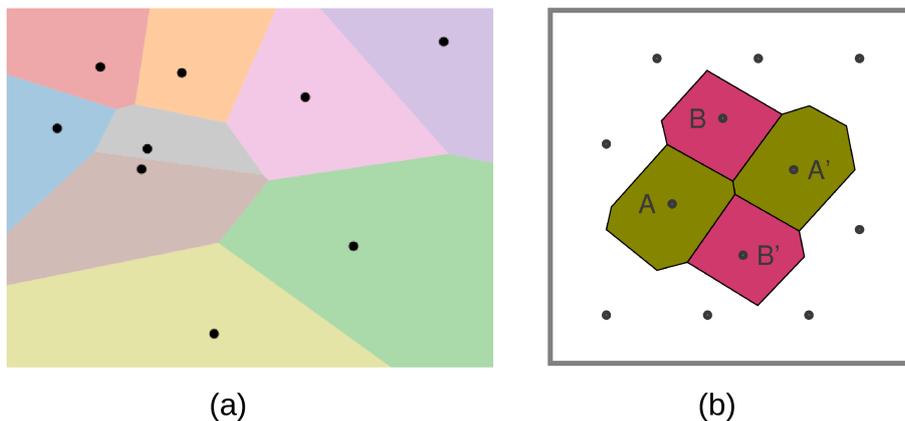
Figure 4.9: (a) The Voronoi tesselation for nine randomly distributed points. Each color presents the closest area to the enclosed point. **(b)** The Voronoi construction for a defected triangular lattice. A disclination defect is a cell with the wrong number of neighbors. The purple/green cells have five/or seven neighbors. These cells, individually, are called disclination defects. A single dislocation consists of a pair of bounded disclinations with a total zero disclinicity. This figure shows a pair of bounded dislocations.

I first introduce the Voronoi tesselation to visualize the other defects, i.e., disclinations. A Voronoi cell defines a region around a given site consisting of all points closer to this site than any other. Hence, the Voronoi tessellation constructs a mosaic of a given distribution of sites. In such a way, two sites are neighbors if their cells share at least one side of their Voronoi cell. Figure 4.9-a presents a Voronoi diagram for a randomly distributed configuration. In a perfect triangular lattice, each site is supposed to have six neighbors. So, a site with, e.g., five or seven neighbors is a defect

called a disclination. Figure 4.9-b shows the Voronoi construction and disclinations in a triangular lattice.

While a Burgers vector characterizes dislocations, a scalar number called disclinicity, $s(\vec{r})$, characterizes disclinations. This scalar is a positive or negative angle, i.e., two sides of the defect core in the ideal crystal must be twisted relative to each other to form disclinations. In a triangular lattice, these angles are $s(\vec{r}) = \pm\frac{\pi}{3}$, where the positive value serves the five-fold cell. Figure 4.9-b also represents an essential concept for creating crystal defects: A single dislocation consists of a pair of disclinations. The appearance of dislocations formed as pair-bounded defects is a consequence of thermal fluctuations at low temperatures.
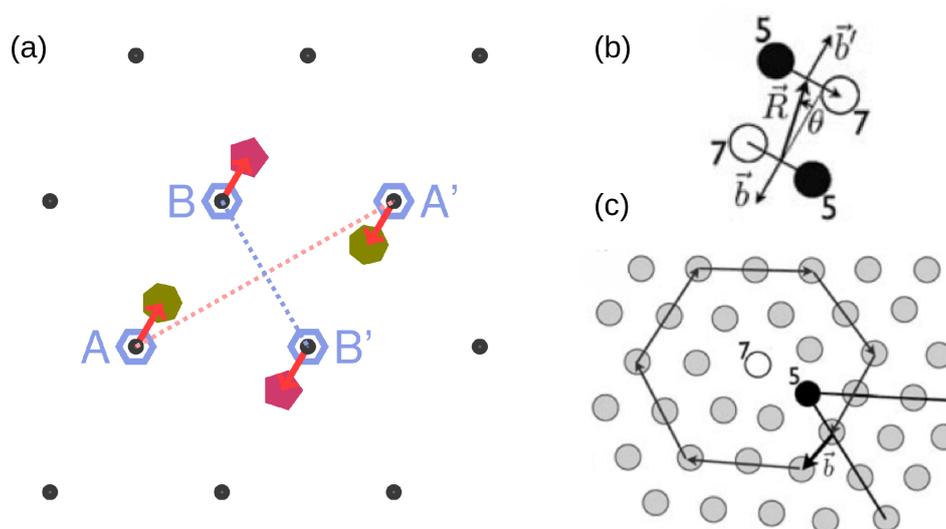
Figure 4.10: (a) Thermal fluctuations form defects. In the perfect crystal, sites A and A' are not neighbors, while B and B' are. A thermal movement displaces two particles in one of the lines in the opposite direction of the two particles in the parallel line in such a way that in each line, one particle loses, and the other one gains a neighbor. The purple particle has five, and the green has seven neighbors. This procedure results in a pair of dislocations consisting of a pair-bounded disclination. **(b)** A bounded dislocation and the Burgers vector describing each dislocation are shown. The black/white particles are five/or seven-fold sites. The vector R gives the orientation of one dislocation with respect to the other. **(c)** Burgers circuit and the Burgers vector in a triangular lattice. **(a)** is reproduced with permission through personal correspondence with the authors. **(b)** and **(c)** reprinted from [13], Copyright (2010), with permission from John Wiley & Sons, permission license Nr. 5117600835308.

Figure 4.10 shows a schematically thermal pathway of defect formation and Burgers vector in a triangular lattice. Four sites in a perfect triangular lattice are labeled. Sites B and B' are neighbors while A and A' are not. Thermal fluctuations form a defect by displacing these sites, as shown in the figure. For a certain amount of displacement, sites A and A' become nearest neighbors instead of B and B'. Hence, a pair of bounded dislocations occur. A higher temperature may form this pair at further separation, i.e., appearing single dislocations, and the crystal loses its translational order.

To summarize this section, a perfect crystal is both a translationally and orientationally ordered system. Thermal fluctuations lead to appearing defected cores. The melting transition begins by losing the translational order. At the first stage, bounded-dislocations of $(5, 7, 5, 7)$-fold sites appear in the ideal crystal. At high enough temperatures, the bounded dislocations appear at further distances and form single dislocations. By occurring single dislocations, the solid loses its global translational order. A single dislocation consists of a pair of bounded disclinations or a pair of $(5, 7)$-fold sites. Increasing temperature will dissociate it into two isolated disclinations. This process destroys the orientational order of the system and completes the melting transition from solid to isotropic liquid.

4.3.4 KTHNY picture of melting

According to the Mermin-Wagner theorem, a 2D solid does not show long-range translational order, evidenced by an algebraic decay in the density-density correlation function. However, it has orientationally long-range order, meaning the bond-orientational correlation function approaches a constant for large r. Thus, a two-dimensional solid has a quasi-long range translational and long-range orientational order [47].

I have discussed that dislocations and disclinations are topological defects regarding a two-dimensional crystal. At low temperatures, thermal fluctuations form dislocations. An argument regarding elastic theory, which is out of the scope of this thesis, reveals that the energy cost of the formation of a single dislocation diverges logarithmically with the size of the system [6]. So analogous to the XY system presented in section 4.2, isolated dislocations are not favored. A pair of dislocations of two opposite Burgers vectors separated by a lattice constant has finite energy. Hence, it is

favored. The Hamiltonian of a system of interacting dislocations is [47]:

$$H_{DL} = -\frac{1}{8\pi} \sum_{\vec{r} \neq \vec{r'}} \left(K_1 \vec{b}(\vec{r}) \cdot \vec{b}(\vec{r'}) \ln\left(\frac{|\vec{r} - \vec{r'}|}{a}\right) - K_2 \frac{\vec{b}(\vec{r}) \cdot \left(\vec{r} - \vec{r'}\right) \vec{b}(\vec{r'}) \cdot \left(\vec{r} - \vec{r'}\right)}{|\vec{r} - \vec{r'}|^2} \right)$$
$$+ E_{DL} \sum_{\vec{r}} |\vec{b}(\vec{r})|^2.$$

Here, the summation runs over $\vec{r}$, the position of individual atoms, $\vec{b}(\vec{r})$ is the Burgers vector at the site of $\vec{r}$, and couplings K_1 and K_2 are related to the elastic constant. E_{DL} is the core energy related to the dislocation, and a is the radius of the defect. The Hamiltonian has a logarithmic form analogous to the XY model. The second term comes from the vector nature of the Burgers vector and depends on the orientation of the Burgers vector relative to the displacement vector between them, i.e., $(\vec{r} - \vec{r'})$. Hence, compared to the XY system, the Hamiltonian of dislocations is more complex.

While dislocations are characterized by vectors, a scalar number, $s(\vec{r})$, describes disclination defect. Hence, the Hamiltonian of a bound disclination is mathematically equivalent to the vortices interaction

$$H_{DC} = -\frac{\pi K_A}{36} \sum_{\vec{r} \neq \vec{r'}} s(\vec{r}) s(\vec{r'}) \ln\left(\frac{|\vec{r} - \vec{r'}|}{a}\right) + E_{DC} \sum_{\vec{r}} s^2(\vec{r}),$$

where K_A is the Frank constant. Since dislocation defects are energetically equivalent to vortices, analogous to the presented XY system, it concludes that the total disclinicity in a melting crystal has to be zero.

The elastic theory considerations indicate that two dislocations at a separation distance r attract each other to lower the strain energy of the system. At sufficiently high thermal energy, the pair may reach higher states of energy and appear at further distances r, meaning there is a probability to see isolated dislocations at high enough temperatures.

A two-dimensional crystal is known by quasi-long-range translational order and long-range orientational order [47]. The global bond-orientational order, $|\psi_6|$, is unit value, and the bond-orientational correlation function, $g_6(r)$, approaches a constant at large separations, as shown in figure 4.11. Applying the KT-transition to a 2D crys-

tal, i.e., unbinding dislocation at temperature T_m, destroys the quasi-long range translational order. The new phase differs from an isotropic liquid by possessing a quasi-long range translational order. The global bond orientational order varies abruptly to fewer values. In a triangular lattice, this phase is known as the hexatic phase. The long-range translational order changes to the quasi-long-range order characterized by a power-law bond-correlation function of the hexatic phase,

$$g_6(r) \sim r^{-\eta(T)},$$

where $\frac{1}{4} \leq \eta(T_m) \leq \frac{1}{3}$; see figure 4.11. A low density of isolated dislocations screens the logarithmic interaction between bounded disclinations. Hence, applying a subsequent KT transition to the hexatic phase would complete the melting transition. At higher temperatures, the unbinding defects process continues, and a single dislocation can dissociate into two isolated five- and seven-fold particles. These newly produced defects, i.e., disclinations, complete the melting process. At temperature T_i, free disclinations appear in the system to destroy the orientational order. Hence, the system switches to an isotropic liquid known for exponentially decaying in translational and orientational order. Figure 4.11 illustrates the melting process regarding the KTHNY-scenario, where changing the order parameters characterizes different phases.
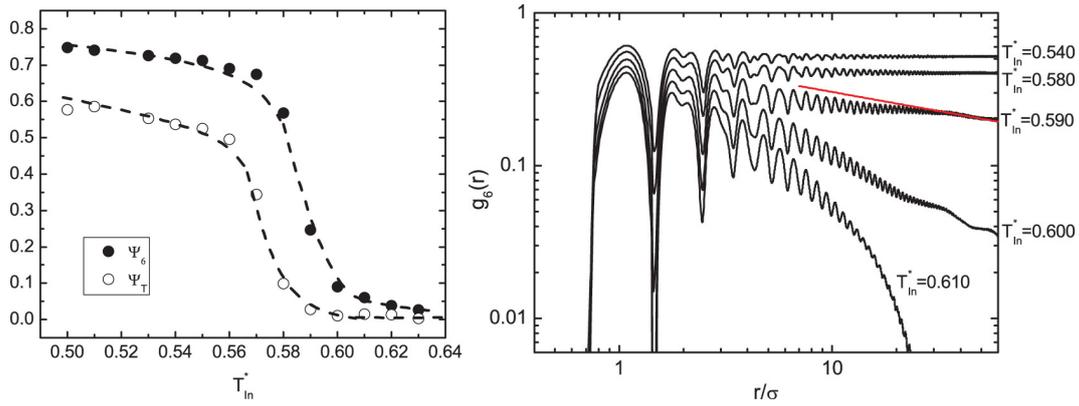
Figure 4.11: Left: global translational order parameter, ψ_T, and global orientational order parameter, $|\psi_6|$. A sharp decrease of ψ_T is a signature of transition to the hexatic or liquid phase. For a solid in a perfect triangular lattice, $|\psi_6| = 1$, i.e., the sixfold orientational symmetry is maintained. Randomly distributed bonds orientation in a liquid leads to $|\psi_6| \approx 0$. Right: the orientational correlation function $g_6(r)$ at different input temperatures. According to the KTHNY theory, $g_6(r)$ is the determinative parameter to distinguish different phases. It approaches a constant in the crystal phase, decays algebraically with a critical exponent $\eta = \frac{1}{4}$ in the hexatic phase and exponentially in the liquid phase. The red line indicates a line with a slope of 3. Hence, the studied system at the underlying temperature reached its hexatic phase. Reprinted from [49], Melting in two-dimensional Yukawa systems: A Brownian dynamics simulation, Copyright (2010), with permission from the American Institute of Physics, License Nr. 5254260817875.

The KTHNY scenario suggests two subsequent KT transitions from a two-dimensional crystal to an isotropic liquid. The first transition converts the crystal to an ordered liquid called the hexatic phase. The touchstone of the hexatic phase is the algebraic decaying of the orientational correlation function. The second transition completes the melting process and provides the isotropic liquid. Topological defects produced by thermal fluctuations in the matter drive these transitions. Hence, both transitions are continuous second-order transitions. Table 1 summarizes the KTHNY scenario of melting.

However, the KTHNY picture of melting, i.e., two subsequent second-order transitions, is one of many possible mechanisms. Another picture, for instance, would be the premature dissociating disclinations before isolated dislocations appear, which result in a first-order transition. An exemplary phase diagram concerning the possibility of either a first-order or KTHNY scenario is shown in figure 4.12.

Phase of Matter	Solid	Hexatic	Liquid
Temperature	$T < T_m$	$T_m < T < T_i$	$T > T_i$
Dislocations	Bound in pairs	Free	Free
Disclinations	Bound in Quarters	Bound in pairs	Free
Positional Correlations	Quasi-Long-Range	Short Range	Short Range
Bond-Orientational Order	Long range	Quasi-Long-Range	Short Range
Elastic Constant	Finite, nonezero	Zero	Zero
Frank Constant	Infinite	Finite, nonezero	Zero

Table 4.1: two-dimensional melting according to the KTHNY-scenario
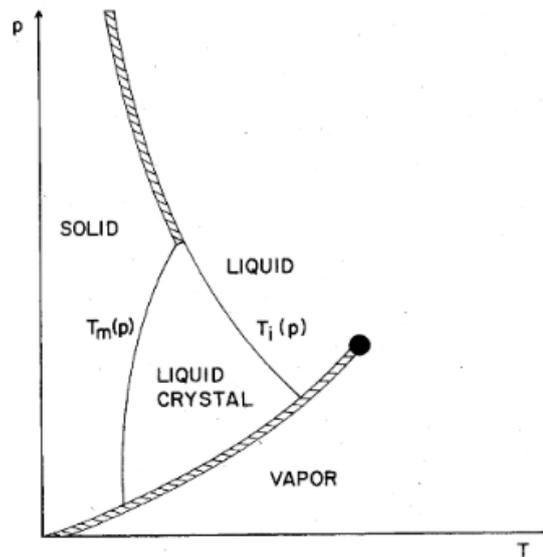
Figure 4.12: The KTHNY scenario proposes a two-stage melting transition involving an intermediated ordered liquid. The first step transforms the solid to a liquid crystal phase characterized by a quasi-long-rang orientational order. The liquid crystal region in the diagram illustrates this phase. The second transition is required to result in an isotropic liquid. T_m and T_i denote the temperatures in which these transitions occur. The hatched lines exhibit the possibility of a first-order transition. Reprinted from [47], Copyright (1979), with permission from American Physical Society.

Several factors, such as the interactions between individual particles, their shape, and the substrate potential, may affect the melting process in atomic systems. Thus, numerical simulations have become an extensive field to investigate and clarify the ambiguities of the melting transition. The last section of this chapter serves the numerical simulations studied two-dimensional systems of hard disks.

4.4 Computational studies

4.4.1 Melting transition in a hard-disk system
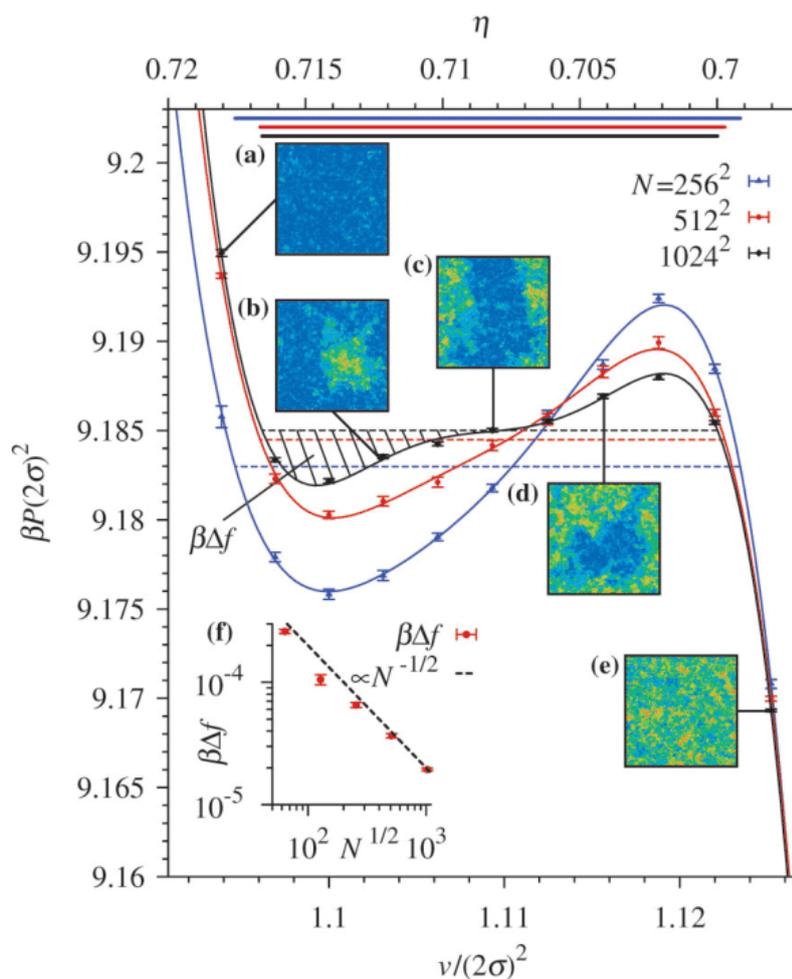
Figure 4.13: The pressure of a hard-disk system driven by the equation of state is plotted vs. density, upper scale, and the volume per disk, lower scale, measured in a NVT ensemble. The equation of state is associated with a loop that comes from the interface free energy effects. The loop tends to be flattened in the larger system, indicating the finite-size effects. The horizontal lines are Maxwell's construction. The hatched area presents the free energy per disk for the larger system size. This area scales as $\frac{1}{\sqrt{N}}$, which provides evidence of a first-order liquid-hexatic transition. Snapshots show two single phases and the coexisting liquid-hexatic phases in between. The color code presents the local orientation, explained in detail in the caption of fig. 4.14. Reprinted from [5], Copyright (2011), with permission from American Physical Society.

A system of hard disks in two dimensions is one of the simplest models of a classical fluid. Besides the simplicity of this system, the nature of the phase transition in this system was unknown for decades. In two dimensions, particles are idealized disks of radius R initialized by their positions and velocities in free space or box. Particles behave freely far from others but obey simple reflection rules during contact with walls or other particles. The following potential introduces such a system

$$V(r) = \begin{cases} 0, & r \geq 2R \\ \infty & r < 2R. \end{cases}$$

Here, r states the radii distance between two disks.

The hard disk system near the melting transition is sensitive to the finite size effects. This sensitivity was the origin of the controversy surrounding this problem.

Bernard et al., see [5], addressed the ambiguity of melting transition using the event chain Monte Carlo algorithm. They studied larger-size systems containing up to $N = 1024^2$ hard disks in the NV ensemble [3]. The visual evidence of coexisting phases obtained by bond orientational order investigation, see fig.4.13. After analyzing the free energy per disk, they concluded that the transition involving the isotropic phase is first-order. The local density averaged over $50R$ exhibits an obvious connection between correlated and high-density regions. This connection provides visual evidence of the liquid-hexatic coexistence, see Fig. 4.14 and table 4.2.

$\rho < 0.700$	$0.700 \leq \rho \leq 0.716$	$0.716 \leq \rho \leq 0.720$	$\rho > 0.700$
Liquid	Coexistence phase	Hexatic	Solid

Table 4.2: Two-dimensional melting transitions in a hard-disk system.

---

[3]For a model with short-range interactions in the thermodynamic limit, the results do not depend on the choice of ensemble.
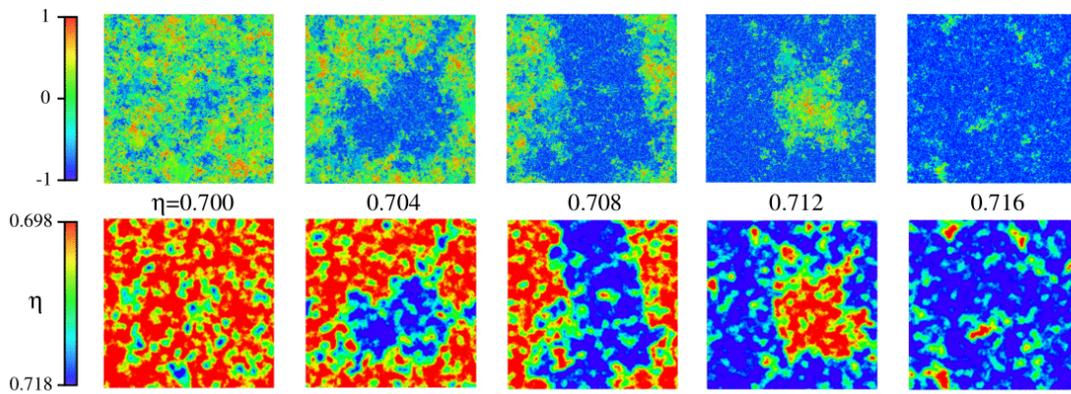
Figure 4.14: Snapshots of color-coded local orientation and the coarse-grained density. Upper snapshots: the complex vector $\Psi_K = \frac{1}{n_l}exp(6i\theta_{kl}) \equiv (\mathrm{Re}\,\Psi_K, \mathrm{Im}\,\Psi_K)$ defines the local bond-orientational order parameter for the disk k, where l stands for the nearest neighbors of k. One can introduce a unit vector $\hat{e}_k$ based on this value. The global orientation of the system is another complex vector defined as $\Psi = \frac{1}{N}\sum_k \Psi_K$. The projection of the local orientation on the sample orientation is applied to the Voronoi construction. The magnitude one, labeled in blue, states that the local and global orientations are aligned. Then, the blue regions present correlated areas. Lower panels show the coarse-grained density. The local density averaged over $50R$ provides visual evidence of coexisting hexatic-liquid phases. An obvious connection exists between correlated and high-density regions in the upper and lower panels. Reprinted from [5], Supplementary Information, Copyright (2011), with permission from American Physical Society.

However, the two-dimensional pair correlation study revealed no doubt that a single phase corresponding to the hexatic exists. The KTHNY theory predicts a power-law function for the orientational order parameter with the exponent $-\frac{1}{4}$ in the hexatic phase. Bernard et al. obtained a negative value for the exponent. However, the exponent was close to zero and far from the expected value by the KTHNY scenario. They concluded that the first-order liquid-hexatic transition might affect the hexatic phase.

In conclusion, because of the appearance of the hexatic phase, the two-dimensional melting can not be classified as a pure first-order transition. Therefore, these observations deviate from the predictions of the KTHNY theory. Bernard et al. concluded that in a hard-disk system, the hexatic phase acts as a mediator for the melting transition. Notably, the liquid-hexatic transition is first-order, while the hexatic-to-solid transition is continuous, presenting a unique behavior.

4.4.2 Melting transition in a soft-disk system
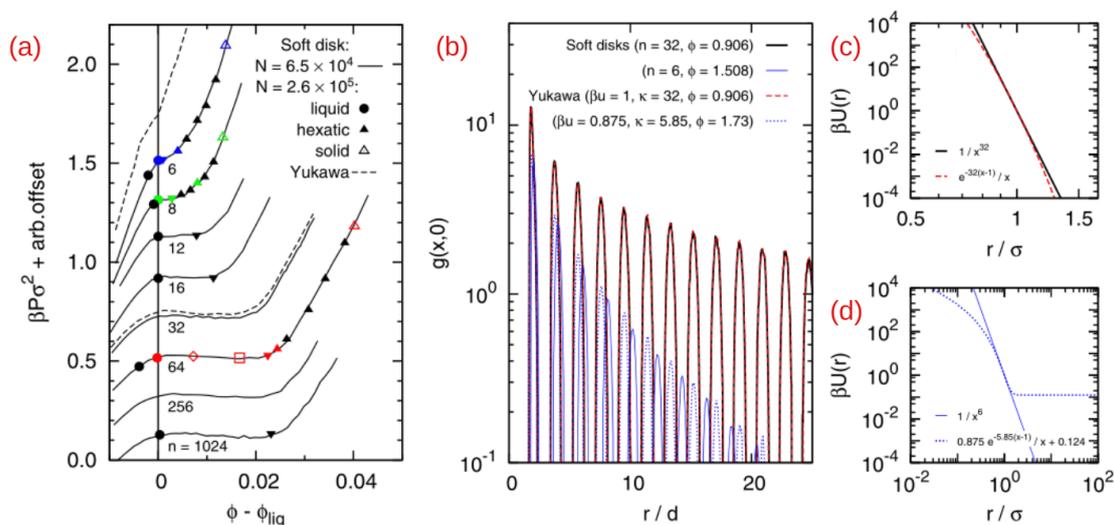
Figure 4.15: (a) Equation of state for a family of repulsive potentials $U(r) = \epsilon \left(\frac{R}{r} \right)^n$. The Mayer-Wood loop, which occurs for $n > 12$, is evidence of a first-order transition. The coexisting interval varies non-monotonically. $\phi = \frac{\sigma^2 N}{V}$ is density, and ϕ_{liq} denotes the liquid density at coexistence. At $n \preceq 6$, the first-order scenario changes to a continuous transition. **(b)** Radial distribution function along the x-axis. Depending on the interaction strength and Debye-Hückel screening length, a Yukawa potential may exhibit a first-order or continuous liquid-hexatic transition. **(c-d)** Interaction potential of the soft disks and the Yukawa particles around $r = \sigma$. Reprinted from [27], Copyright (2015), with permission from the American Physical Society.

A study of hard disks interacting via soft-repulsive potential $U(r) = \epsilon \left(\frac{R}{r} \right)^n$, established the two-step transition including the hexatic phase [27]. Columb forces and dipole interactions with $n = 3$ are a few examples of such interactions. However, depending on parameter n, continuous and first-order liquid-hexatic transitions were observed.

A system with a larger n recovered the hard-disk scenario with the first-order hexatic-liquid transition. The hexatic-liquid coexistence width varies non-monotonically. The coexisting interval vanishes at $n = 6$. The systems with $n \preceq 6$ showed the two-step continuous transition, see Fig.4.15-a.

Having both continuous and first-order liquid-hexatic transition led the authors to a conclusive hypothesis: each particle explores a small part of interactions, i.e., any interacting system in which the interactions match the soft-disk potential in the range

of inter-particle distances, follows the soft-disk scenario, which was confirmed by using a Yukawa interaction $U(r) = u \left(\frac{2R}{r} \right) exp^{\left[\kappa \left(1 - \frac{r}{2R} \right) \right]}$, see figure 4.15-(b-d). This study suggests that tuning Debye-Hückel screening length will result in both first-order and liquid-hexatic scenarios.

Chapter 5

Adsorption kinetics

Class II hydrophobins are amphiphilic proteins produced by filamentous fungi. The adsorption kinetics measured for the wild-type proteins HFBI and HFBII are astonishingly robust. Several experiments have been performed but still show strongly consistent features; the adsorption kinetics for self-assembly at the air-water interface follows linear kinetics up to saturation. This feature distinguishes these proteins from, e.g., Lysozyme, in which the adsorption kinetics follow the conventional Langmuir adsorption model. This chapter will investigate the origin of this behaviour by increasing knowledge of the dynamic behaviour of these proteins at air-water interfaces.

To this end, I present the experiments the Jacobs group has performed at Saarland University. Those experiments were carried out on wild-type proteins HFBI and HFBII and a few selected mutants of HFBI. These observations will set the subsequently developed model into context.

In the second section, I introduce the stochastic model, which studies the dynamic behaviour of class II hydrophobins.

The last section states the simulation results and discussion.

This chapter is based on the article **Dynamic assembly of class II hydrophobins from T. reesei at the air-water interface** published in American Chemical Society, Langmuir, 2019, 35, 9202-9212 by **Hendik Hähl** *et al.* [20].

5.1 Experimental section

The hydrophobins HFBI, HFBII, HFBI-dCBM, HFBI-D30N-K32Q, and FpHYD5, introduced in section 2.1.6, were the subjects of our investigation. We used non-destructive ellipsometry measurements to quantify the protein adsorption kinetics at the air-water interface [20]. This method, which is based on the refractive index variation and provides the adsorbent's dry mass, was instrumental in our research. The following subsection presents the experimental results and discussion of these measurements. The adsorption experiments, coupled with atomic Force Microscopy (AFM) inspection, offered unique insights into the structure of the film. This method allowed us to probe both the hydrophilic and hydrophobic sides of the film, a distinction that is crucial for our conclusions. Therefore, I will elaborate on the procedure of these measurements at the outset of section 5.1.2.

5.1.1 Adsorption kinetics

Hydrophobins are self-assembled at air-water interfaces due to their amphiphilic structure. The adsorption kinetics for wild-type proteins HFBI and HFBII for protein concentrations of 0.1, 0.2, 0.5, 1, and 2 μM are depicted in figure 5.1. As shown in the figure, the adsorbed protein increases linearly until saturation. The exception is concentration 0.1 μM, where the saturation was not achieved in the experimental time windows.

As shown in figure 5.1, all measurements show deviations from the exponential Langmuir kinetics; the adsorption rate, defined as the average mass of protein adsorbed at the interface in the unit of time, is constant up to saturation. In particular, for the wild-type proteins HFBI and HFBII, the adsorption rate is independent of the already adsorbed amount. The linear adsorption kinetics contrasts with the usually observed Langmuir kinetics in which the adsorption rate decreases as the adsorbed mass increases, see section 2.2.6.
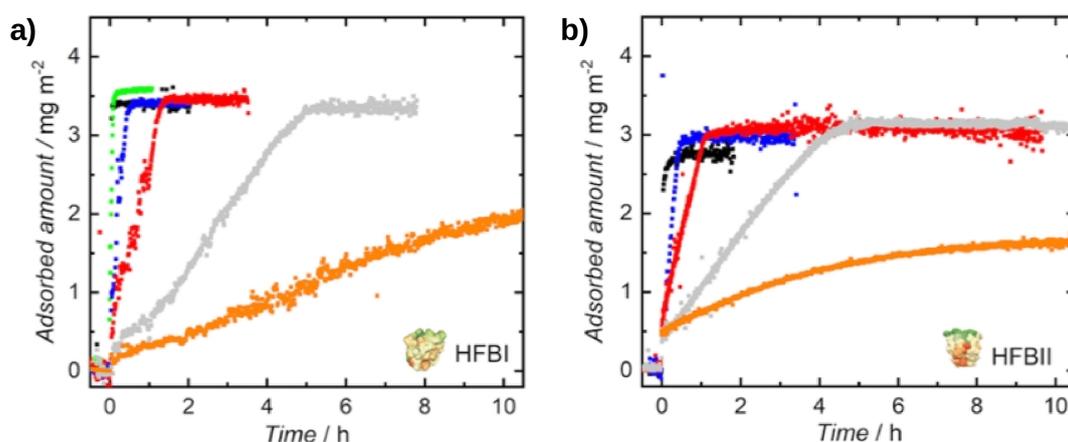
Figure 5.1: Wild-type hydrophobins adsorption kinetics; The adsorbed amount of **(a)** HFBI, and **(b)** HFBII as a function of time. The dynamics of the protein assembly were measured at concentrations of 0.1 (orange), 0.2 (gray), 0.5 (red), 1.0 (blue), 2.0 (green), and 5.0 (black) μM. The adsorbed protein increases linearly in time until saturation. The saturation is not achieved for the concentration of 0.1 μM in the experimental time windows. The figure is reprinted from **Hähl *et al.* [20]**, with the permission of **American Chemical Society**.

While the adsorption rate depends on the concentration, i.e., a higher concentration results in quicker film formation and a higher adsorption rate, the saturation value does not; it reaches a constant value independent of the protein's concentration; the final recorded mass at the interface is 3.4 and 3.0 $mg\ m^{-2}$ for HFBI and HFBII, respectively. As I will discuss later, those values correspond to the formation of a monolayer at the interface.

Figure 5.2 sketch the adsorption kinetics of HFBI as a reference and three mutations of it, HFBI-D30N-K32Q, HFBI-dCBM, and FpHYD5, for different concentrations. For the mutant HFBI-D30N-K32Q, for large concentrations, the kinetics is similar to HFBI, i.e., the kinetics is linear. However, the adsorption rate is larger compared to HFBI. For a lower concentration of 0.2 μM, The adsorption rate changes gradually from a constant none-zero to zero. Remarkably, HFBI-D30N-K32-Q reaches saturation even for the lowest concentration of 0.1 μM in the time window of the experiment. The adsorption rate of the lowest concentration shows time dependency.

For the mutant HFBI-D30N-K32Q, for large concentrations, the kinetics is similar to HFBI, i.e., the kinetics is linear. However, the adsorption rate is more significant compared to HFBI. Here, two exceptions meet; the adsorption rate changes gradually

from a constant none-zero to zero for a lower concentration of 0.2 μM. Remarkably, HFBI-D30N-K32-Q reaches saturation even for the lowest concentration of 0.1 μM in the time window of the experiment; moreover, the adsorption rate of the lowest concentration shows time dependency.
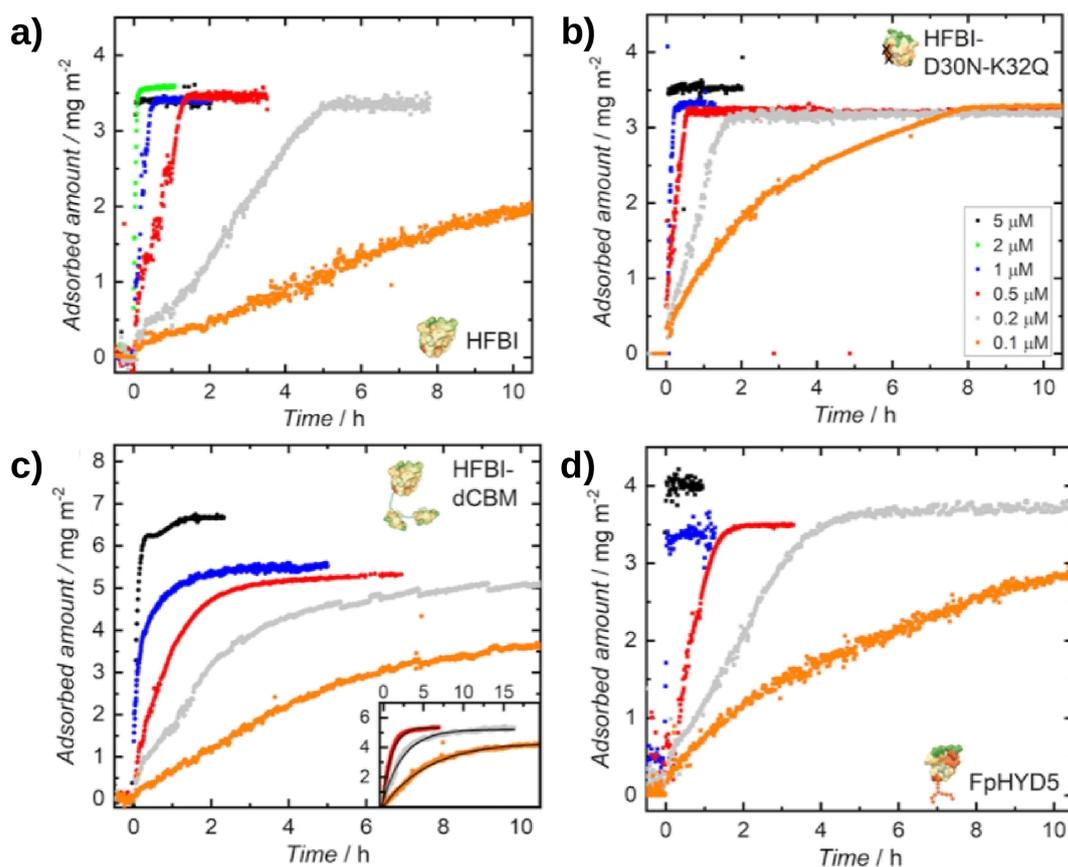
Figure 5.2: Adsorption kinetics; The adsorbed amount of **(b)** HFBI-D30N-K32Q, **(c)** HFBI-dCBM, **(d)** FpHYD5 as a function of time. The adsorption kinetics of HFBI are shown as a reference in **(a)**. The dynamics of the protein assembly were measured at concentrations of 0.1 (orange), 0.2 (gray), 0.5 (red), 1.0 (blue), 2.0 (green), and 5.0 (black) μM. The inset of (c) shows that the data fit the first-order kinetics model of the experiments in the 0.1-0.5 μM. The figure and caption are reprinted from **Hähl _et al._ [20]**, with the permission of **American Chemical Society**.

The adsorption kinetics for mutant HFBI-dCBM, shown in panel-c, completely differ from wild-type hydrophobins. The kinetics does not exhibit a constant adsorption rate. In contrast, the adsorption rates in all probed concentrations are perfectly described by a Langmuir adsorption kinetics in which the adsorption rate decays exponentially, as shown in the inset of this panel. Moreover, HFBI-dCBM displays a higher

saturation value at a concentration of 5 μM.

The adsorption kinetics of FpHYD5, shown in panel-d, is laid between purely linear wild types and the pure exponential of HFBI-dCBM. The adsorption kinetics of the lowest concentration follows the Langmuir kinetics. In contrast, the adsorption rate for higher concentrations is constant, differing from wild-type profiles by gradually decreasing to zero rates upon reaching the saturation value.

The film formed at the interface is likely monolayer; the formation of a multilayer is not suspected according to the following calculations performed for HFBII.

The mean saturated adsorbed mass of HFBII is 3.0 $mg\ m^{-2}$. The molecular weight of HFBII is 7.2 $k\ Da$. Hence, the adsorbed mass in the unit of mole is

$$\text{Adsorbed mass} = \frac{3.0 \times 10{-3}\ [\text{gr}\ m^{-2}]}{(6.02 \times 10^{23})\,(7.2 \times 103)\,(1.66 \times 10{-24}\ [\text{gr}])}$$
$$= 0.41 \times 10^{-6}\ [mol\ m^{-2}]\,.$$

Which gives the mean area per protein of $\frac{1}{(0.41 \times 10^{-6})(6.02 \times 10^{23})} = 4.0\,[nm^{-2}]$. This experimentally calculated value is comparable to the measured amount of $3.2\ nm^{-2}$ [22]. Therefore, the formed film is likely a monolayer.

Similarly, the mean area per protein for other examined proteins is calculated and presented in table 5.1. While almost all mutants occupy a similar area per protein, the mean area of bulky mutant HFBI-dCBM is significantly larger, which is suspected to be due to the role of the attached domains.

protein	$\Gamma\ [\mu M\ m^{-2}]$	area per protein $[nm^{-2}]$
HFBI	0.45 ± 0.01	3.7 ± 0.1
HFBII	0.41 ± 0.02	4.0 ± 0.1
FpHYD5	0.40 ± 0.03	4.2 ± 0.3
HFBI-dCBM	0.27 ± 0.02	6.1 ± 0.5
HFBI-D30N-K32Q	0.44 ± 0.02	3.8 ± 0.1

Table 5.1: Saturation value of the adsorbed amount of used protein. The table is adapted from **Hähl *et al.* [20]**, with the permission of **American Chemical Society**.

The only difference between HFBI and its variant HFBI-D30N-K32Q are the charged amino acids exchanged by the neutral ones, which suggests that Coulomb interactions play a crucial role in self-aggregation. The role of electrostatic interactions is investigated by varying the ionic strength of the buffer. The ionic strength is adjusted

by adding Sodium Chloride to the solution. As shown in figure 5.3 for HFBII, the ionic strength variation does not change the saturation value; however, it does modify the adsorption rate; the higher ionic strength results in a faster-saturated interface. In addition, the adsorption rate deviates from a constant to a decaying amount.
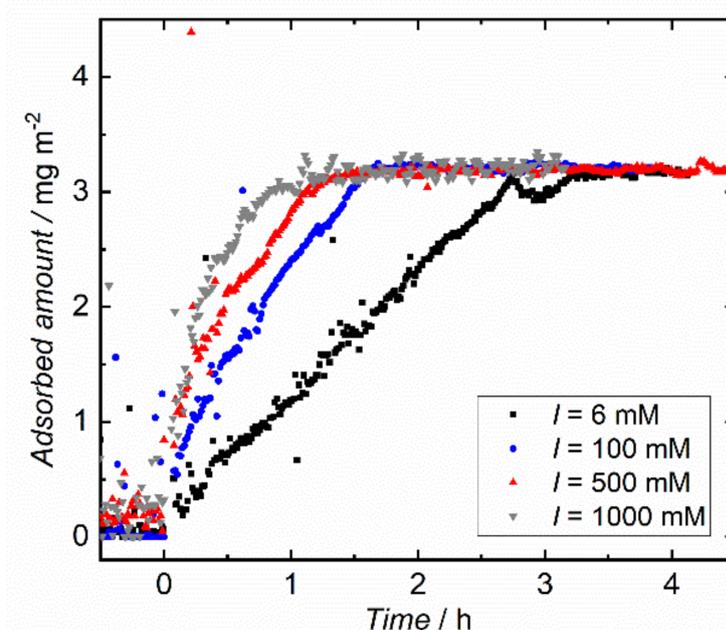
Figure 5.3: HFBII (concentration $= 0.2\,\mu M$) adsorption kinetics at the ionic strength of 6 (gray triangles), 100 (red triangles), 500 (blue dots), and 100 mM (black squares). Increasing ionic strength converts the kinetics from unusual linear to the conventional Langmuir adsorption kinetics. The figure is reprinted from **Hähl *et al.* [20]**, with the permission of **American Chemical Society**.

To summarize, the distinct feature of wild-type assembly is the linear adsorption kinetics until saturation is reached. In the case of wild-type hydrophobins, the maximum adsorbed amount is independent of the bulk protein concentration as is observed in other studies, e.g., [59, 69]. Hydrophobins gain energy upon the adsorption to the interface; moreover, these proteins do not change their conformation. These facts explain the irreversible adsorption with maximum saturation. However, the adsorption of proteins typically follows Langmuir kinetics, in which the adsorption rate decays with the contraction of the available vacancies at the interface. Without more knowledge of the underlying mechanism, a cooperative adsorption process may explain the constant adsorption rate in which an already adsorbed protein guides the coming one to a free space. I introduce a model that describes the mechanism to gain a deep knowledge of the mechanism underlying

the adsorption process. The model shows that a high diffusive motion next to the interface and a relatively high enough adsorption to the interface are responsible for the observed linear adsorption.

5.1.2 Structure of the interface film
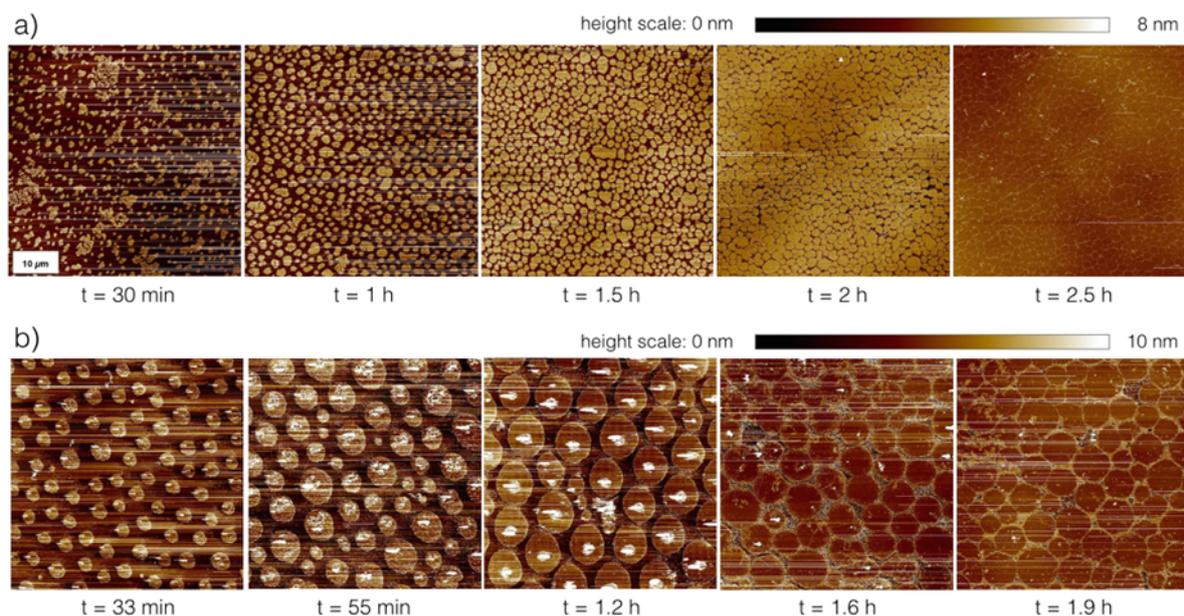
Figure 5.4: AFM height images of HFBI films from the air-water interface. The films were transferred to a hydrophobic solid substrate. Protein concentration was $0.2 \quad \mu M$, ionic strength was **(a)** $6 \quad mM$ and **(b)** $100 \quad mM$. The time after protein injection is given below the images. Images in **(b)** have the same lateral scale. Increasing the ionic strength speeds up the adsorption rate; hence, identical coverage was chosen to compare the two experiments of different ionic strengths rather than identical time points. At the same coverage of adsorbent, increasing the ionic strength led to less numerous clusters but in a larger size. The figure is reprinted from **Hähl *et al.* [20]**, with the permission of **American Chemical Society**.

The structure was probed during an adsorption course to comprehensively study the formed film, meaning that the film's hydrophilic and hydrophobic faces were examined using Atomic force microscopy (AFM). To this end, samples are prepared and transferred to a solid substrate based on the Langmuir- Schaefer (LS) method, in which the substrate moves vertically toward the interface and subsequently lifts off.

As a substrate, a piece of a silicon (Si) wafer was used to image the hydrophilic face

of the film. The substrate was hydrophobized with a self-assembled monolayer of octadecyl trichlorosilane in advanced [35]. The substrate contacts the film from above immediately after being removed; then, it is raised until it fully detaches from the interface. In the last step, the sides of the silicon wafer were approached with an absorbent paper to drain off the residual water on the film.

The hydrophobic face of the film is examined using a hydrophobized mica as a substrate. This method is more invasive than the usual Langmuir-Schaefer method described above; the film is introduced into the solution to bring the hydrophilic side in contact with the substrate.

The criterion for comparing samples is the adsorbed amount at the interface rather than the same time points. For HFBI without Sodium Chloride at a concentration of 0.2 μM, the transfer and imaging were performed every 30 minutes (see figure 5.4-a). Since the absorbed amount depends on the ionic strength and at a higher ionic strength, the adsorption rate increases, the measurement time point was adjusted to represent the same amount of adsorbed material, as shown in figure 5.4-b.

The most particular feature observed is the formation of two-dimensional grain boundaries; for all ionic concentrations, two-dimensional grain boundaries are visible. The individual domains, which, from here on, I call clusters, grow until they get in touch. Cluster domains do not fuse even if they touch others. This remarkable feature suggests that the proteins possess a preferred orientation in each domain, which differs from domain to domain.

In the previous section, I presented the idea that the ionic strength speeds up the adsorption. As it is evident in figure 5.4, it also controls the formation of clusters; at an identical surface coverage, a higher ionic strength results in less numerous domains but significantly larger ones.

The observed solid material in the center of clusters in figure 5.4-b is most likely a crystal of Sodium Chloride, meaning that samples are transferred to a highly hydrophobized substrate. It is plausible that samples trap a droplet of the solution during the transformation. Since even the hydrophobic face of the film has less hydrophobicity than the surrounding environment, the drop of buffer stays on the

protein layer. This drop contains NaCl. Since the samples are not washed up before imaging, after water evaporation, crystals of NaCl remain in the middle of the clusters.

It was questionable whether cluster formation is spontaneous through film formation or if water evaporation induces cluster formation. To address this, the formed film was imagined after 22 hours; the two-dimensional clusters with well-defined borders were found. This evidence ensures that the cluster formation is a spontaneous process, as shown in figure 5.5-a.
The examination of the hydrophobic face, where the protein domains are also observed, is a complex and invasive process. As illustrated in the inset of figure 5.5-b, the substrate is carefully introduced into the solution to bring the hydrophilic side in contact. This is followed by a vertical lift of the substrate, resulting in the formation of a film. The intricate nature of this procedure leads to a heavily distributed interface.
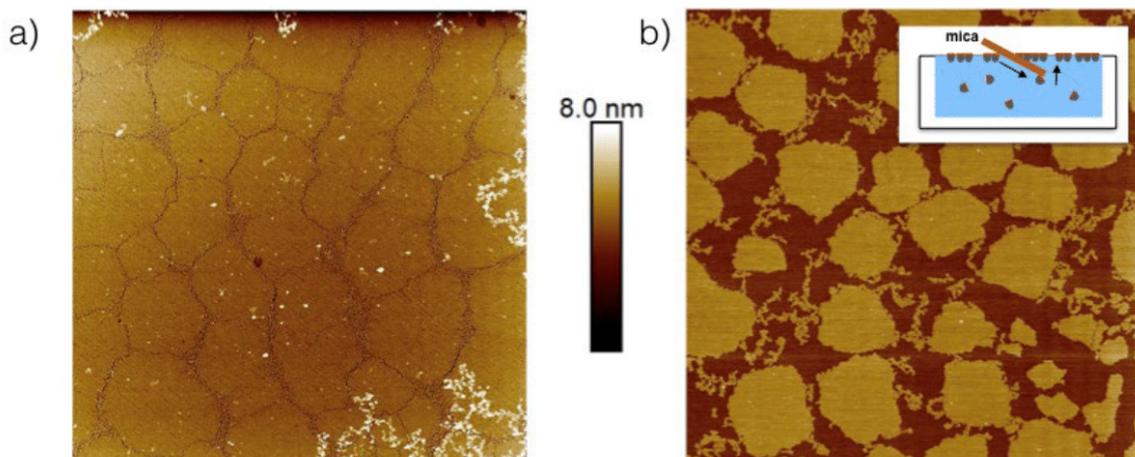
Figure 5.5: AFM images (size: $10 \ \mu m \times ten \ \mu m$) of Langmuir-Schaefer films of HFBI more than 22h after protein injection into the solution ($c = 0.2 \ \mu M$). **(a)** The film was prepared as the films shown in figure 5.4. A completely saturated film is observed, still showing boundaries between individual clusters. **(b)** This film was prepared on a freshly cleaved mica sheet, which was first introduced in the solution and then lifted to bring the hydrophilic parts of the hydrophobins in contact with the mica (see inset sketch). Due to this procedure, the interface is heavily disturbed, and clusters are separated. Compared to the images in figure 5.4, the hydrophobic and hydrophilic regions are inverted: Here, the cluster surfaces are hydrophobic, whereas the substrate around the clusters is hydrophilic. The figure is reprinted from **Hähl *et al.* [20]**, with the permission of **American Chemical Society**.

In contrast to the cluster formation of WTs, the AFM examination of variant HFBI-dCBM shows, rather than protein domains, a homogeneous film (see figure 5.6). The

mutant HFBI-dCBM has 18.5 $k\,Da$ molecular weight compared to the WT protein; it has 11.5 $k\,Da$ molecular weight and increased size, which reduces its translational and rotational diffusion at the interface. A comparable high molecular weight, in addition to the large size, influences its mobility at the interface. Hence, it is likely that neighbouring proteins are not able to rotate and match their directionality, resulting in a homogeneous film.
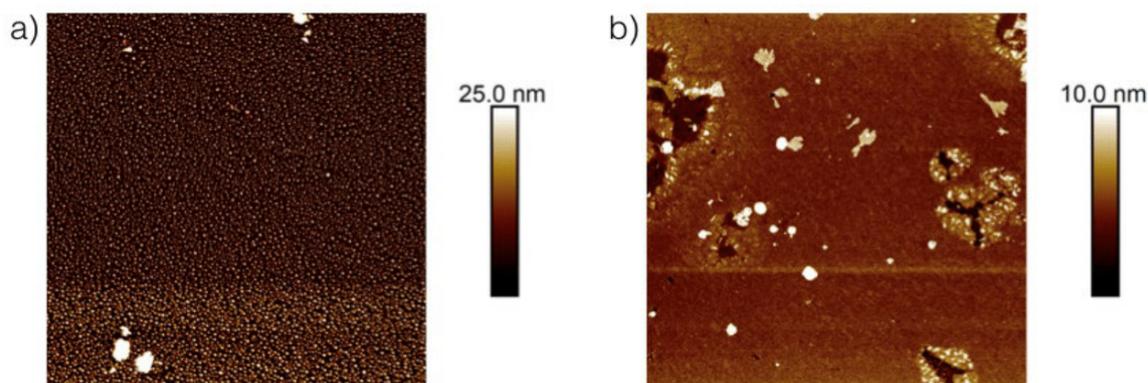
Figure 5.6: AFM images (size: $10\quad\mu m \times 10\quad\mu m$) of Langmuir-Schaefer films of HFBI-dCBM **(a)** ca. 15 min. and b) 10 h after the start of the adsorption measurements. The films were prepared as the HFBI films presented in figure 5.4. In contrast to the HFBI films, no cluster growth can be observed. In **(b)**, agglomerations of proteins are likely formed during the drying process. The figure is reprinted from **Hähl *et al.* [20]**, with the permission of **American Chemical Society**.

5.1.3 Summary

Protein adsorption to the air-water interface usually follows the Langmuir adsorption kinetics, where the effective adsorption rate depends on the vacancies in the film. In contrast, the hydrophobins HFBI and HFBII show, up to saturation, a constant adsorption rate. As revealed by AFM measurements, the adsorption process is accompanied by the formation of stable grain boundaries.

The origin of the observed structure and unusual linear dynamics has been investigated by changing the ionic strength of the solution, as well as by considering modified proteins.

The mutation affects the kinetics and cluster formation. The bulky mutant HFBI-dCBM adsorption kinetics is well-fitted by the Langmuir adsorption kinetics. Moreover, the grain boundaries have not been observed for both weighted mutants HFBI-dCBM and FpHYD5.

Experimental observations ensure that the electrostatic interactions play a role in

both observed phenomena; screening electrostatic interactions by increasing ionic strength or mutation changes the linear kinetics to the conventional Langmuir adsorption kinetics. The increment in the ionic strength also speeds up the effective adsorption rate and controls the size and number of grain boundaries.

Without more profound knowledge, cooperative behaviour controls the assembly. In the following section, I aim to light up the ambiguities in this issue through a theoretical study with the help of introducing a stochastic model.

5.2 Model

Class II hydrophobins are small globular-solvable proteins that, in a solution, may form a multimer complex or be found as monomers depending on the concentration. Hydrophobins are surface-active proteins that self-assemble in the air-water interfaces to shield their hydrophobic parts, as shown in figure 2.12. Conformational changes are not expected to occur during the adsorption or at interfaces.

Hydrophobins can form a complete crystalline monolayer at the air-water interface. Experimental observations confirmed long-ranged grain boundaries that are growing over time without fusing. The grain boundaries show a crystal structure comparable to a two-dimensional honeycomb lattice.

Next to the features mentioned above, the adsorption kinetics of class II hydrophobins HFBI and HFBII show a remarkable feature: the adsorption kinetics follow a linear kinetics up to saturation. This aspect has distinguished these proteins from, e.g., lysozyme, in which the adsorption rate shows dependency on vacancies.

The main goal is to develop a simplified model system to expose the underlying mechanisms of the unusual hydrophobin's adsorption kinetics. The model is restricted to the most elementary processes where tuning their relevance allows for in-depth knowledge of their contributions to the observed linear kinetics.

I model the adsorption as a lattice gas model in two dimensions, not including details of protein-protein and protein-surface interactions. Those calculations, even if possible, are usually time-consuming and are only feasible for part of the systems. Specifically, the model is constructed to limit the number of free parameters. However, to model the mechanical response of the system under external stress, as discussed in chapter 6, I will introduce an off-lattice approach.

Experimental results with mutants and variable ionic strength suggest the relevance of long-range electrostatic interactions. In addition to the van der Walls interactions between proteins in a layer, the possibility of an energy barrier and a second minimum in the surface potential energy leads to modelling the reality as a bilayer system in which the DLVO theory describes the interactions between particles.

The adsorption kinetics is modeled by Monte Carlo dynamics based on the Metropolis algorithms. Monte Carlo methods allow us to characterize the relative time scales

of stochastical processes. The connection to the real-time scales can be taken, e.g., from the initial slope of the adsorption kinetics. The stochastic approach's inherent noise avoids introducing any artificial deterministic parameters. The whole dynamics are governed by the choice of rates for elementary processes and their acceptance probabilities, which are given by the defined particle-particle interactions. The selection of these processes is implemented through tower sampling.

The following subsections describe the model and used methods in detail.

5.2.1 Model: Experimental setup
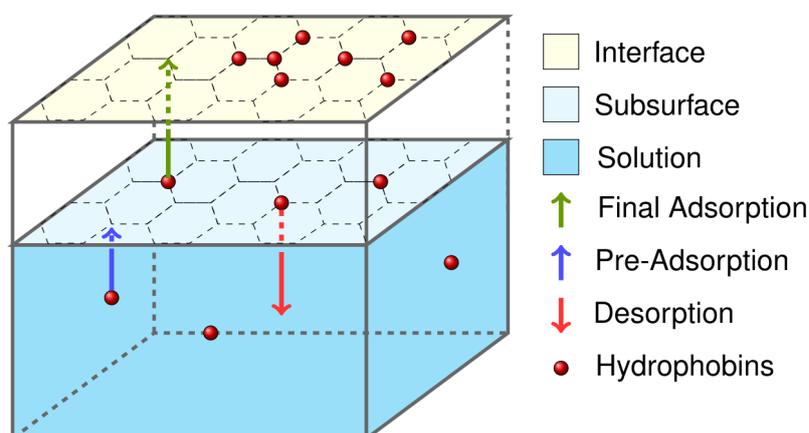
Model structure

Figure 5.7: Basic model structure. Top: air-water interface. The protein solution is not considered explicitly; instead, I introduce a subsurface layer below the interface and the solution as bulk. Honeycomb lattices represent both subsurface and interface layers. The pre-adsorption to the subsurface is reversible, while the final adsorption to the interface is not. The protein concentration in the solution dictates the adsorption rates to the subsurface layer and interface. The figure is adapted from **Hähl** *et al.* **[20]**, with minor changes, with the permission of **American Chemical Society**.

The experimental setup consisting of a protein solution and the air-water interface is modeled as a system of three parts: the interface, subsurface, and solution, as shown in Fig. 5.7. The upper layer models the liquid-air interface. The subsurface layer is just close to the interface. A honeycomb lattice is used for both the interface and subsurface in line with the observed crystal structure for wild-type hydrophobins.

The solution in figure 5.7 represents the bulk. Since multimerization is not expected for the given experimental conditions, the bulk solution is assumed to consist of uniformly distributed monomers of hydrophobins. An effective particle flux from the bulk solution to the interface allows adsorption.

The Langmuir adsorption kinetics describes particles as ideal gas molecules adsorbing to a lattice layer, see 2.2.6. Particles encounter lattice sites randomly. Experiencing an already occupied site is punished by rejection. Similar to the Langmuir adsorption kinetics, particle-particle exclusion is included. However, the model introduced here differs in several aspects from the Langmuir adsorption kinetics: Introducing the subsurface is the first extension of the Langmuir adsorption kinetics.

Moreover, adsorption takes place in two steps. First, a particle enters randomly to the subsurface. The pre-adsorbed particle returns to the bulk or reaches the interface via final adsorption. According to the experimental observations, the final adsorption is an irreversible process. Further extensions of the Langmuir adsorption kinetics are the diffusive motion and the interactions between particles in each layer, which I introduce in the following.

5.2.2 Model: hydrophobins and interactions

The sizeable hydrophobic patch of individual hydrophobins defines a preferred axis oriented perpendicular to the interface after adsorption such that the hydrophobic patch faces towards the liquid. Furthermore, anisotropy in the interface is caused by the location of charged amino acids inside the protein. Hence, the interactions between the adsorbed proteins at the interface should depend on their relatively in-plane orientations. Moreover, structural imaging of fixated films and molecular dynamic docking simulations of several monomers suggest the formation of multimers such as trimers and hexamers at the interface for HFBII [42]. This local in-plane ordering is supposed to be the dominant reason for the stable clusters observed in wild-type hydrophobins.

Those characteristics of the experimental system led us to model proteins as particles with an internal degree of freedom, whose property represents the relative in-plane orientation of proteins. To care for this anisotropy in the simulations, I employ a spinlike multistate model comparable to Pott's model, presented in 2.2.7. Therefore, each protein is defined by a discretized orientation called "spin"

as an internal degree. The spin state is restricted to six different discrete numbers according to the coordination number of the honeycomb lattice. The internal degree defines an anisotropic interaction between proteins. In an ideal configuration, proteins in a cluster share the same spin orientation.

Further internal states, such as conformational changes, are negligible due to the high number of disulfide bridges and hydrogen bonds in class II hydrophobins.

Here, two parts of the particle-particle interactions are contributed: isotropic and anisotropic interactions. The layers' proteins interact by van der Waals, Coulomb interactions, and steric repulsion. This part is described by the DLVO energy, discussed in 2.2.1, with a cut-off radius.

In addition to the isotropic part, the model incorporates the protein orientation by including an energetic benefit to the configurations in which neighbouring hydrophobins i and j share an identical spin. The anisotropic contribution stats

$$U_{ij}(spin) = \begin{cases} -J_0, & \text{if } S_i = S_j \\ 0, & \text{otherwise}, \end{cases}$$

where S_i is the orientation of proteins, and J_0 is a coupling constant that we adjust to a physically reasonable value consistent with the experimental observations.

5.2.3 Model: bilayer model and molecular processes

Since there is little to nothing known about the surface potential of the air-water interface regarding hydrophobins or their protein-protein interactions, I chose a rate-based statistical model with the simplicity of underlying processes as a critical concern.

In Langmuir model kinetics, the initial adsorption rate is restricted only by diffusive motion. The accessibility of free sites also limits the adsorption rate at later stages. A dynamic effect can exert an influence on this fact.

As a dynamic effect, the effective concentration of proteins next to the interface does not necessarily equal the protein concentration in bulk. Typical surface potential allows a secondary minimum next to the interface with a resulting energy barrier that has to be crossed to adsorb to the interface. The secondary minimum could lead to the buildup of a particle reservoir that provides a steady particle flux to the interface. The continuous flux supplies the adsorption at a constant speed. Hence, the availability of free sites in the interface does not restrict the adsorption rate, and

the inverse of the flux rate determines the adsorption rate. This dynamic effect is the main reason for considering a bilayer model. Hence, the adsorption is a two-step process. Each of them takes place at a defined transition rate. These transition rates have the unit of the inverse of time and determine the effective acceptance probability of each event.

A protein undergoes at least two processes to adsorb to the interface. First, it appears at a random site in the subsurface. This process is called the pre-adsorption process and happens by the pre-adsorption transition rate of μ_A. The protein concentration in bulk determines the pre-adsorption transition rate. However, the relativity between pre-adsorption probability and protein concentration is non-linear. The adsorption process is completed if the pre-adsorbed protein disappears in the subsurface and appears at the same position at the interface. The latter process, called final adsorption, takes place with the final adsorption transition rate of λ_A. It is experimentally evident that the final adsorption is irreversible; however, the pre-adsorption process is not. A protein in the subsurface has a chance to return to the bulk. This event defines the desorption process, which is possible by the desorption transition rate of μ_d.
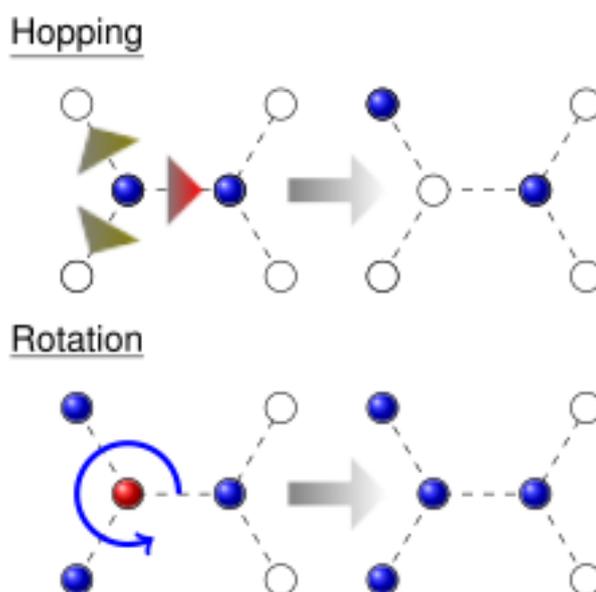
Figure 5.8: Illustration of intra-layer particle dynamics. **(Hopping)** The hopping process describes diffusive motion in the subsurface and interface layers. A particle is chosen randomly and jumps to one of its nearest neighbouring sites. The exclusion is included in the hopping process; hopping to an occupied site is forbidden. The left figure shows the configuration before the process. The green triangles show the allowed position to jump. Jumping to the site shown by the red triangle is not allowed due to exclusion. The right-hand side figure shows a possible configuration resulting from the acceptance of the hopping process. **(Rotation)** Each protein has a specific orientation as an internal degree of freedom. Different colors show different orientations. A configuration in which all particles point in an identical direction is energetically favored. Through the rotation process, a particle with a different orientation has the probability of orienting and matching its direction with neighbouring particles. The left-hand side figure shows a configuration in which the orientation of particles is not similar. The right-hand side figure shows the configuration after the rotation process. The figure is reprinted from **Hähl *et al.* [20]**, with the permission of **American Chemical Society**.

Diffusive motion in both layers dynamically influences the adsorption. The diffusive motion is hopping to one of the nearest neighbouring sites by a specified probability of μ_M and λ_M for diffusion in the subsurface and interface, respectively. The diffusive motion in the subsurface extends the searching time for a pre-adsorbed particle to complete adsorption. Suppose the final adsorption is rejected due to overlapping in the interface; in that case, the diffusive motion enables the pre-adsorbed particle to hop to a neighbouring site and try to complete the adsorption. As I will state in the following subsection, the hopping and final adsorption are the critical processes determining the kinetics. Moreover, I will show that the interface's diffusive motion

is essential to having regular roundish clusters, too.

Finally, a rotation process is possible in the interface. Two neighbouring proteins of identical spin are energetically favoured. Hence, a probability of rotation for proteins is considered to match the orientation of neighbouring proteins. The rotation process occurs only in the interface layer with the probability of λ_R.

Figure 5.8 illustrates the diffusive motion and rotation. A brief description of simulation parameters is summarized in table 5.2.

subsurface	interface	description
μ_A	λ_A	a site is chosen randomly and a particle is adsorbed
μ_d	-	one particle in the subsurface layer is chosen randomly and desorbed to the solution
μ_M	λ_M	one particle is chosen randomly, it moves randomly to one of its nearest empty neighbors
-	λ_R	one particle is chosen randomly and rotated

Table 5.2: Simulation parameters

5.3 Results and discussion

5.3.1 Results: determining the reference parameters set

The easiest way to understand how this model results in linear kinetics is to use the diffusive-sorting process in adsorption; the density of particles and, therefore, the number of free sites at the interface are time-dependent. If the diffusive motion in the subsurface and the transition to the interface take place fast enough compared to the subsurface supply, the accessibility of vacancies at the interface will not control the adsorption rate; thus, the effective adsorption rate remains constant, and the supply rate to the subsurface determines it.

A high trail transition rate from subsurface to interface ensures that a free site in the interface is readily occupied once found by a diffusive particle in the subsurface. Thus, the only requirement for linear kinetics is to have diffusive particles in the subsurface, which sweeps a large enough area between two subsequent transitions to the interface. As soon as the diffusive area covers at least one free site in the interface, the diffusive particle will occupy the free site, and the adsorption rate will be

dominated by the supply rate of particles to the subsurface. To achieve this, a low particle density in the subsurface is necessary.

It only remains to find the relevant ratio between the transition rates of different microscopic processes to reproducing the linear kinetics observed experimentally.
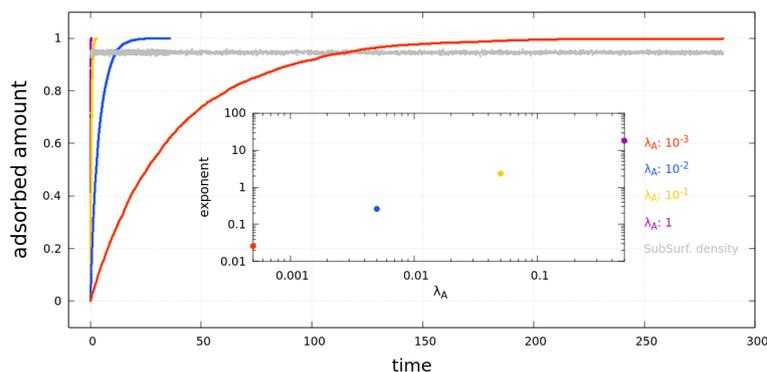
Figure 5.9: Adsorption kinetics under the high-density subsurface layer condition: Model rates are identical with the exceptions of a relatively high amount of μ_A to ensure a high density in the subsurface and a variable value of λ_A to investigate the kinetics. The grey plot shows the subsurface density. A high adsorption rate provides a high density in the subsurface. The adsorption kinetics for λ_A from more minor are shown in purple, yellow, blue, and red, respectively. Each graph fits well with an exponential function of $g(t) = 1 - exp(-\alpha t)$; hence, the kinetics follows the Langmuir adsorption model. The dots in the smaller plot are derived exponents, α. The exponent is the initial adsorption rate. The high density in the subsurface layer restricts the effective diffusive motion, which dynamically affects the effective adsorption rate, leading to the conventional Langmuir adsorption kinetics.

As shown in figure 5.9, a fully occupied subsurface generates Langmuir adsorption kinetics. In the simulations, microscopic transition rates were chosen identically, except for the pre-adsorption transition rate to the subsurface and the final adsorption transition rate. The high pre-adsorption transition rate provides a high density of particles in the subsurface, as shown by the grey plot in figure 5.9. The high density in the subsurface can be interpreted as a reservoir of particles which are "waiting" for adsorption. Hence, the accessibility of free sites in the interface controls the adsorption, and the kinetics follows the Langmuir kinetics. Each diagram is very well fitted by an exponential function of $\rho(t) = \rho_0(1 - exp(-\alpha t))$, as expected in Langmuir kinetics, where the exponent states the initial adsorption rate. The Langmuir kinetics is presented even if the final adsorption rate is increased. Therefore, we identify the pre-adsorption as a critical process.
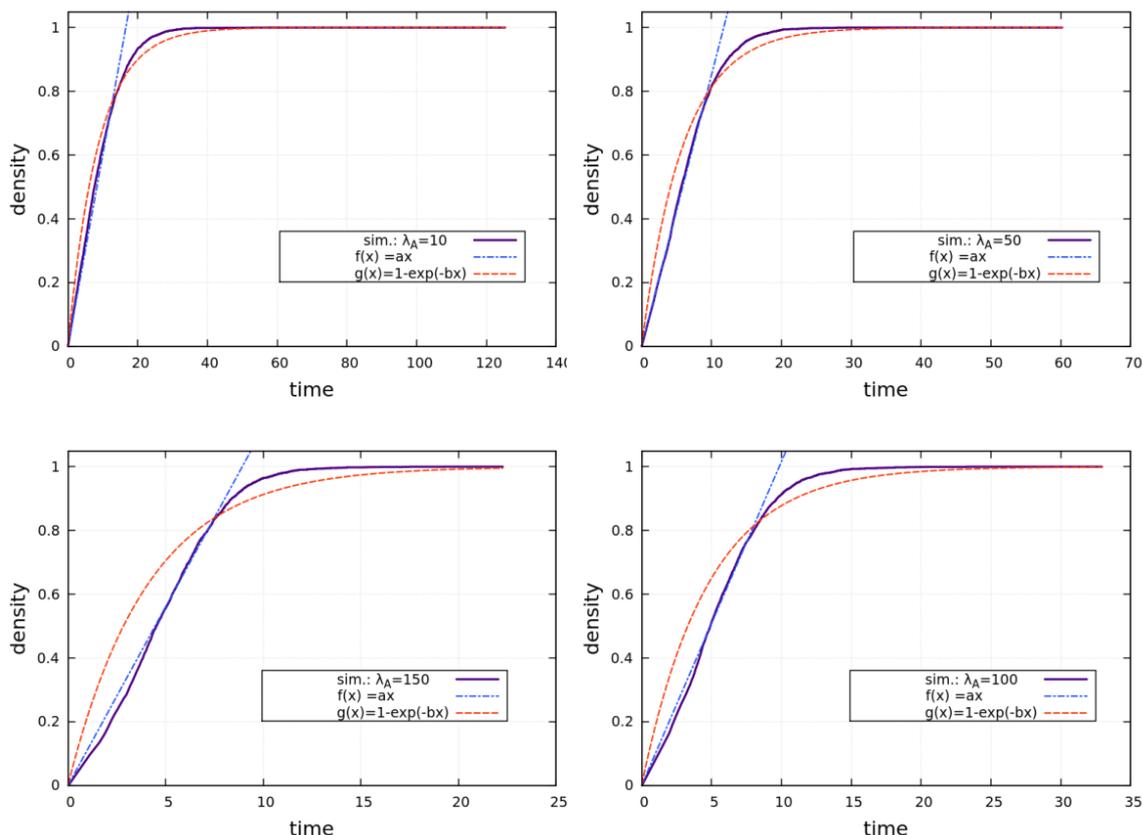
Figure 5.10: Adsorption kinetics under the low-density subsurface layer condition: Model rates are identical except for a smaller value of μ_A to ensure a relatively empty subsurface layer. The value of λ_A is a variable to study the kinetics. Each graph is fitted by a linear and exponential function, shown by dashed light blue and red, respectively. A higher final adsorption rate makes more deviation from Langmuir kinetics.

The simulation results for a low supply to the subsurface are shown in figures 5.10 and 5.11. A low pre-adsorption transition rate leads to an almost empty subsurface. Each diagram has been fitted with linear and exponential functions to guide the eyes in clarifying kinetics. The light blue and red dashed lines are fitted with the function of $f(t) = at$ and $g(t) = 1 - exp(-bt)$, respectively. Microscopic transition rates are identical to those resulting in figure 5.9, except for a low pre-adsorption transition rate and varying final adsorption transition rate. The simulations suggest that the transition from Langmuir to linear kinetics by increasing the final adsorption transition rate.
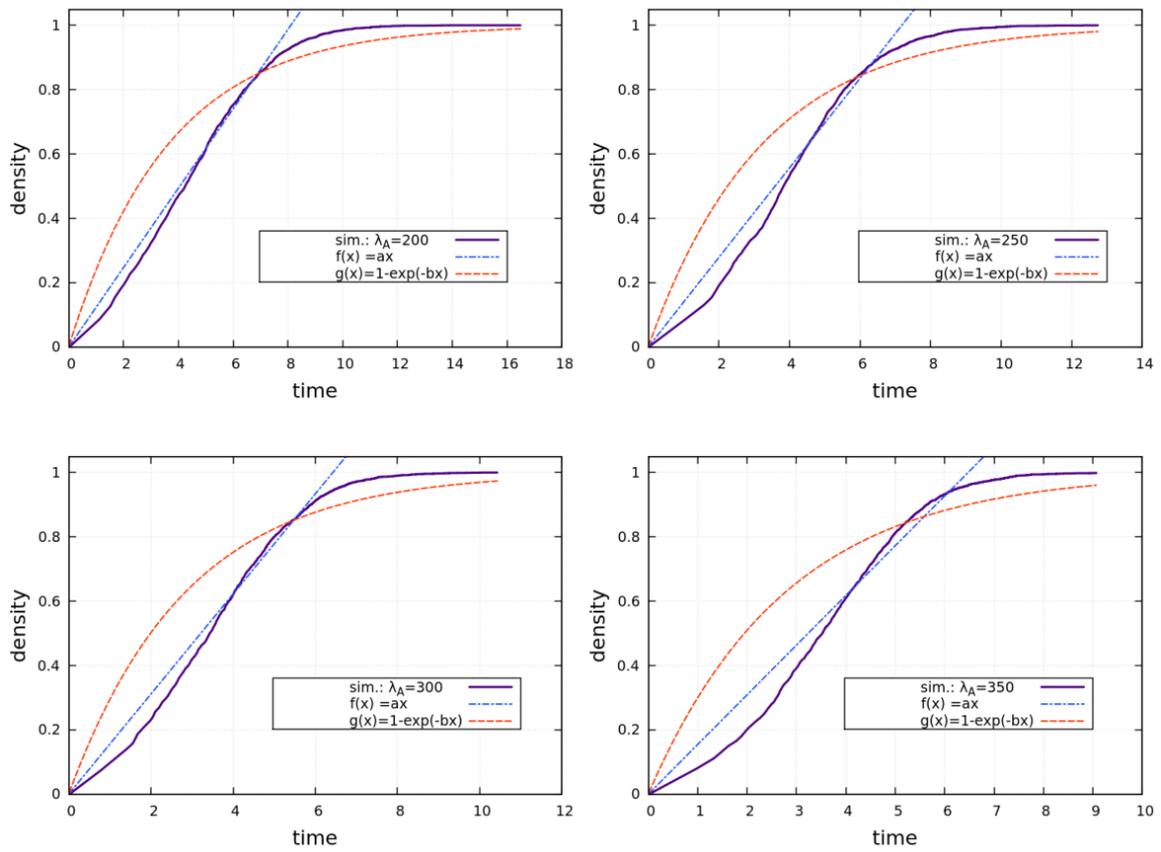
Figure 5.11: Adsorption kinetics under the less-density subsurface layer condition: Model transition rates are identical except for a minor amount of μ_A to ensure a relatively empty subsurface layer. The value of λ_A is a variable to study the kinetics. Each graph is fitted by linear and exponential functions, shown by dashed light blue and red. A higher final adsorption transition rate changes the shape of the diagrams to a concave plot.

A higher final adsorption rate, shown in figure 5.11, results in a concave shape in the adsorbed profile, i.e., the density of particles at the interface is not enough to support linear kinetics. In other words, there are not enough particles in the subsurface to adsorb; a pre-adsorbed particle leaves the subsurface before completing the adsorption. I have shown that a high density of particles in the subsurface leads to Langmuir kinetics; hence, to maintain the subsurface empty and give the particles a chance to stay longer in the subsurface, a solution is to decrease the desorption transition rate of the subsurface.
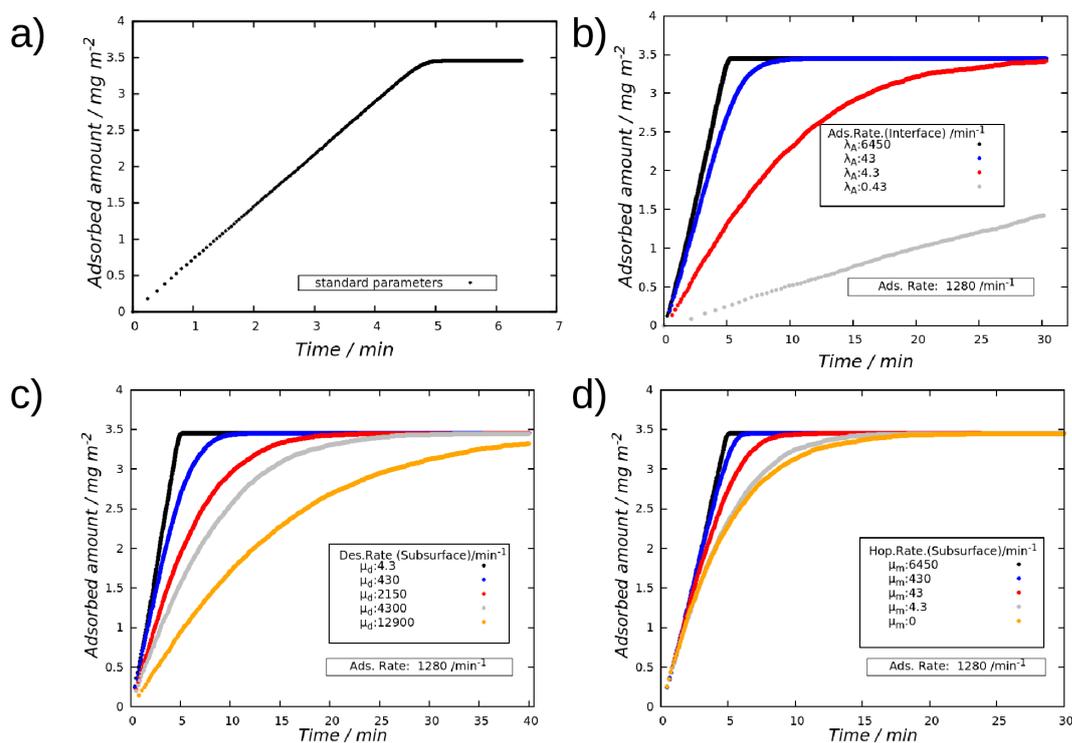
Figure 5.12: Adsorption kinetics: **(a)** Adsorption kinetics using the standard set of transition rates given in table 5.3. Effectively low pre-adsorption transition rate and low desorption transition rate to bulk, in addition to relatively high diffusive motion and high final adsorption processes, support a constant effective adsorption rate. **(b)** Reducing the final adsorption transition rate converts the kinetics from linear to the Langmuir. **(c)** Increasing the probability of desorption resembles the Langmuir adsorption kinetics. **(d)** Decreasing the diffusivity of particles in the subsurface layer affects the kinetics. A less diffusive particle in the subsurface covers a smaller area in the interface; hence, the free sites can not be found efficiently, and the adsorption kinetics convert to Langmuir kinetics. The figures **(b)-(d)** are reprinted from **Hähl *et al.* [20]**, with the permission of **American Chemical Society**.

The resulting simulation using the transition rates called hereon standard transition rates is shown in figure 5.12-a: Simple Langmuir kinetics fails to explain the experimental observations as the effective adsorption rate slows down due to the limited number of empty sites. Our approach to address this issue involved proposing a two-step adsorption process: (i) pre-adsorption to the subsurface and (ii) final adsorption to the interface. The low pre-adsorption transition rate facilitates the final adsorption process. The final adsorption and diffusive processes happen much faster compared to the pre-adsorption transition rate, allowing the pre-adsorbed particle sufficient time to find a vacant site, resulting in an effectively

constant adsorption rate. The standard transition rates in table 5.3 reproduce the experimental data of concentration $5\mu M$ with physically reasonable parameters.

The results for variation of the model transition rates in figures 5.12-b-d reveal the origin of the unusual adsorption kinetics. As illustrated in 5.12-b, reducing the final adsorption transition rate resembles the Langmuir adsorption kinetics. This choice of parameters reduces the effective vacancies search at higher densities at the interface.

Increasing the desorption transition rate recovers Langmuir kinetics, as shown in figure 5.12-c: the pre-adsorbed particle stays shorter in the subsurface, leading to a reduction of its searching area.

The diffusivity reduction in the subsurface leads to Langmuir adsorption kinetics, as shown in figure 5.12-d. Less diffusive particles cover smaller interface areas; hence, the effective search for unoccupied sites decreases significantly.
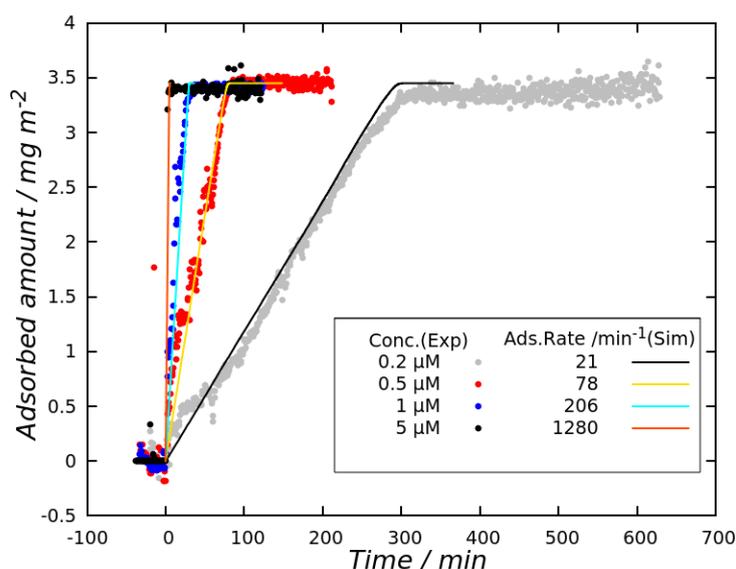
Figure 5.13: Adsorption kinetics for various concentrations. The solid lines and the dots represent the simulation and experimental results of different concentrations of HFBI, respectively. Each colour corresponds to a specific concentration, as described in the legend. The simulation results are in excellent agreement with the experimental data. The figure is reprinted from **Hähl *et al.* [20]**, with the permission of **American Chemical Society**.

In experiments, the density of proteins in bulk, meaning the protein concentration, uniquely determines the initial adsorption rate. In this model, the pre-adsorption transition rate corresponds to the bulk concentration. Thus, only varying this single

parameter is enough to produce all experimental data quantitatively. The results are shown in figure 5.13.

molecular process transition rate	symbol	standard value $[min^{-1}]$
pre-adsorption	μ_A	1280
desorption	μ_d	43
particle diffusion	μ_M	6450
final adsorption	λ_A	6450
particle diffusion	λ_M	6450
particle rotation	λ_R	6450

Table 5.3: simulation rates-Reference parameters- The table is reprinted from [20].

5.3.2 Results: effects of short-Range interactions in cluster formation

One robust feature observed in AFM images of wild-type hydrophobins is the formation of grain boundaries, see 5.1.2. These grain boundaries are visible on both the hydrophilic and hydrophobic sides of the film. Notably, the ionic strength is crucial in controlling this characteristic feature.

As the ionic strength increases in experimental conditions, fewer but more substantial grain boundaries become apparent. This differs from experiments involving HFBI-dCBM and FpHYD5, where grain boundaries are not seen; see 5.1.2.

A common observation in AFM images is not fusing clusters. The boundaries between grains become well-defined once full coverage is achieved. This preference for not fusing is attributed to the unlikelihood of neighbouring domains sharing the same protein orientation. To model this feature, I assign a spin property to proteins, as mentioned in 5.2. It is energetically favoured if two adjacent proteins share an identical spin. The parameter of J_0, the coupling constant, adjusts the strength of the spin interaction in this model. A high value of J_0 indicates a high attraction, while a zero value switches the model to the isotropic system. The spin-interaction effects are illustrated in figure 5.14.

The isotropic limit is shown in figure 5.14-a; in contrast to experimental observations, individual clusters easily fuse, forming a single domain; hence, no large individual clusters are visible. In the solid spin-interaction limit shown in figure 5.14-b, the clusters formed at the interface resemble domains typically observed in diffusion-limited aggregation [63]. Strong attraction in this scenario inhibits diffusive motion

at the interface, resulting in dendritic structures that differ from the well-defined roundish grain boundaries. Figure 5.14-c illustrates the configuration of the tuned interactions. Here, similar to experimental observations, individual stable clusters are formed and grow until their borders touch.
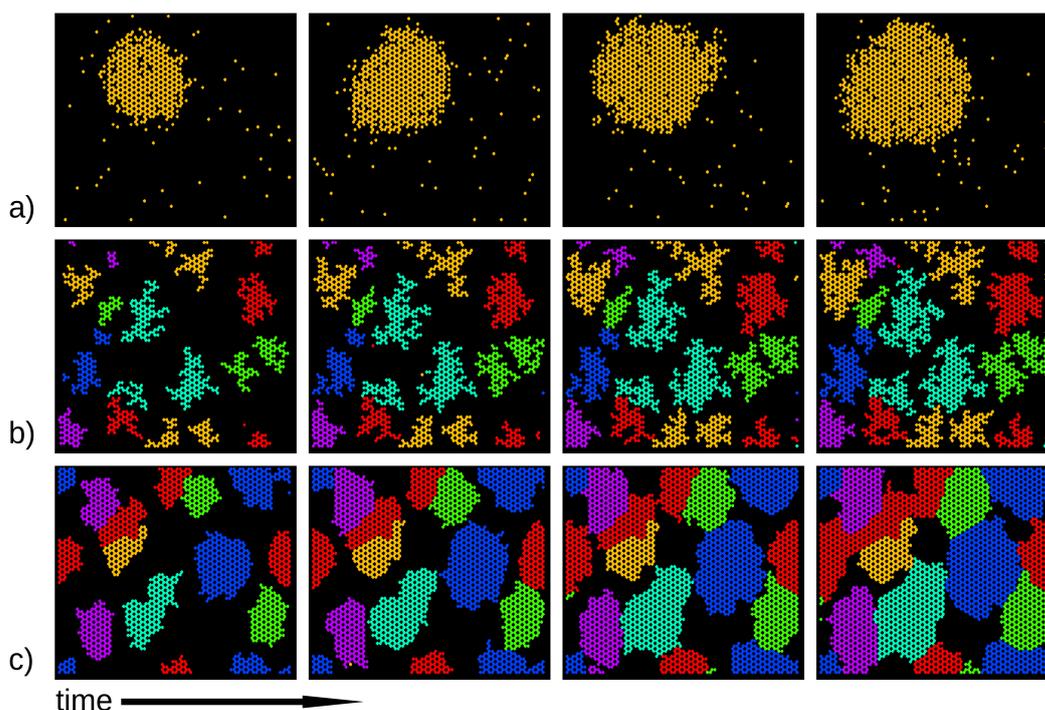
Figure 5.14: Snapshots of the surface layer for different model transition rates. **(a)** System without spin interactions. **(b)** Limit of strong spin interactions. **(c)** Standard model transition rates. The simulation time is increased from left to right. The figure and the caption are adapted from **Hähl *et al.* [20]**, with the permission of **American Chemical Society**.

The choice of parameters leads to three distinct regimes: (i) a zero coupling constant results in a single domain; (ii) a non-zero coupling constant leads to different domains that grow until they touch and remain stable; and (iii) medium interaction strength results in roundish clusters, akin to the experimentally observed structures.

This model suggests that the roundish clusters observed experimentally result from lateral short-ranged interactions, which indicate a preferred orientation. A prerequisite for cluster formation is close contact between the hydrophobins. In the case of bulky hydrophobins such as HFBI-dCBM, the attached domains prevent close affinity, resulting in a homogeneous film with no cluster formation.

5.4 Summary and Conclusion

This chapter delves into the fascinating world of class II hydrophobins, specifically HFBI and HFBII, and their unique adsorption kinetics at the air-water interface. The experimental observations reveal intriguing behaviour: these hydrophobins self-assemble irreversibly at the interface, with linear kinetics and a constant effective adsorption rate up to saturation, abruptly changing to zero. In contrast, the mutant HFBI-dCBM adheres to Langmuir adsorption kinetics, an exponentially increasing profile in such cases.

The key surprise here is the role of repulsive electrostatic interactions in controlling the kinetics supported by the restoration of Langmuir adsorption kinetics at higher ionic strengths. This chapter also explores how ionic strength affects the initial adsorption rate.

Atomic force microscopy imaging illustrates that wild-type protein film formation results in grain boundary formation at the interface, a feature not seen with bulky mutants. This observation suggests that close contact is a prerequisite for protein domain formation. Electrostatic interactions profoundly affect cluster formation, leading to fewer, larger grain boundaries at higher ionic strengths, even with the same mass adsorbed.

A Monte Carlo scheme was employed to develop a stochastic lattice gas model and explore the experimental observations. This model suggests a two-stage adsorption process: reversible pre-adsorption near the interface followed by irreversible final adsorption. The pre-adsorption phase keeps proteins close to the interface, allowing them to dynamically search for vacancies in the interface, rendering the adsorption coverage independent and aligning with experimental results.

Additionally, the stable visual clusters in experiments imply a strong short-ranged interaction preventing domain fusion. A spin interaction model is introduced to capture this interaction, as neighbouring proteins of the same spin construct clusters. Diffusive motion and rotational processes enable proteins to locate their domains at the interface.

Importantly, this model may have broader applications, particularly in describing the adsorption of other surface-active proteins where short-ranged and electrostatic interactions play a significant role, especially when denaturation at the interface is negligible.

Chapter 6

Mechanical Response

Interfacial films formed from hydrophobins are known for their robust elastic response. Shear elastic moduli measured for films formed by pure or a mixture combination of hydrophobins are higher than other investigated proteins, e.g., lysozyme. The solidification of the film at the interface causes elastic behavior. Moreover, the film formed by the class II hydrophobins interested in this thesis shows a repeating pattern close to the hexagonal structure, indicating lateral interactions between proteins at the interface, see section 2.1.8.

Here, I aim to answer the question of whether the mechanical stability of the structure is related to the postulated angular interaction. To this end, I develop a stochastic model in the continuum space to obtain a structural insight into the mechanical properties of those films. I quantify the mechanical response of the film under the quasi-static tension process by measuring the two-dimensional order parameters of the system.

This chapter contains the model and results. The first section presents the fundamentals of the model, including the modeling procedure and introducing measurements. The second section presents the simulation results of modeling thermal fluctuations as the first step of the transition to the continuum space. The results of modeling quasi-static stretching and the response of different model systems are presented in sections three to five. In the last section, I examine if this approach can detect a transition from a disordered to an ordered phase.

6.1 Methods and measurements

6.1.1 Thermal fluctuations
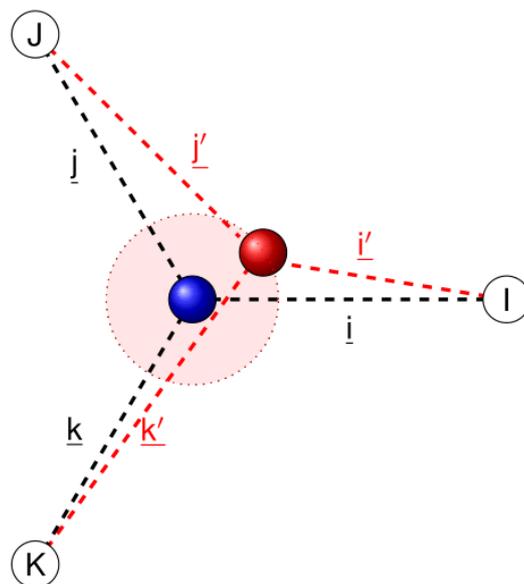
Figure 6.1: Thermal fluctuations. A particle shown in blue is chosen randomly. The black dashed lines represent the physical bonds connecting the referenced particle to its nearest neighbors. A new position inside a circle centered around the current site, shown in light pink, is proposed. The red ball in the figure shows the suggested position. The bonds connecting the suggested position to the nearest neighbors are depicted by red dashed lines. The proposed position is accepted according to the Metropolis probability, i.e., $min\left(1, \exp\left(-\frac{\Delta E}{k_B T}\right)\right)$.

Lattice models efficiently study adsorption processes in which the observation and the adsorption time scales are in the same order as the lattice formation. The time scale of the adsorption processes is in a few minutes, as observed experimentally; see section 5.1.1. However, the mechanical response of the system happens in more minor time scales, i.e., in the scale of nanoseconds; hence, a study in the framework of lattice models is not practical and continuous models are essential to study the mechanical response of the system. We elaborate our lattice model to the continuum space by implementing thermal fluctuations. To this end, the model system is initialized as a perfect honeycomb lattice. Afterwards, the Monte Carlo scheme, as I present in the following, is used to apply thermal fluctuations. As shown schematically in figure 6.1, a particle labeled by blue is chosen randomly. A new position inside the circle centered on the current position is proposed. The new position consists of

two independent, normally distributed random numbers. The proposed position is accepted according to the Metropolis probability, i.e., $min\left(1, \exp\left(-\frac{\Delta E}{k_B T}\right)\right)$, where ΔE is the energy variation due to movement, T is temperature, and k_B is the Boltzmann constant. Choosing a high number of Monte Carlo sweeps ensures that all particles fluctuate at least once. Hence, the entire system reaches a new equilibrium state in contact with a thermal bath of temperature T. The number of Monte Carlo sweeps is a size-dependent factor in the simulation. Since I use a different number of particles in various studies, I will present the exact number of particles and the number of used sweeps at the beginning of each relevant section.

6.1.2 Hamiltonian and quasi-static stretching
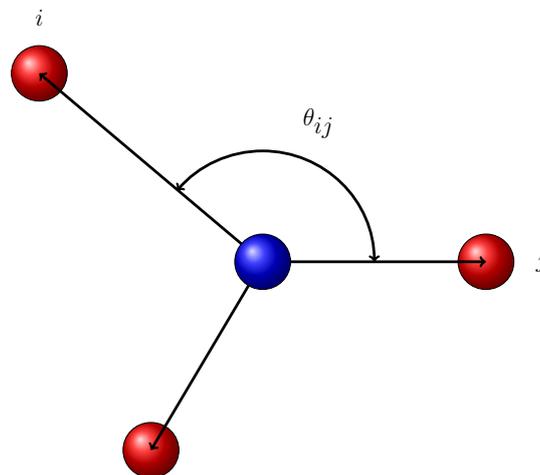
Figure 6.2: Angle-dependent potential. The angle-dependent potential considers the orientational ordering in the system. A configuration in which two adjacent bonds make a particular angel are energetically favored. The reference particle is shown in blue. The solid black lines connect the reference particle to its nearest neighbors, depicted by red balls. The angle between two adjacent bonds, θ_{ij}, is shown with an arc.

As I presented in 2.1.8, the films formed by hydrophobins, particularly HFBI, are known for exceptionally high shear moduli. Previous studies suggested the lateral interaction between proteins at the interface can be responsible for this feature [65]. The charged amino acids inside the protein can cause those interactions. Hence, the arrangement of amino acids generates an anisotropy in the orientation of proteins at the interface. This anisotropy leads to the formation of crystal structure. In our

lattice model presented in 5.2, the interactions of the system were considered in both long- and short-ranged interactions. The long-ranged interactions provided the isotropic part of the Hamiltonian and were considered in the framework of DLVO theory, in which the radial distance between proteins determines the interaction. On the other hand, the short-ranged interaction affects only the nearest neighbor proteins. This interaction provides the anisotropic orientation and is modeled by introducing the "spin" property for proteins. I showed that the spin property is responsible for cluster formation at the interface; see 5.3.2.

In this chapter, the stability of the structure under thermal fluctuations and external stresses is investigated. Hence, an off-lattice model is essential. The isotropic interactions are independent of the nature of the model; thus, the isotropic interactions in this off-lattice model are implemented exactly similarly to the lattice model. However, the off-lattice model has to include the honeycomb structure observed experimentally. To this end, a short-ranged angular interaction potential is introduced, where it can retain the domains of the honeycomb structure and cause anisotropy in the orientation. Figure 6.2 shows schematically how the angular interaction works; each particle has a certain number of nearest neighbors. A link between a given particle and each neighbouring particle is assumed. These links are physical bonds. Hereon, for simplicity, I will call them "bonds." Two adjacent bonds have energetic benefits in a specific range of angles. The angle-dependent potential is defined by

$$U(\theta) = \begin{cases} -J_0 & \text{if} \quad 115° < \theta < 125° \\ 0 & \text{otherwise}, \end{cases}$$

where θ is the angle formed between two adjacent bonds, depicted schematically in figure 6.2, and J_0 is a coupling constant that takes the values zero to $16.0 \ U_{min}$, and U_{min} is the minimum value of the isotropic potential in the optimum configuration of two particles, i.e., $U_{min} = U(r_{min})$, introduced in section 5.2.2. Hence, the Hamiltonian is given by

$$U(r) = \frac{1}{2} \sum_{k=1}^{N} \left(\sum_{\substack{i \neq k}}^{r_{ik} < r_{cutoff}} U(r_{ik}) + \sum_{\{i,j\}} U_k(\theta_{ij}) \right),$$

where the first sum is taken over all particles in the system, the second sum is taken over neighbors in a defined cutoff radius, and the third sum is taken over the nearest neighbors of the given particle at arbitrary distances in a cutoff radius

of angular interaction. Periodic boundary conditions have been considered in all energy calculations.

Furthermore, to understand the importance of the angular interaction in the stability of the hydrophobin membranes, I study the mechanical response of the film to the external stresses; the film is exposed to quasi-static stretching. The initial state of the system is the equilibrated state after implementing thermal fluctuations. The following procedure is used to implement quasi-static stretching. The procedure involves stretching the system, followed by a relaxation process. The stretching process involves rescaling the position of particles, x_i, and the length of the film, l, by the scaling factor α, i.e.,

$$x_i \rightarrow (1+\alpha)x_i,$$
$$l_i \rightarrow (1+\alpha)l_i,$$

where the value of α is set to $0.005\, l_0$ and l_0 is the length of the film before rescaling and x_i is a vector component. A relaxation process follows each stretching step to achieve equilibrium: According to the Monte Carlo procedure introduced in section 6.1.1, the particles fluctuate thermally around their current positions. In this part, I study a system of 384 particles. Each simulation includes 70 cycles of stretching-relaxation, where each relaxation part consists of 1.8×10^6 Monte Carlo sweeps. By the end of each simulation, the surface area and the surface density fraction reach the values $2.02\, A_0$ and $0.49\, \rho_0$, respectively, where A_0 and ρ_0 denote the initial area and surface number density fraction, for N proteins at the interface of the initial length of l_0, i.e., $A_0 = l_0^2$, and $\rho_0 = \frac{N}{A_0}$.

6.1.3 Analysis and measurements

The response of the system is quantified by the order parameter of the system as well as by examining the film structure. As mentioned in 4.3.1, a solid in a crystal phase is orientationally long-range ordered. The global bond-orientational order parameter, $|\psi_6|$, quantifies the orientational correlation of the bonds connecting nearest neighbors, see 4.3.2.

Additionally, I measure the radial distribution function (RDF), $g(r)$, of the system. As mentioned in 4.3.1, $g(r)$ is related to the structure factor of the system and deter-

mines how the density varies as a function of distance from a reference particle. Suppose a two-dimensional system of N particles of a total area A. The radial distribution function is obtained by choosing an arbitrary particle as the origin and counting the number of particles whose centers lie in the shell of Δr of a circle of radius r of the origin, as shown in figure 6.3. In a simple algorithm to determine $g(r)$, the distance between all particle pairs is calculated and is binned into a histogram of bins of width Δr. The average density number of particles, $\rho_0 = \frac{N}{A}$, normalizes the histogram. Hence, g(r) is given by

$$g(r) = \frac{(\text{Number of bonds of length}(r, r + \Delta r))}{2\,(2\pi r\,\Delta r)\,\rho_0},$$

where the first factor of two in the denominator comes from double counting the distance between two particles, and $2\pi r\,\Delta r$ is the area of the shell highlighted in pink in figure 6.3. The width of the bins, Δr, is considered as $0.03\,a_0$ where a_0 denotes the constant of the ideal honeycomb lattice.
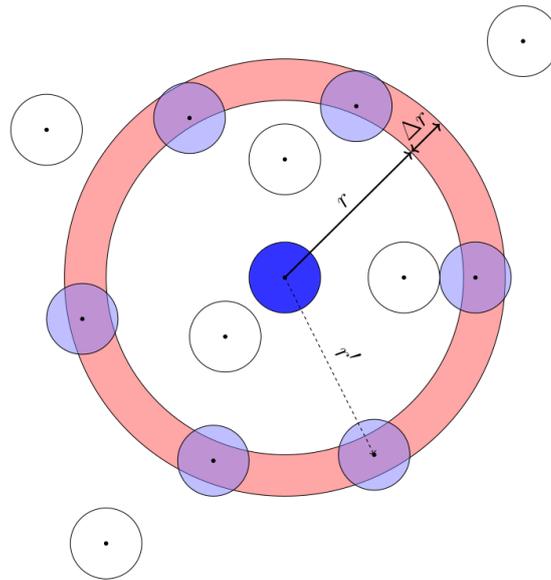
Figure 6.3: Radial distribution function (RDF) is obtained by choosing a particle as the origin and counting the number of particles whose centers are positioned in distance $(r, r + \Delta r)$ of the origin. The figure is inspired by [6].

Additionally, the Voronoi tessellation, introduced in 4.3.3, is used to visualize defects in the structure. The Voronoi tessellation determines the number of nearest neighbors. Instead of showing the Voronoi cell, I illustrate each particle as the center of the cell marked in a specific color representing the number of nearest neighbors.

I use the phrase "the nearest neighbor visualization" or, as the abbreviation, n.n.v hereon for it. From the visual inspection, appearing particles in similar colors may represent an ordered structure.

In contrast, the coexistence of different colors is evidence of a disordered phase. This visualization leads eyes to see the order in the system; however, it is inadequate; additionally, measuring the bond orientational order parameter, $|\psi_6|$, determines the exact ordering of the system. Two exemplary n.n.v are shown in figure 6.4. The structure shown on the left panel is an ordered system with an observable honeycomb structure; each particle has three nearest neighbors colored in cyan. The right panel depicts a disordered system. Each color represents a particle with a certain number of neighbors. Red, dandelion, green, cyan, blue, purple, and orange represent particles with 0 to 6 neighbors for $r < r_{\text{cutoff}}$, respectively.
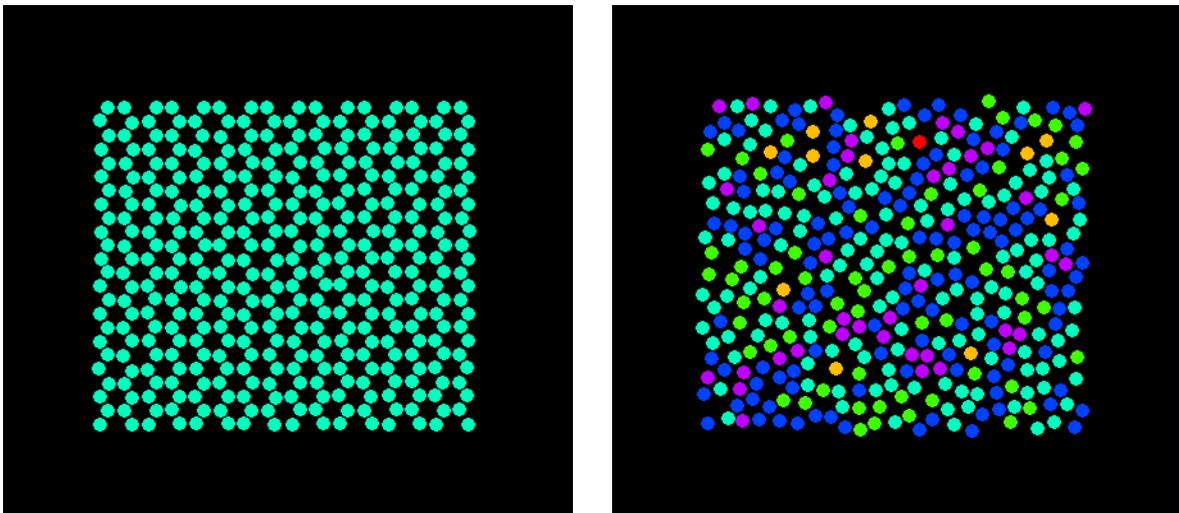
Figure 6.4: Comparison of n.n.v in ordered and disordered structures; instead of showing the area of the Voronoi construction, for better visualization in the presence of the high number of particles, each particle has been marked in a particular color representing the number of neighbors. The left panel shows the n.n.v of an ordered phase where each particle has three nearest neighbors colored in cyan, and the right panel represents a disordered phase. Each color represents a particle with a certain number of neighbors. Red, dandelion, green, cyan, blue, purple, and orange represents particles with 0 to 6 neighbors for $r < r_{\text{cutoff}}$, respectively.

6.2 Results: stability of structure

6.2.1 Thermal fluctuations
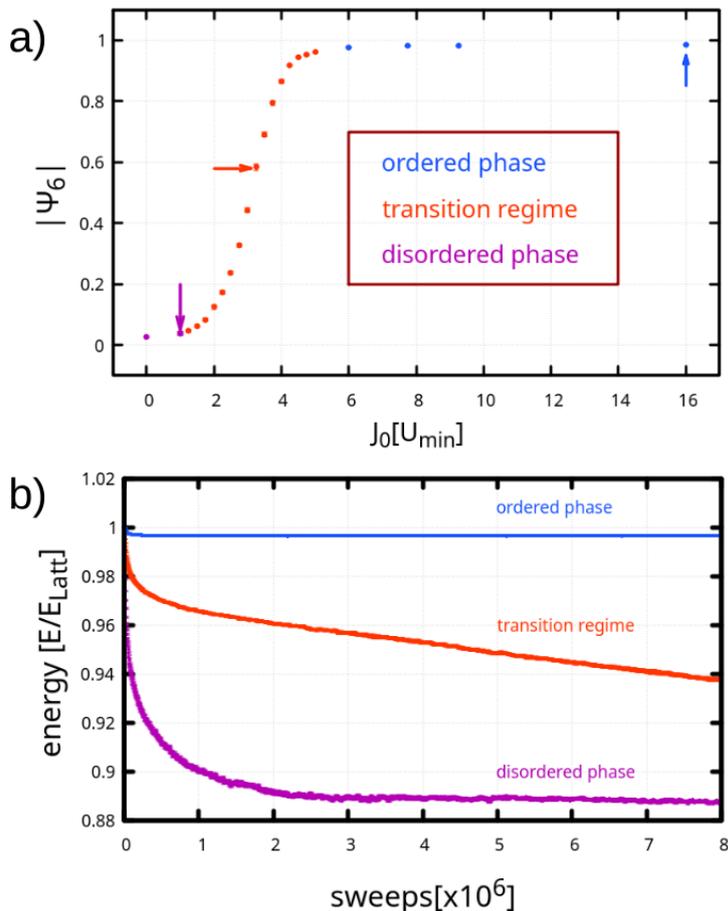
Figure 6.5: (a) The global bond orientational order parameter, $|\psi_6|$, dependent angular inter-action coupling constant, J_0. The blue dots present the regime in which the system shows long-range order. The purple dots represents the disordered phase. For the amplitudes shown in orange, the system could not reach its equilibrium in the time window of the simulation. **(b)** The energy evolution of the system during the Monte Carlo simulation of amplitudes pointed by arrows. The energy is normalized to the energy of a lattice of the same number of particles. Identical colors, as in the panel **(a)**, are used to present the energy evolution of the system.

In this part, the stability of the structure under a slow relaxation process is probed. The studied model is a system of 1536 particles positioned on ideal sites of a honeycomb lattice. The system is subjected to Monte Carlo motion. Hence, particles move freely from their current positions on a continuum space. According to the size of the system, 8×10^6 Monte Carlo sweeps are chosen. Here, the rule of angular

interaction in the stability of the structure is investigated by measuring the global orientational order parameter, $|\psi_6|$, as well as by the radial distribution function, $g(r)$.

As mentioned in section 4.3, an ideal crystal is orientationally ordered, and the bonds connecting nearest neighbors are correlated at long distances. The bond orientation order parameter, $|\psi_6|$, quantifies this correlation. $|\psi_6|$, measured for varying coupling constant, is shown in figure 6.5. The coupling interaction constant is characterized by J_0. The simulation results show three different regimes for the thin film depending on the choice of coupling constant: (i) The disordered phase in which the ordered phase is unstable within the slow relaxation process; the bond orientational order parameter is close to zero. This phase is achieved for coupling constant of zero and one, which are shown by purple dots in the figure; (ii) The ordered phase, shown by blue dots in the figure. For a high enough coupling constant, thermal fluctuations do not destroy the order of the system; and (iii) The transient regime, shown by orange dots. In this phase, the coupling constant is not high enough to maintain the structural order of the system. However, the simulation time window was short, and the disordered phase could not be achieved.

The panel (b) of figure 6.5 shows the energy evolution of the system during the Monte Carlo process. The energy is normalized to a lattice system of the same number of particles. Both disordered and crystal phases reach the stationary state, while the energy plot of the non-equilibrium phase does not within the accessible simulation time. Here, it is worth emphasizing that we are interested in the stability of the phase rather than the possibility of the phase transition.
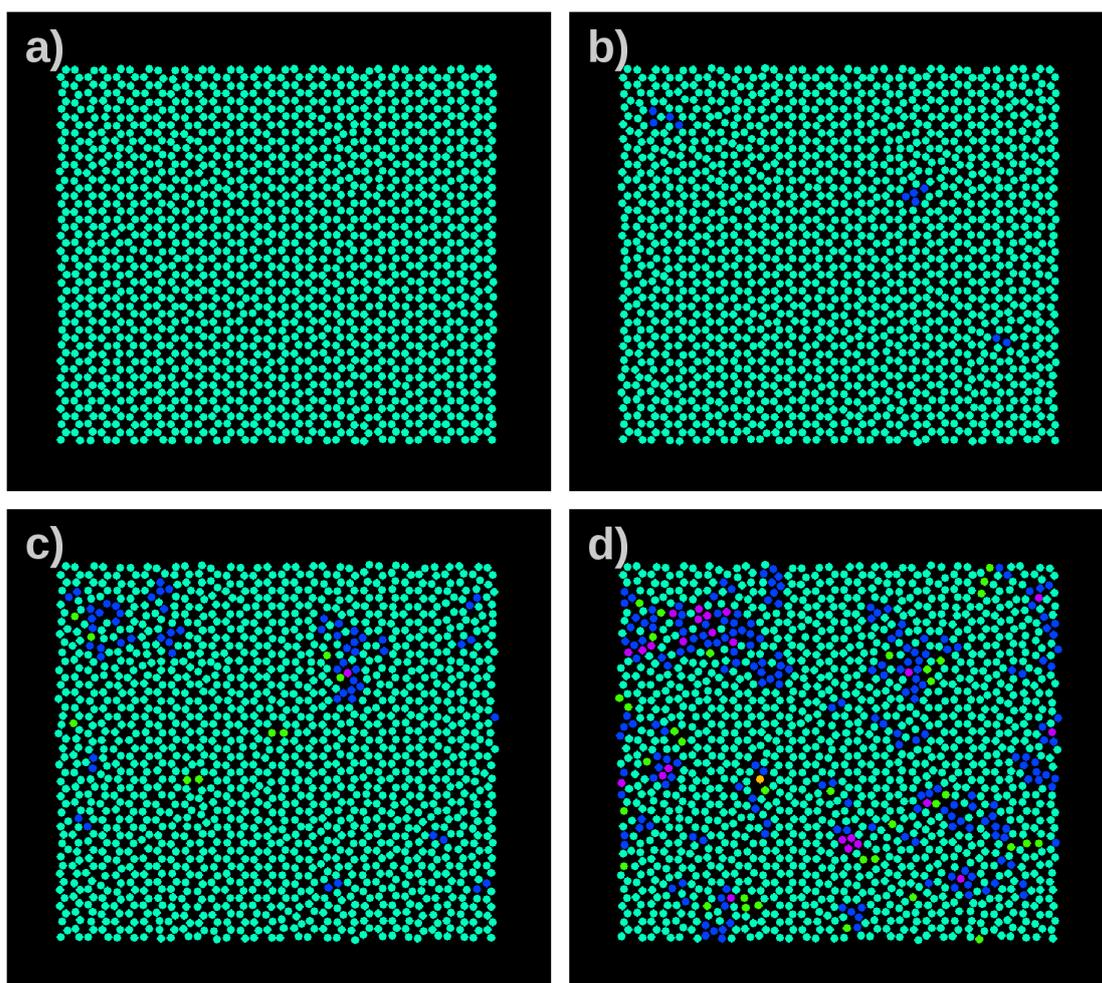
Figure 6.6: Evolution of the n.n.v after **(a)** 2×10^6, **(b)** 4×10^6, **(c)** 6×10^6, and **(d)** 8×10^6 sweeps for the coupling angular interaction constant $J_0 = 3.5U_{min}$. The disordered nuclei appear in the system and grow over time. Each color represents a particle with a certain number of neighbors. Red, dandelion, green, cyan, blue, purple, and orange represents particles with 0 to 6 neighbors, respectively.

The nearest neighbor visualization reveals more details of the film structure. To better understand the structural evolution in the non-equilibrium phase, I present the nearest neighbor visualization during a simulation with coupling constant $J_0 = 3.5U_{min}$ in figure 6.6. The simulation results show the formation of disordered nuclei and their growth over time in the non-equilibrium phase. Hence, the simulation suggests that the system will reach a disordered phase after sufficient Monte Carlo sweeps.
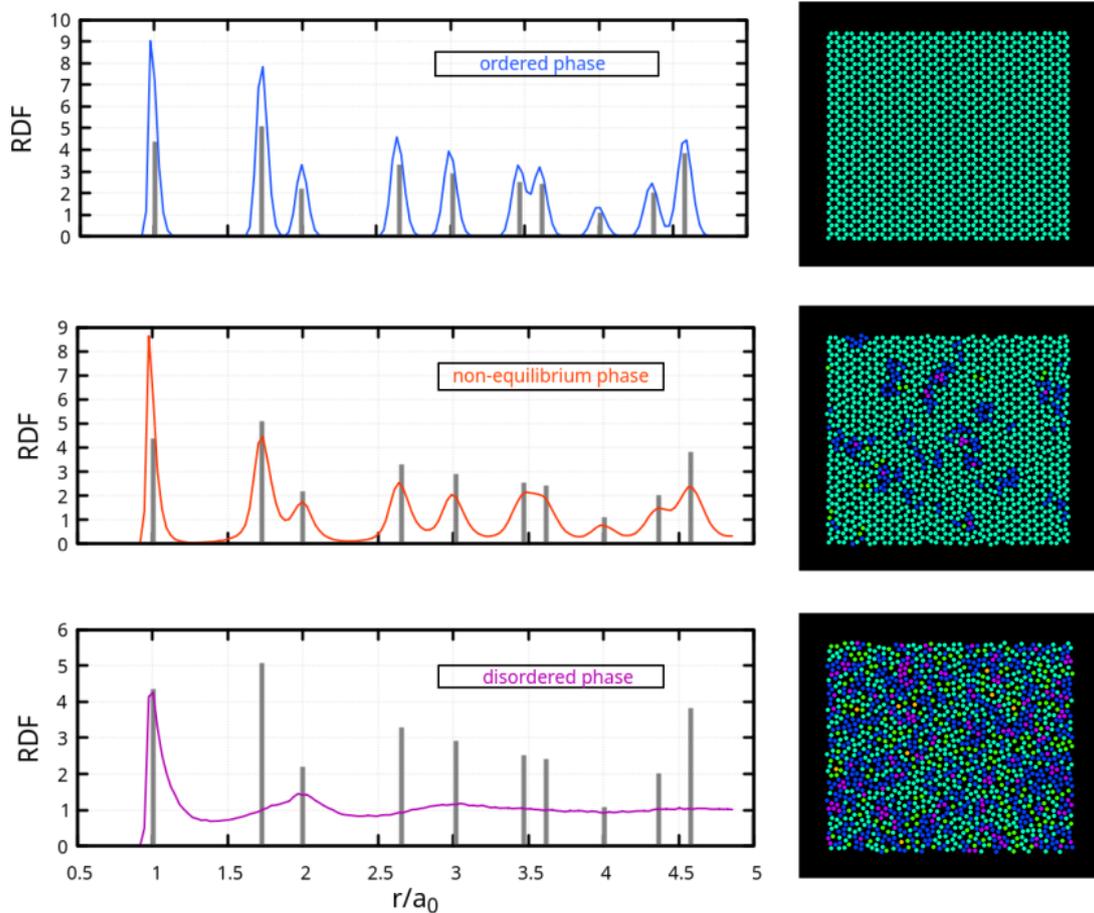
Figure 6.7: Radial distribution function, $g(r)$, and the nearest neighbor visualization of different phases. The first row is the ordered phase. The radial distribution function is shown in blue. The positions of peaks in $g(r)$ follow the positions of the neighbors in a honeycomb structure shown by gray bars in the diagram. The perfect crystal structure can be observed. The other extreme, plotted in the third row with the purple radial distribution function, presents a disordered phase in which $g(r)$ exhibits a prominent peak for the first nearest neighbors and a secondary peak. The second row corresponds to a transient regime. The radial distribution function shows peaks that follow the honeycomb structure but with significant variance, indicating losing the order. Several disordered cores appear in the nearest neighbor visualization, confirming that the condition of the system can not be maintained. Identical colors, as in the figure 6.5, are used to present the radial distribution function of the system. In the nearest neighbor visualization, each color represents a particle with a certain number of neighbors. Red, dandelion, green, cyan, blue, purple, and orange represents particles with 0 to 6 neighbors, respectively.

The radial distribution function, $g(r)$, and the nearest neighbor visualization of the film for particular coupling constants specified by arrows in figure 6.5 are shown in figure 6.7. Identical colors, as in the figure 6.5, are used to present the radial

distribution function of the system. The disordered phase is shown in the third row; in addition to the radial distribution function with a significant primary peak followed by a shorter secondary peak, the nearest neighbor visualization confirms the disordered phase.

Conversely, in the first row, the nearest neighbor visualization of the phase plotted in blue shows perfect order. Each particle has three neighbors, the honeycomb structure can be observed, and the position of the honeycomb structure peaks matches the peaks of $g(r)$ of the system.

In the middle, the orange diagram of the non-equilibrium phase shows features between two extreme phases: The formation of disordered nuclei differentiates this phase from the crystal. The radial distribution function peaks match a honeycomb structure; however, these peaks, compared to an ideal arrangement, have a more significant variance, which indicates losing the order in this system.

6.3 Modeling the mechanical response to the external stress

The membranes formed by hydrophobins are known for significantly high shear elasticity and buckling strength. To understand the origin of these features, we study the mechanical response of the film composed of 384 particles under quasi-static stretching, see Sec. 6.1.2. The rigidity of the system depends on the two inter-particle interactions: the Lenard-Jones interaction, which is isotropic, and the anisotropic-proposed angular interaction. The Lenard-Jones potential goes through a minimum, U_{min}, at the radial distance of $r_{U_{min}} = 1.012a$, where a is the lattice constant of an ideal crystal. We chose a cutoff radius of $r_{cutoff} = 1.3\ r_{U_{min}}$ of angular interaction. We quantify the response by the bond orientational order parameter,$|\psi_6|$, measured for varying coupling angular interaction constant, J_0, for a system under stretching tension. Depending on the choice of the coupling constant, we identify three parameter regimes: (i) The ordered phase. This phase is achieved for the coupling constant higher than $12\ U_{min}$. As I will discuss in section 6.3.1, the angular interaction is the most crucial factor in stabilizing the crystal structure; tension results in a rupture in this phase. The fracture is evident by appearing a sharp increment in the energy profile; (ii) The disordered phase is achieved for a coupling constant of less than four. $|\psi_6|$ is very low, and the system is disordered. The energy of the system increases smoothly by stretching the film; (iii) The intermediate phase in which the system loses order. A more detailed discussion of these phases is presented in the following.

6.3.1 Ordered phase
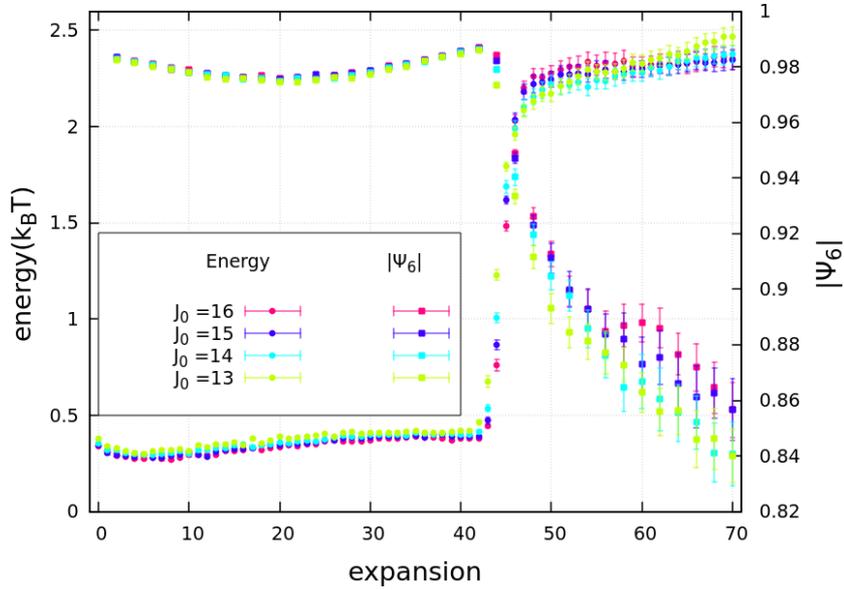
Figure 6.8: Ordered phase. Energy and the bond-orientational order parameter, $|\psi_6|$, are shown in terms of expansions for the ordered phase. The energy is normalized to the number of particles in the system. There exists a sharp jump in the energy profile, indicating the rupture in the system. The order parameter of the system, $|\psi_6|$, shows a remarkable feature; it has a local minimum followed by a local maximum, after which the order parameter decreases monotonically. The local maximum is a significant feature of this phase, indicating that the angular interaction is dominant before rupture. Dots and squares show the energy and $|\psi_6|$, respectively. Simulated data and error bars represent the mean and standard error of the mean, SEM, taken over 20 simulations. Individual values of the coupling constant of angular interaction, J_0, are reported in the legend.

The ordered regime observed in our simulations is characterized by two dominant features: (i) There exists a sharp jump in the energy profile in terms of expansions, and (ii) the bond orientational order parameter shows a local minimum followed by a local maximum. The sharp jump presents a rupture in the system. Our results detect the ordered regime for the coupling constant of the angular interaction, J_0, higher than $12\,U_{min}$. Figure 6.8 shows the simulation results in the mentioned range of coupling constant. The local maximum happens in the distance between two particles close to the critical distance of the angular interaction. Hence, according to the simulation results, the strong-short range interaction, i.e., the angular interaction, drives the rupture. In the following, I present a detailed discussion of the results obtained in this regime for the coupling constant of $16\,U_{min}$.
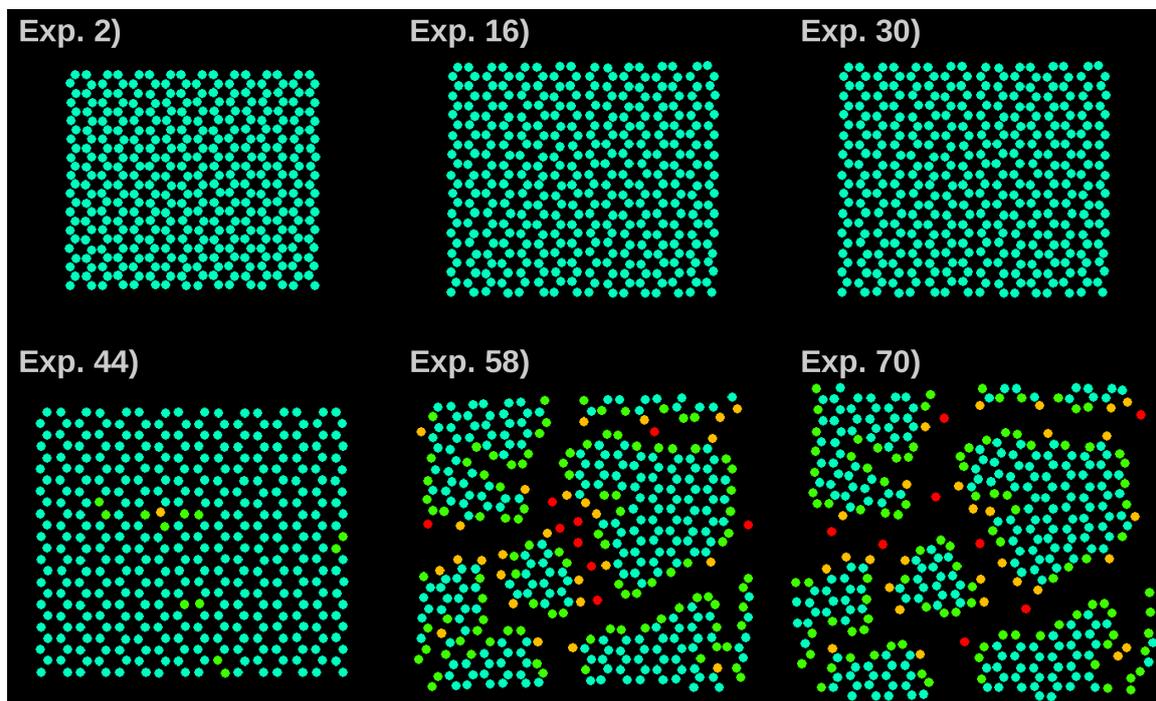
Figure 6.9: The nearest neighbor visualization of a stretching system. The initial state is the honeycomb-ordered system in thermal equilibrium with a bath. The coupling constant of the angular interaction, J_0, is $16\ U_{min}$. Each expansion follows a relaxation process. The crystal structure of the system is observable. The rupture happens during courses of 42 to 50, shown in figure 6.10. After the rupture process, each cluster shows a honeycomb structure; see panels of expansions 58 and 70. Each color represents a particle with a certain number of neighbors. Red, dandelion, green, cyan, blue, purple, and orange represents particles with 0 to 6 neighbors, respectively.

The evolution of the nearest neighbor visualization of a film of a coupling angular interaction constant of $16\ U_{min}$ during the stretching process is shown in figure 6.9. The honeycomb structure is observable; the cyan color for all particles indicates that before expansion 40, each particle is tightly linked to three neighboring sites. The coupling constant, $J_0 = 16\ U_{min}$, is high enough to maintain the crystal structure. Further stress from expansion leads to a rupture in the interface. Occurring particles in color different from cyan means some particles are losing ideal three-fold neighboring; hence, fracture in the film is visually evident. The defective sites spread through the system, and more stress leads to widening gaps and causes an observable rupture, as is evident in the panel of expansion 58. After the fracture, separated clusters are distinguishable; each cluster shows a well-defined honeycomb structure. The position of initial defects and propagation through the surface layer is a stochastic

process leading to different patterns in individual realizations. A step-by-step rupture process is shown in figure 6.10.
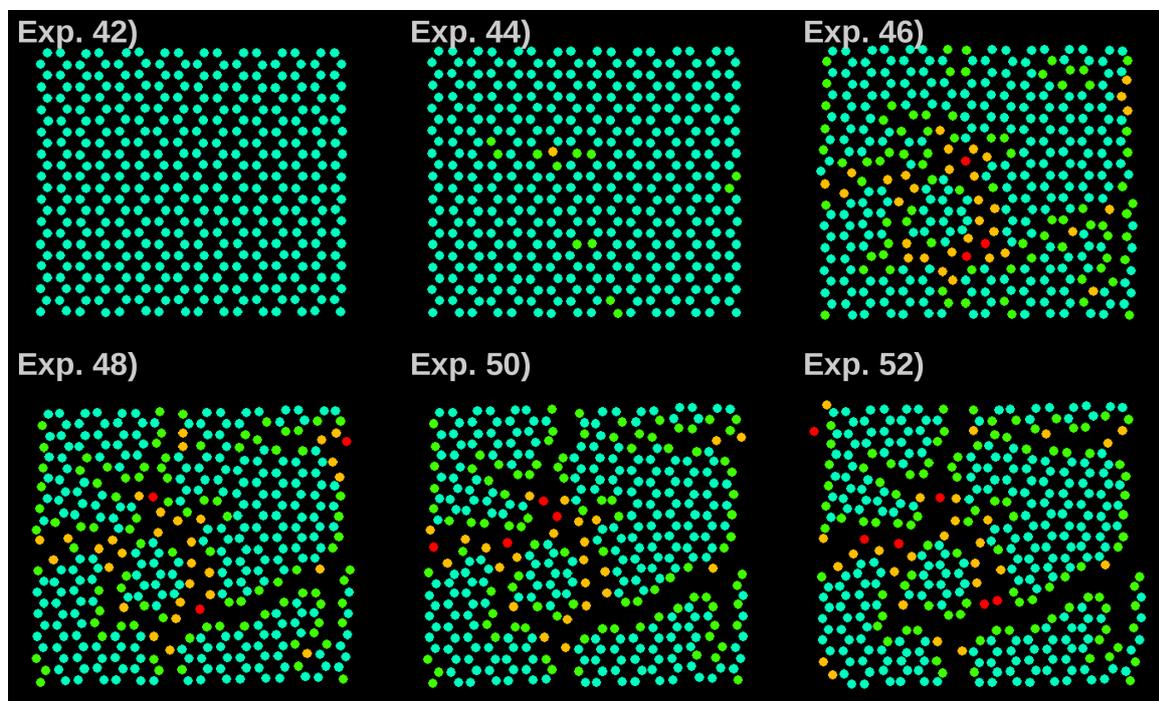
Figure 6.10: Stepwise rupture process. Although it is not visible, the rupture process starts at the 42nd expansion. The stretching stress results in defects in the system, which are observable by changes in the color of particles representing the number of nearest neighbors. More tension causes the defective sites to spread throughout the system until the isolated clusters appear. Each color represents a particle with a certain number of neighbors. Red, dandelion, green, cyan, blue, purple, and orange represents particles with 0 to 6 neighbors, respectively.

The radial distribution function, $g(r)$, shown in figure 6.11, confirms the ordered structure observed in the nearest neighbor visualization. The first panel shows $g(r)$ before starting the expansion process. The following panels show the evolution of the radial distribution function in terms of expansions. The stretching process shifts the position of each peak to the right (see Fig.6.11), meaning that on average, each bond length elongates consistent with the current accessible area. However, the $g(r)$ of the system almost matches the undisturbed honeycomb lattice. In each panel, $g(r)$ of a perfect honeycomb lattice corresponding to the current accessible area is shown as a reference by the gray vertical lines. The stretching process does not disturb the mean position of the peaks in a specific accessible area; however, it leads to the peaks of smaller heights but with more significant variance (see figure 6.11).
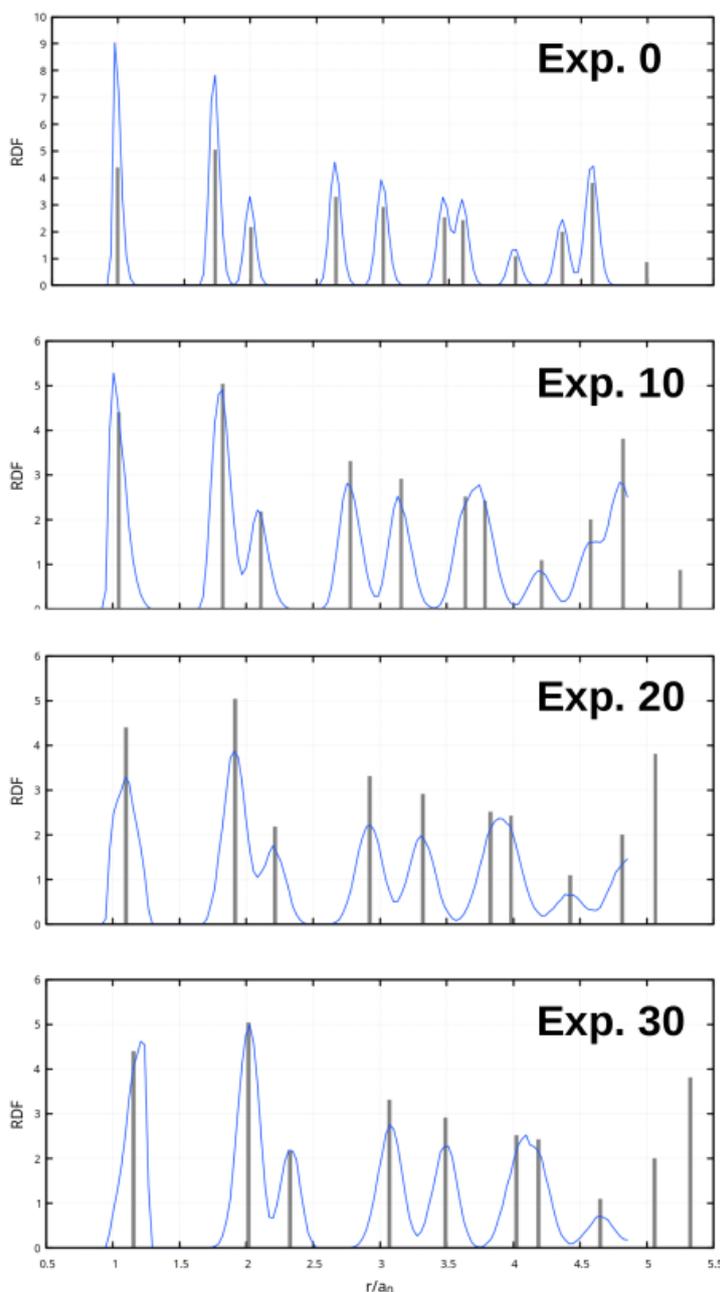
Figure 6.11: Radial distribution function, $g(r)$, before rupture process for the coupling constant of the angular interaction, $J_0 = 16\ U_{min}$. The first panel, i.e., Exp.0, shows $g(r)$ for a honeycomb structure in thermal equilibrium with bath. Next panels present the evolution of $g(r)$ in terms of the expansions. The position of peaks of a perfect honeycomb lattice are shown as the reference by the solid gray bars. The position of peaks follows the ideal honeycomb structure. The stretching process elongates bonds that are matched to the current accessible area. The first peak corresponds to the nearest neighbor shell, the second peak to the second one, etc. Presented data of $g(r)$ is the mean value taken over 20 simulations.

The evolution of energy and bond orientational order parameter, $|\psi_6|$, of the film during the stretching process, taken over 20 simulations, are shown in figure 6.12. The purple dots and error bars represent the mean and the standard error of the mean energy of the system at the end of each expansion-relaxation process. Before 40 and after 48 expansions, the energy increases consistently with the energy change of increasing the radial distance between the particles. The sharp increment in the energy profile, started at the 42nd expansion, is the dominant sign of fracture. Figure 6.13 illustrates the energy evolution in terms of the Monte Carlo sweeps during each expansion process of 40-50, assuring that despite causing a rupture, the system reaches a steady state (Fig. 6.12).
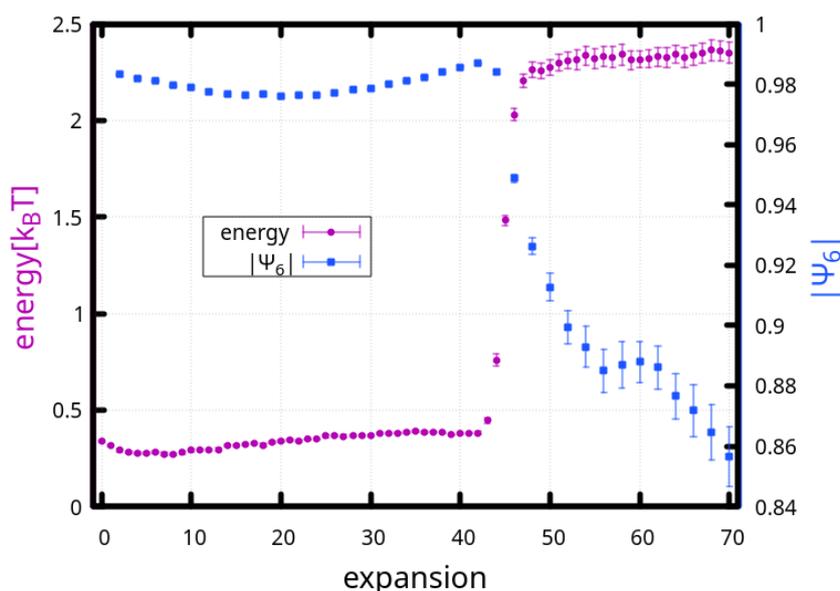
Figure 6.12: The bond orientational order parameter, $|\psi_6|$, and the energy of a film in thermal equilibrium with a bath are shown in blue squares and purple dots, respectively. The coupling constant of the angular interaction, J_0, is $16\ U_{min}$. The system undergoes several expansions, each of which follows a relaxation process. By the end of each expansion-relaxation process, the system reaches its stationary state, as shown for expansion 42-50 in figure 6.13. The sharp increment in the energy diagram starting at the 42nd expansion is the signature of the rupture process. The energy increment before and after the rupture follows the stretching. $|\psi_6|$ shows a peculiar feature; the initial reduction of the order parameter follows by an increment before the fracture. The existence of a local maximum just before the rupture indicates the critical role of the angular interaction in stabilizing the crystal structure in this range. Dots and error bars represent the mean and standard error of the mean, SEM, taken over 20 simulations.
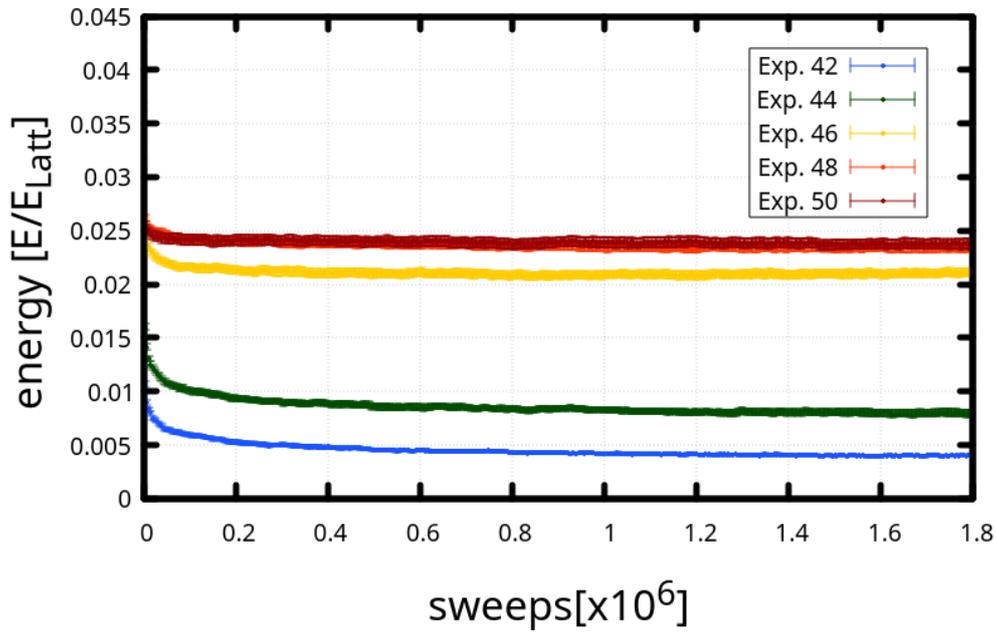
Figure 6.13: Energy evolution in terms of the MC sweeps.

The blue dots, purple squares, and corresponding error bars represent the mean and standard error of the mean bond-orientational order parameter, $|\psi_6|$, and the energy of the system in terms of expansion, respectively. $|\psi_6|$ does not show a monotonic behavior. It first decreases, passes through a local minimum, increases symmetrically, reaches a maximum just before rupture, and decreases monotonically. A close examination of both energy and $|\psi_6|$ indicates the critical role of angular interaction just before the fracture; despite the expansion, the energy remains constant during expansions 39-42. At the same time, $|\psi_6|$ increases to a maximum just before the rupture, see Fig. 6.12, indicating that the angular interaction is the dominant interaction in this range and is responsible for maintaining the ordered structure of the system.

6.3.2 Disordered phase
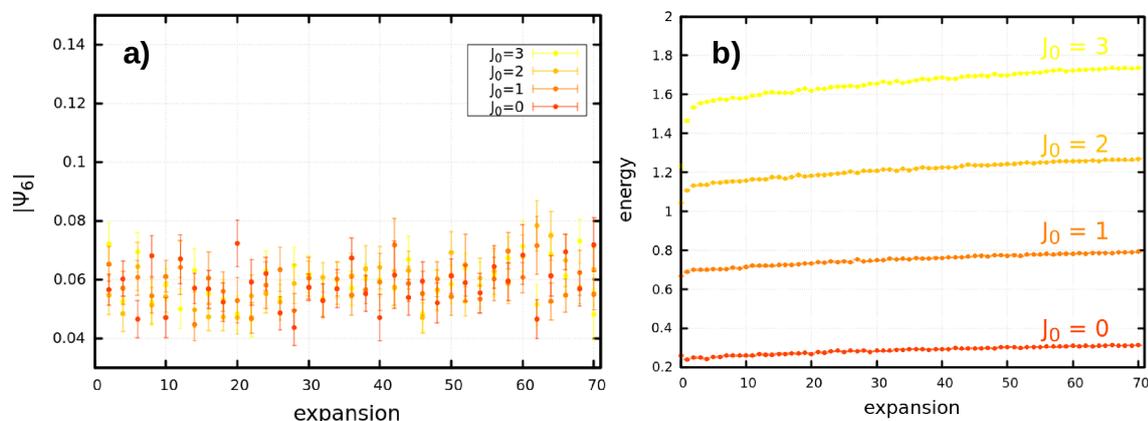
Figure 6.14: (a) Bond-orientational order parameter, $|\psi_6|$, of system with coupling constant, J_0, less than four. A low value of $|\psi_6|$ implies that the system is disordered. **(b)** Evolution of energy due to stretching. In all cases, energy increases smoothly, corresponding to enlarging the system. The individual value of J_0 is reported in the legend. Identical colors, as the panel **(a)**, are used to present the energy evolution of the system.

The simulation results indicate a disordered phase for a coupling constant of less than $4\ U_{min}$. The bond-orientational order parameter measured for these coupling constants is low, as shown in figure 6.14-a. The evolution of energy due to stretching, shown in figure 6.14-b, behaves similarly for all J_0 in this range; it increases smoothly as the area of the film is stretched. However, in an identical expansion, the energy required to break the angular bond of a system of a higher coupling constant is higher; hence, the system gains more energy for larger values of J_0. A system of zero angular coupling constant interaction well-located in this regime is discussed in the following.
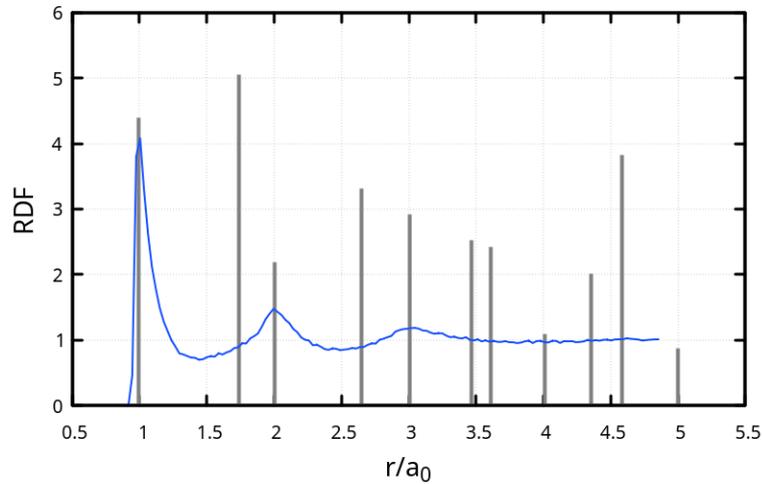
Figure 6.15: Radial distribution function, $g(r)$, of a system before expansion. The angular coupling constant, J_0, is set to zero. Appearing a primary peak followed by a shorter secondary is the feature of a liquid phase. The position of peaks of an ideal honeycomb lattice is shown as a reference by the solid gray bars.
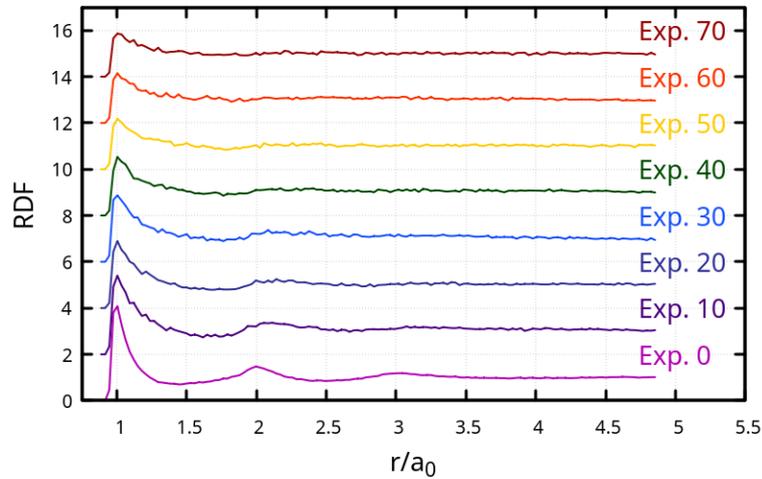
Figure 6.16: Radial distribution function, $g(r)$, of a system in different expansions process. The coupling constant, J_0, is set to zero. Appearing a primary peak followed by a shorter secondary is the feature of a liquid phase. The second peak is flattened at higher expansions, and the system behaves as a disordered gas phase. The plots are shifted vertically by two for better visualization. The individual number of expansions is reported in identical colors to the plot.

The radial distribution function, $g(r)$, of zero coupling constants before expansion is displayed in figure 6.15. The valley between the first and second peaks is not deep.

After the second peak, $g(r)$ goes uniformly to one, which is the significant feature of a disordered liquid phase. At higher expansions shown in figure 6.16, the height of the first peak decreases and shifts to larger r. The second peak is not pronounced anymore. Here, the plots are shifted vertically by two for better visualization.

The evolution of local packing investigated by the nearest neighbor visualization is shown in figure 6.17. The disordered phase after the second expansion is apparent. At higher expansions, the accessible area per particle increases; hence, the number of free particles, shown in red, increases dramatically.
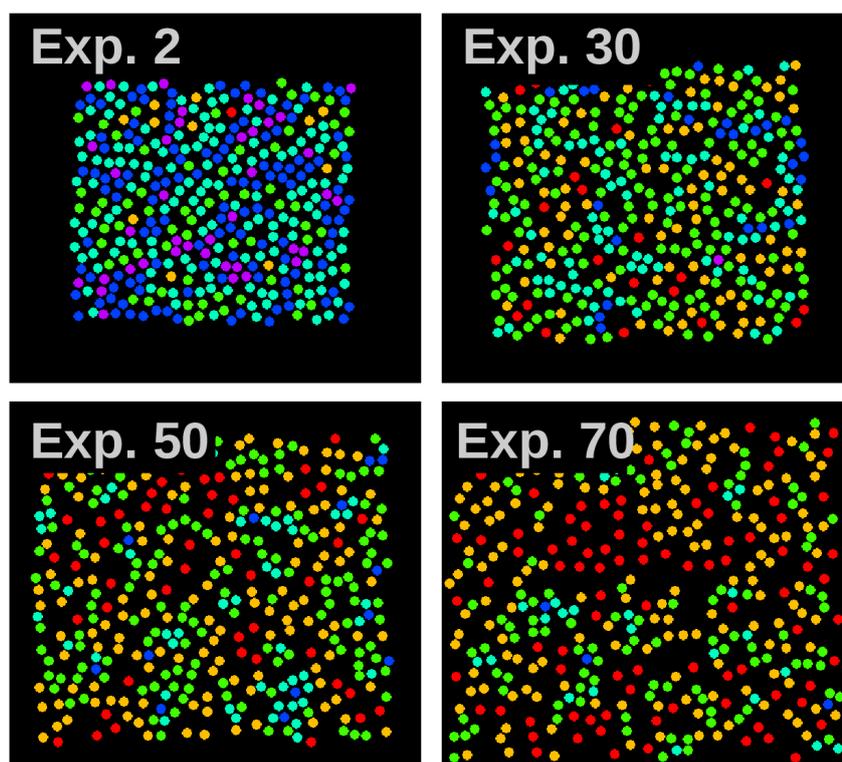
Figure 6.17: The nearest neighbor visualization of a stretching system. The initial state was the fluctuating honeycomb-ordered system. The coupling constant of the angular interaction, J_0, is set to zero. Each expansion follows a relaxation process. The system is disordered. At higher expansions, the accessible area per particle increases; hence the number of free particles, shown in red, increases dramatically. Each color represents a particle with a certain number of neighbors. Red, dandelion, green, cyan, blue, purple, and orange represents particles with 0 to 6 neighbors, respectively.

6.3.3 Intermediate phase
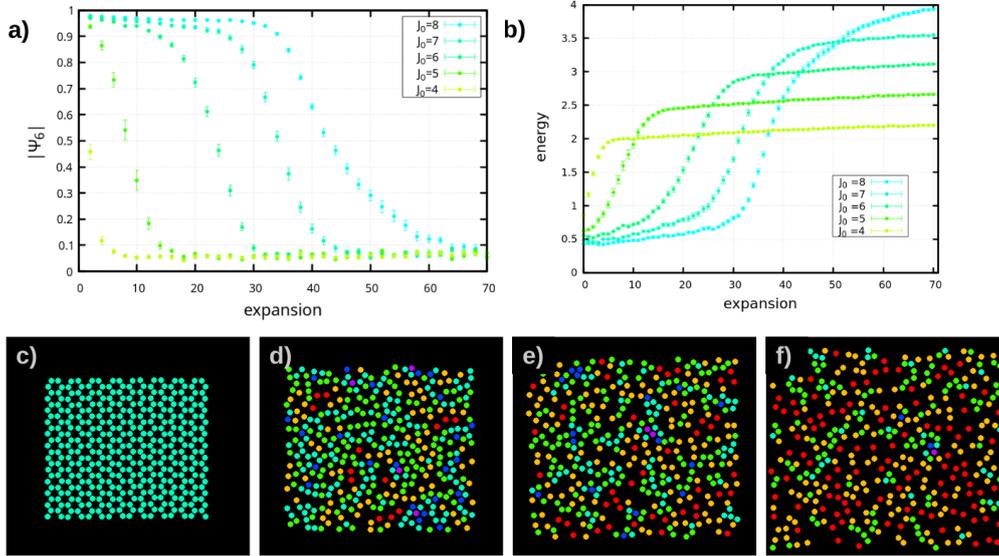
Figure 6.18: Intermediate phase. Panels **(a)** and **(b)** show the evolution of the bond-orientational order parameter, $|\psi_6|$, and the energy of a system with coupling constant, J_0, of $4 - 8$ U_{min} in terms of expansions, respectively. The initially high value of $|\psi_6|$ differentiates this phase from the disordered phase. However, $|\psi_6|$ decreasing monotonically due to expansions, reordering is not observed for this phase. The increment of the energy is not sharp compared to the ordered phase. Dots and error bars represent the mean and standard error of the mean, (SEM), taken over 20 simulations. The individual values of J_0 are reported in the legends. Identical colors, as in the panel **(a)**, are used to present the energy evolution of the system. Panels **(c)-(f)** show the evolution of the nearest neighbor visualization of the system after **(c)** 2, **(d)** 30, **(e)** 50, and **(f)** 70 expansions. In panels **(c)-(f)**, each color represents a particle with a certain number of neighbors. Red, dandelion, green, cyan, blue, purple, and orange represents particles with 0 to 6 neighbors, respectively.

The measured bond-orientational order parameter, $|\psi_6|$, of the intermediate phase illustrated in figure 6.18-a, distinguishes this phase from disordered and ordered phases; A high value of $|\psi_6|$ at the early stages differentiates this phase from the disordered phase. However, $|\psi_6|$ decreases monotonically, implying that the angular interaction is insufficient to maintain the crystal structure. Furthermore, the reordering of the system, as visualized by a local maximum followed by a monotonically decreasing of $|\psi_6|$ in the ordered phase, can not be seen in this phase. The evolution of energy with expansions is shown in panel 6.18-b. The increment in the energy is not sharp compared to the ordered phase, implying that a rupture in the system is not expected. The nearest neighbor visualization of a system with coupling angular

interaction constant, $J_0 = 5 \ U_{min}$, shown in 6.18-(c-f), indicates an ordered phase at the beginning. However, the system becomes disordered after 30 expansions in panel d.

In conclusion, the ordered regime in our results is determined by two distinct features: A characteristic behavior of the order parameter, $|\psi_6|$, accompanied by a sharp jump in the energy quantified in terms of expansions. On the opposite extreme, the disordered regime is characterized by the energy that varies smoothly with the stretching process. The sharp variation in the energy, a representative sign of rupture, is not seen. Moreover, a very low bond orientational order parameter value with the characteristic plot of $g(r)$ quantifies this phase as a disordered-liquid phase.

6.4 Hexagonal structure
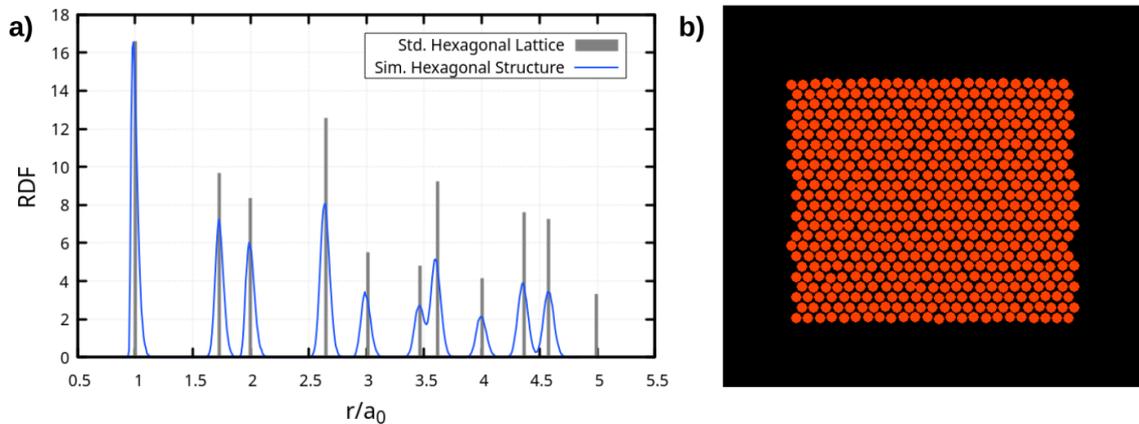
Figure 6.19: Radial distribution function (RDF), $g(r)$, and the nearest neighbor visualization of a hexagonal structure. **(a)** $g(r)$ of a hexagonal system in equilibrium with a thermal bath is shown in blue solid lines. The gray peaks show $g(r)$ of an ideal hexagonal lattice. The position and the height of peaks in both structures are consistent. The plot represents the mean value taken over 20 simulations. **(b)** The nearest neighbor visualization shows a hexagonal lattice in equilibrium with the thermal bath. In panel **(b)**, each color represents a particle with a certain number of neighbors. Red, dandelion, green, cyan, blue, purple, and orange represents particles with 0 to 6 neighbors, respectively. All particles are orange, meaning that six neighbors surround each particle.

The honeycomb lattice observed experimentally for the membrane of class II hydrophobins is a particular case of the hexagonal lattice with a two-atom basis. In this section, to complete our study, I present the development of our model for a hexagonal arrangement. Thus, the initial state differs from the honeycomb by having a hexagonal structure in thermal equilibrium with a bath. Furthermore, the interactions between particles are considered isotropic in the framework of DLVO theory. We do not assume any angular interactions between particles. Hence, comparing the results obtained for a honeycomb structure and the general hexagonal lattice will light up the role of the postulated angular interaction in the stability of the film.

The initial state of the system is characterized by the radial distribution function, $g(r)$, and the nearest neighbor visualization as shown in figure 6.19, panels (a) and (b), respectively. The solid blue curve in panel (a) illustrates $g(r)$ for a system of isotropic-interaction particles in a hexagonal structure, followed by the position of the peaks of a perfect hexagonal lattice shown in solid gray-vertical lines in this panel. $g(r)$ of our

hexagonal structure compared to the ideal lattice shows a shorter height of peaks with variance due to thermal fluctuations.
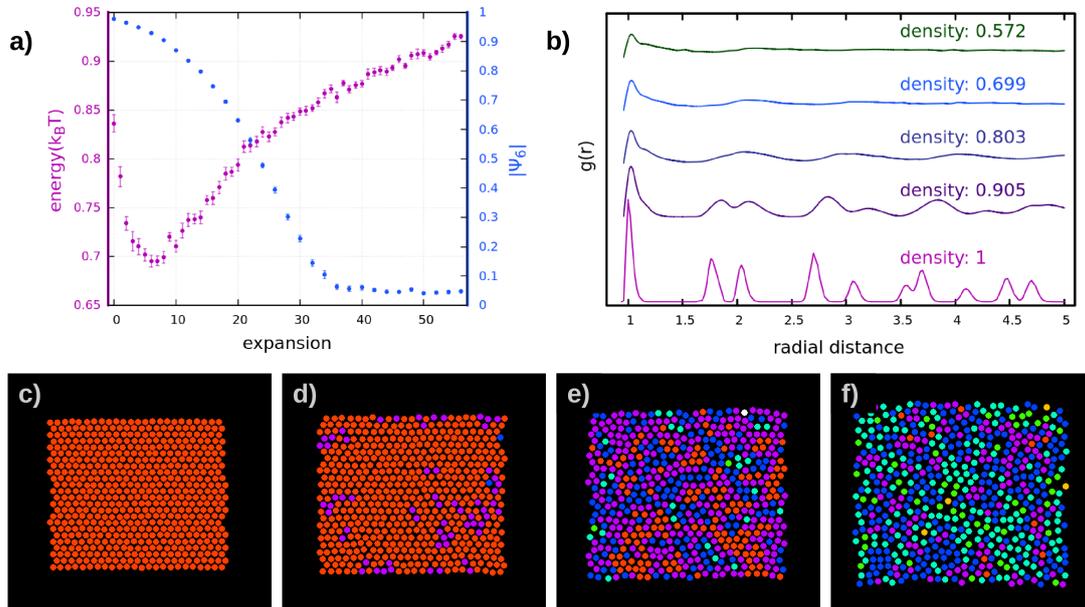
Figure 6.20: The hexagonal structure under the stretching. **(a)** The bond-orientational order parameter, $|\psi_6|$, shown in blue decreases due to stretching the system. The purple dots present the energy of the system. The change in energy is matched by the disordered regime observed for a honeycomb structure, see 6.3.2. Dots and error bars represent the mean and standard error of the mean, SEM, taken over 20 simulations. **(b)** The radial distribution function, $g(r)$, of a hexagonal structure at different stages of stretching. The individual density reported at identical colors represents the number density after 0, 10, 22, 36, and 56 expansion-relaxation processes. The plots are shifted vertically for better clarification. **(c-f)** The nearest neighbor visualization of the system. Panel **(c)** shows the initial equilibrium state of the system of density one. Six neighbors well surround each particle. **(d-f)** By the stretching process, the density of the system decreases. The following panels **(d-f)** correspond to the stretched systems of fewer densities, i.e., the number density in each panel, ρ, is 0.905, 0.803, and 0.699, respectively. In panels **(c)-(f)**, each color represents a particle with a certain number of neighbors. Red, dandelion, green, cyan, blue, purple, and orange represents particles with 0 to 6 neighbors, respectively.

The bond-orientational order parameter, $|\psi_6|$, is high at the initial state, shown by blue dots in figure 6.20-a, presenting the closed-packed ordered system. However, the order is not stable in this system; The stress caused by stretching decreases $|\psi_6|$ monotonically. The local minimum representative signature of reordering in the system, see 6.3.1, is not observed. The evolution of the energy after the expansion-relaxation process shown in purple dots in the plot is comparable with the results

obtained in the disordered regime of the initial honeycomb structure: A global minimum followed by a monotonical energy increase due to enlarging bonds connecting particles observed. The evolution of $g(r)$, shown in panel (b), illustrates converting the system from ordered to disordered. The initial state is ordered, as evidenced by the well-defined peaks; however, these peaks are no longer pronounced after stretching the system. The individual number density of the system is written in identical colors on the plot. The number of expansions matched with a specific density shown in the figure is presented in the caption. The plots are shifted vertically for better visualization. The transition from ordered to disordered is observed in the nearest neighbor visualization shown in panels (c) to (f).

6.5 Defect structure

The mechanical properties of the film can be affected significantly due to defects in such systems. Defects are an indispensable part of systems in equilibrium. In general, in a 2D material, two types of defects exist: (i) point defects, such as voids and topological defects introduced in 4.3.3, and (ii) line defects, such as grain boundaries. Here, as a point defect, we investigate the influence of the voids in the film on the maximum coupling constant of angle-dependent potential, i.e., $J_0 = 16\, U_{min}$.
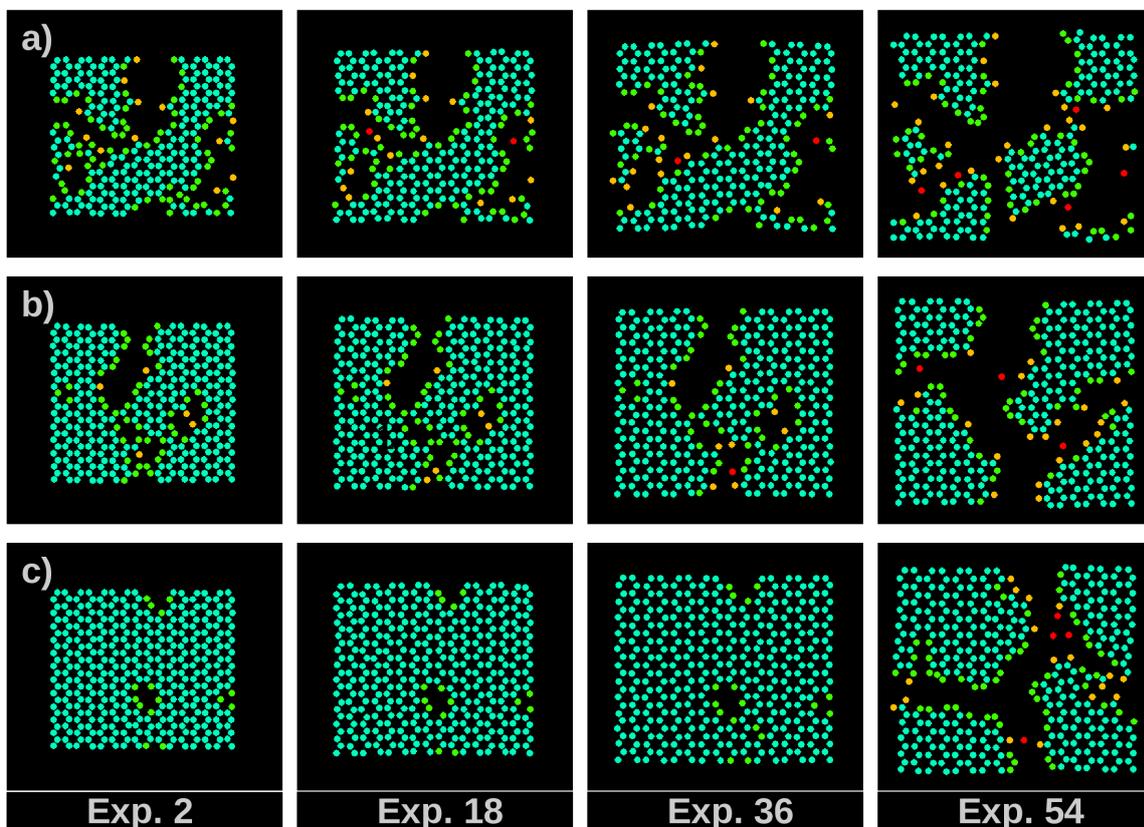
Figure 6.21: Nearest neighbor visualization of a defect film. The position with a lower density, i.e., defect in the system, is sensitive to tension. The defects are the core of starting fractures in the film. This figure shows the evolution of films of different initial number densities, ρ_0. Initial number density: **(a)** $\rho_0 = 0.716$, **(b)** $\rho_0 = 0.872$, **(c)** $\rho_0 = 0.976$. Each column represents a certain stage of expansion reported at the bottom.

Due to voids in the film, the initial number density, defined as $\rho_0 = \frac{N}{N_0}$, is not one, where N and N_0 denote the number of particles on film, and the number of available sites on an ideal honeycomb lattice of the same length, respectively. An identical protocol of expansion-relaxation process is used, i.e., the initial state of the system

is a film of the honeycomb structure in thermal equilibrium with the bath. The only difference is that the film is incomplete and consists of randomly distributed voids. The nearest neighbor visualization, shown in figure 6.21, reveals a rupture in the system. The position of voids was determined stochastically; however, the simulation results show that the defect point is the core of starting fracture in the system.

The evolution of energy due to stretching, shown in figure 6.22, with a sharp energy variation, confirms the fracture in the system under tension by appearing a dominant jump in the energy panel. The results show that a more homogenous film with fewer defects resists more stress, as evident by a higher jump in the energy panel for higher initial densities.
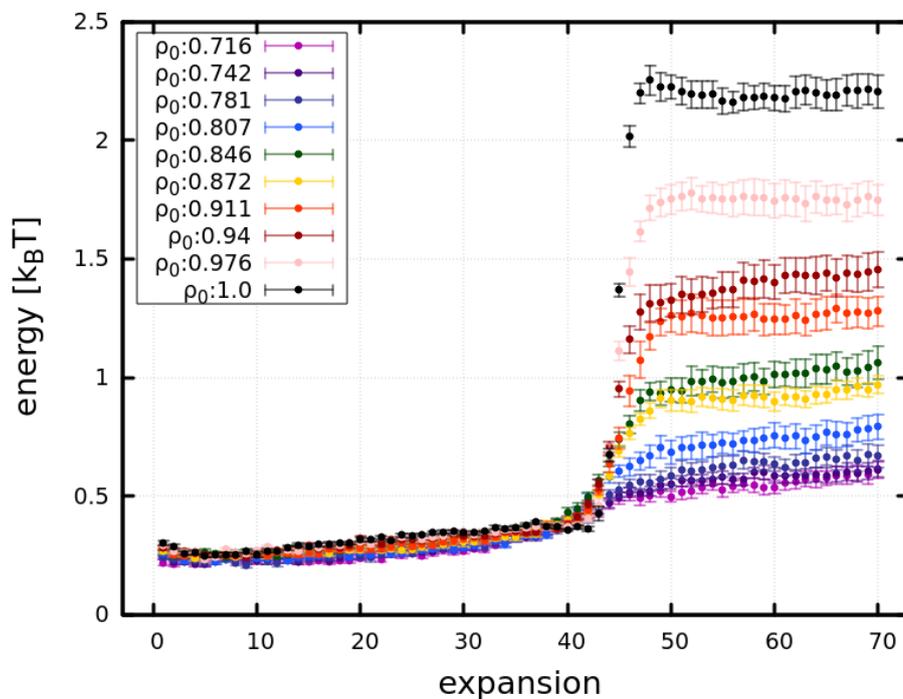
Figure 6.22: Energy profile in terms of expansion of a defect film. The sharp variation in the energy profile after expansion of 40 indicates a rupture in the system. The energy required to cause a rupture in the system increases for higher initial number densities, ρ_0. Dots and error bars represent the mean and standard error of the mean, SEM, taken over 20 simulations. The individual number densities are reported in the legend.

6.6 Disorder to order transition
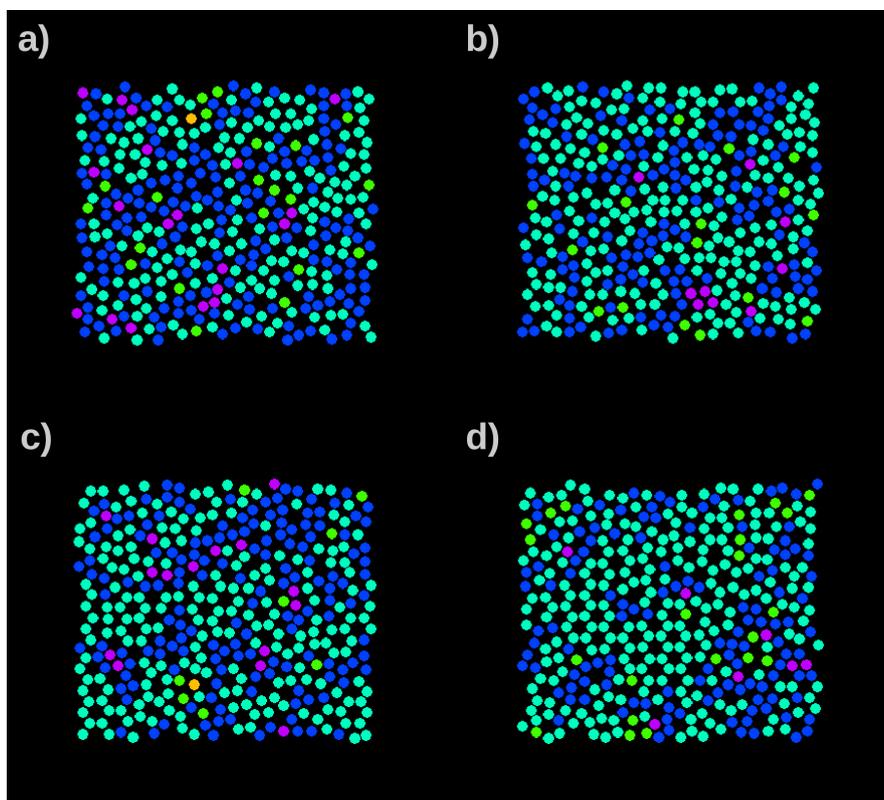
Figure 6.23: Disordered to ordered transition. The initial state is a disordered phase of particles interacting via isotropic interactions. The system is in thermal equilibrium with a heat bath. To study whether the transition can be obtained, we abruptly switched the coupling constant of angular interaction, J_0, from zero to the maximum studied value, i.e., $16.0\ U_{min}$. Panels **(a)-(d)** show the evolution of the system examined by the nearest neighbor visualization after **(a)** 2×10^6, **(b)** 20×10^6, **(c)** 80×10^6, **(d)** 96×10^6 sweeps. Each color represents a particle with a certain number of neighbors. Red, dandelion, green, cyan, blue, purple, and orange represents particles with 0 to 6 neighbors, respectively.

In this section, I aim for the ordered phase to be obtained starting from a disordered phase. The initial model state is a system of particles interacting in the framework of DLVO. At the initial state, I ignore the angular interaction between particles, meaning that the coupling constant of angular interaction, J_0, is set to zero, and the disordered phase is in thermal equilibrium with a heat bath. In section 6.3, I concluded that the angular interaction stabilizes the crystal structure; thus, I investigate if the angular interaction might be sufficient to order the system. Hence, the coupling constant parameter is set abruptly to the maximum studied value, i.e., $J_0 = 16.0\ U_{min}$. Afterwards, the system is let to find the optimum energy structure through Monte Carlo

sweeps. The nearest neighbor visualization, shown in 6.23, indicates the presence of separated clusters with a honeycomb structure. The area of the separated cluster increased by increasing the simulation time, i.e., the number of Monte Carlo sweeps and isolated clusters percolated in panel d. Despite the crystalline structure in the layer, the measured bond-orientational order parameter, $|\psi_6|$, states the resulting structure is not an ordered phase, as shown in figure 6.24. Unfortunately, comparable experimental data for hydrophobins have not been reported at the current time. However, this problem has been observed experimentally in the free-standing monolayer of graphene, which has an identical honeycomb structure [36, 45, 55, 68]. Several crystallographic defects are seen in the configurations shown in figure 6.23. In the case of graphene, reported studies revealed that the activation energy compared to the formation energy of these defects is significantly high [36, 55]; hence, restoring a perfect honeycomb structure is challenging with a relatively high restoration barrier energy. These are stable topological defects in the system, which, in the case of graphene, their formation is the termination step for graphene growth.
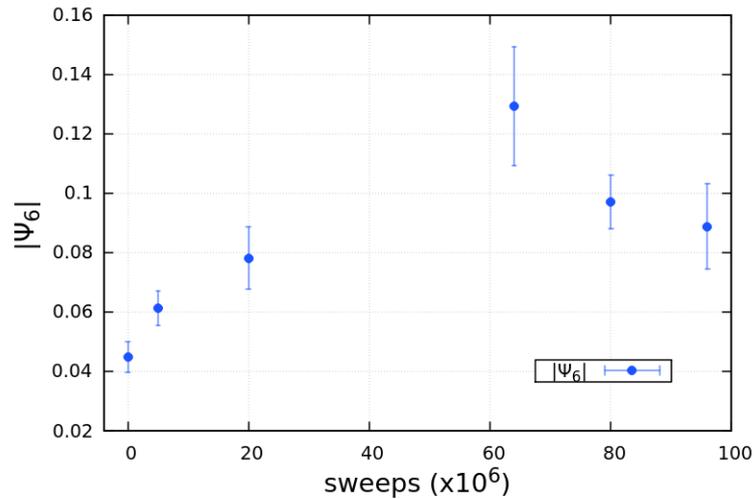
Figure 6.24: Evolution of bond orientational order parameter, $|\psi_6|$, in terms of the Monte Carlo sweeps. The system starts from a disordered phase without considering angular interaction. Afterward, the potential angular interaction is considered and set to the maximum studied value of $J_0 = 16\ U_{min}$. $|\psi_6|$ is measured at different stages of simulation. Dots and error bars represent the mean and standard error of the mean, SEM, taken over 20 simulations.

Theoretically, choosing a different approach may lead to the perfect ordered phase. In these systems, two stable states are separated by an energy barrier, as reported for graphene. Both have low free energy but high and low values in an order parameter, e.g., bond-orientational order parameter, $|\psi_6|$, in this case. Because of the energy landscape, the standard Boltzmann-weighted sampling cannot produce a Markov chain in which the ergodicity is fulfilled. A solution would be using umbrella sampling, an efficient technique to overcome poor sampling caused by the potential barrier. In this method, the potential energy is replaced by a chosen biasing potential, e.g., a function of the order parameter of the system that flattened the present energy barrier. Generating this potential and using the umbrella sampling method is beyond the scope of this thesis. However, the results of the present study can be used as a starting point for future works.

6.7 Summary and conclusions

The hydrophobins represent a class of proteins self-aggregated at air-water interfaces. The interfacial formed film possesses an extraordinarily high shear elasticity and mechanical strength, a testament to the remarkable properties of these proteins. The developed film of HFBI and HFBII proteins shows a honeycomb structure, a particular form of a two-dimensional hexagonal structure known for its high buckling strength. This observation leads to two questions: first, which kind of interactions result in the observed honeycomb structure, and second, if the honeycomb structure dictates the elastic behavior of the films. To answer these questions, I introduced a novel approach, a stochastic model in continuum space. This unique model allows for a deeper understanding of the behavior of the system.

In this model, hydrophobins are considered soft particles interacting in the framework of DLVO theory. Furthermore, I emphasized the crucial role of the honeycomb structure, which is guaranteed by a lateral interaction between proteins. This lateral interaction, a short-ranged angular interaction, is considered beneficial to a configuration in which the adjacent bonds make a certain angle, i.e., 120 degrees. This finding underscores the importance of the lateral interaction in stabilizing the honeycomb structure. Furthermore, a quasi-static stretching process is designed to investigate the mechanical response of the system. The bond orientational order parameter quantifies the response of the system. Moreover, the structure of the system is visualized.

The measured order parameter, shown in figure 6.25, suggests that a high enough lateral interaction between proteins is essential to stabilize the ordered honeycomb structure. This finding further emphasizes the importance of the honeycomb structure in the system's behavior. The feature of the ordered phase, shown in green, is reordering the system detected by a local maximum in the order parameter. In the hexagonal structure, where the angular interaction is switched off, the order parameter decreases monotonically as shown in orange. Visualizing the structure confirms the above. Due to the short-ranged angular interaction, the stable domains in the film were observed. The short-ranged interaction also drives a rupture in the film at high enough tension. This demonstrates the significant impact of the honeycomb structure on the mechanical response of the films.

Moreover, the effect of the point defects, i.e., voids, on the mechanical properties of the films is studied. The results revealed the significant behavior of an ordered

phase at an identical strength of angular interaction; however, the voids are the weak points of the film, which rupture starts from it.
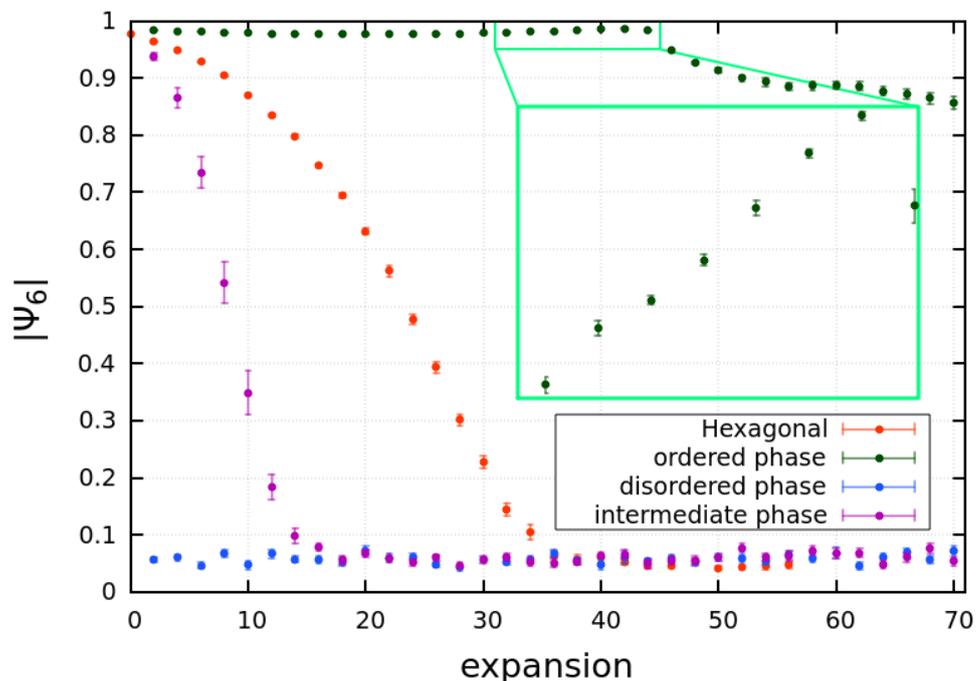
Figure 6.25: Representative behavior of different systems. Bond-orientational order parameter, $|\psi_6|$, of initially honeycomb and hexagonal structures are shown. For the honeycomb structure, depending on the value of the constant parameter, J_0, three phases are observed: (i) the ordered phase, shown in green. The representative sign of this phase is reordering the structure just before rupture, as shown in the zoom-in diagram. (ii) The disordered phase with a low value of $|\psi_6|$ is presented in blue. (iii) The intermediate phase is shown in purple. This phase is differentiated from the disordered phase by the initially high value of $|\psi_6|$. The monotonic decrease of the order parameter distinguishes this phase from the ordered phase. The hexagonal structure, shown in orange, is studied as a reference structure without including angular interactions. Dots and error bars represent the mean and standard error of the mean, SEM, taken over 20 simulations.

Finally, I aimed to investigate if a disordered phase would transform into an ordered phase by considering the angular interaction. The initial state of the system was an entirely disordered phase. The strength of the angular interaction abruptly switched to the highest value, and the system searched for the optimal configuration via Monte Carlo movement. The simulation results show the formation of domains. However, as observed experimentally, the domains cannot fuse to make a mono

structure, and the ordered phase could not be achieved. Even though the failure of the procedure due to stable topological defects appeared in the film, the result could be important for modeling the transition by a different simulating approach.

Our result would also be significant in investigating phase transition in two-dimensional systems. Two-dimensional melting can be a first or second-order transition depending on the interactions of the system. Studying two-dimensional melting is out of the scope of this thesis. Although we did not study two-dimensional phase transition in our study due to the small size of the system, the effect of the anisotropic interaction, such as the angular interaction we introduced here, has yet to be studied in the field. Hence, our results and modeling can be a starting point for future investigations.

Chapter 7

Summary and conclusions

In this thesis, I studied two aspects of films at the liquid-air interfaces by hydrophobins, i.e., the adsorption kinetics and the mechanical response of the thin film to external stress. The characterization of the adsorption kinetics has been performed using a combined approach of theoretical modelling and experiments. I modelled the adsorption kinetics and performed simulations to understand the underlying mechanism observed in experimental results obtained in the Jacobs group at Saarland University. In particular, the experiments and simulations studied selected proteins of class II hydrophobins and a few engineered variants of those proteins. Hydrophobins, characterized by their amphiphilic structure, self-aggregate at air-water interfaces. The time-dependent behaviour of the adsorbed amount, as measured using ellipsometry, revealed intriguing kinetics. Adsorption followed a linear pattern for the wild-type proteins until saturation, which abruptly ceased. Mutant hydrophobins exhibited different kinetics, with the bulky hydrophobins conforming to the Langmuir adsorption model and the charge-mutated protein exhibiting kinetics dependent on solution concentration.

A stochastic model was introduced to unravel the adsorption mechanism. This model considered the experimental setup to consist of three main components: the solution, containing a homogenous distribution of proteins, and two lattice layers, one at the air-water interface and the other in the subsurface. The subsurface layer was essential due to a secondary minimum close to the interface in colloidal systems, which leads to an energy barrier. The energy barrier must be overcome for adsorption. Proteins were treated as particles with an internal degree of freedom referred to as "spin," simulating the anisotropy of the protein patches caused by charged amino acids within the protein structure. The spin property influenced the pattern of the two-dimensional aggregation of proteins, i.e., favouring clusters with similar spin. Finally, the adsorption process was modelled as a rate-based statistical model, with

varying rates explaining the observed adsorption kinetics.

The simulation results showed that the adsorption process is a two-stage process: (i) a pre-reversible-adsorption to the subsurface and (ii) the final adsorption, which is irreversible. A high diffusive motion in the subsurface and a high enough final adsorption rate are two essential factors that result in the linear kinetics observed for wild-type proteins. In the bulky proteins, the diffusive motion is limited; hence, available free sites can not be found efficiently, and the adsorption kinetics becomes Langmuir type with a coverage-dependent behaviour.

The model also elucidated the origin of stable two-dimensional roundish clusters observed in films formed by wild-type proteins. The spin property influenced cluster formation such that the shape transitioned from dendritic to roundish, depending on the strength of the spin interaction. Hence, the larger proteins hindered close affinity, and short-range spin interactions facilitated cluster diffusion and formation into a single cluster.

In the second part of this study, I delved into the mechanical response of the film under quasi-static stretching. Unlike the adsorption process, which operates on longer time scales, the mechanical response requires consideration of shorter time scales. Consequently, the lattice model was replaced by a continuum space model incorporating thermal fluctuations. An angle-dependent potential was also introduced to account for the anisotropy arising from the position of amino acids within the protein. This angular interaction played a crucial role in maintaining the honeycomb structure in the system.

Simulation results demonstrated that the angle-dependent potential stabilized and maintained order in the system. Here, quasi-static stretching of the system was followed by a relaxation process, revealing the role of the angle-dependent potential. The system underwent a rupture for a sufficiently high angular interaction constant. However, the evolution of the order parameter showed a remarkable feature: during the stretching process, the order parameter first decreased and then increased just before rupture, and after that, it decreased monotonically. This reordering process before rupture occurred without any change in the energy profile, emphasizing the dominance of the angle-dependent potential in stabilizing the structure.

Comparatively, the ordered structure could not be maintained in systems of particles initially arranged in a hexagonal structure and interacting isotropically, i.e., only DLVO interactions. The results indicated that purely isotropic interactions couldn't create a stable crystal structure. Hence, the angle-dependent potential introduced in this thesis was identified as the primary contributor to explaining the crystal structure in

the studied hydrophobins.

Furthermore, my Simulation also showed that the growth of an ordered parameter is challenging to achieve for an off-lattice model. Here, local defects are persistent and perturb the development of large-scale ordered parameters.

Bibliography

[1] N. A. Alexandrov, K. G. Marinova, T. D. Gurkov, K. D. Danov, P. A. Kralchevsky, S. D. Stoyanov, T. B. Blijdenstein, L. N. Arnaudov, E. G. Pelan, and A. Lips. Interfacial layers from the protein hfbii hydrophobin: Dynamic surface tension, dilatational elasticity and relaxation times. *Journal of colloid and interface science*, 376 (1):296–306, 2012.

[2] N. W. Ashcroft, N. D. Mermin, et al. Solid state physics, 1976.

[3] V. Berezinskii. Destruction of long-range order in one-dimensional and two-dimensional systems having a continuous symmetry group i. classical systems. *Sov. Phys. JETP*, 32(3):493–500, 1971.

[4] V. L. Berezinsky. Destruction of Long-range Order in One-dimensional and Two-dimensional Systems Possessing a Continuous Symmetry Group. II. Quantum Systems. *Sov. Phys. JETP*, 34(3):610, 1972.

[5] E. P. Bernard and W. Krauth. Two-step melting in two dimensions: First-order liquid-hexatic transition. *Phys. Rev. Lett.*, 107: 155704, Oct 2011. doi: 10.1103/PhysRevLett.107.155704. URL `https://link.aps.org/doi/10.1103/PhysRevLett.107.155704`.

[6] P. M. Chaikin and T. C. Lubensky. *Principles of Condensed Matter Physics*. Cambridge University Press, 1995. doi: 10.1017/CBO9780511813467.

[7] M. A. Clark and J. Choi. [etextbook] biology-2e, 2018.

[8] A. R. Cox, F. Cagnol, A. B. Russell, and M. J. Izzard. Surface properties of class ii hydrophobins from trichoderma reesei and influence on bubble stability. *Langmuir*, 23(15):7995–8002, 2007.

[9] M. L. de Vocht, I. Reviakine, H. A. Wosten, A. Brisson, J. G. Wessels, and G. T. Robillard. Structural and functional role of the disulfide bridges in the hydrophobin sc3. *Journal of Biological Chemistry*, 275(37):28428–28432, 2000.

[10] L. Devroye. An automatic method for generating random variates with a given characteristic function. *SIAM journal on applied mathematics*, 46(4):698–719, 1986.

[11] C. Domb and R. Potts. Order-disorder statistics iv. a two-dimensional model with first and second interactions. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 210(1100):125–141, 1951.

[12] C. Garidiner. Handbook of stochastic methods: For physics. *Chemistry and the Natural Sciences Springer*, 1996.

[13] U. Gasser, C. Eisenmann, G. Maret, and P. Keim. Melting of crystals in two dimensions. *ChemPhysChem*, 11(5):963–970, 2010.

[14] L. Gazzera, C. Corti, L. Pirrie, A. Paananen, A. Monfredini, G. Cavallo, S. Bettini, G. Giancane, L. Valli, M. B. Linder, et al. Hydrophobin as a nanolayer primer that enables the fluorinated coating of poorly reactive polymer surfaces. *Advanced Materials Interfaces*, 2(14):1500170, 2015.

[15] M. F. Gebbink, D. Claessen, B. Bouma, L. Dijkhuizen, and H. A. Wösten. Amyloids—a functional coat for microorganisms. *Nature Reviews Microbiology*, 3(4): 333–341, 2005.

[16] M. A. Gibson and J. Bruck. Efficient exact stochastic simulation of chemical systems with many species and many channels. *The journal of physical chemistry A*, 104(9):1876–1889, 2000.

[17] D. T. Gillespie. Stochastic simulation of chemical kinetics. *Annu. Rev. Phys. Chem.*, 58:35–55, 2007.

[18] H. Hähl. *Proteinadsorption an Grenzflächen: Bestimmung von Adsorptionskinetik und Adsorbatstruktur mittels Reflexionsmethoden*. Logos-Verlag, 2012.

[19] H. Hähl, J. N. Vargas, A. Griffo, P. Laaksonen, G. Szilvay, M. Lienemann, K. Jacobs, R. Seemann, and J.-B. Fleury. Pure protein bilayers and vesicles from native fungal hydrophobins. *Advanced Materials*, 29(1):1602888, 2017.

[20] H. Hähl, A. Griffo, N. Safaridehkohneh, J. Heppe, S. Backes, M. Lienemann, M. Linder, L. Santen, P. Laaksonen, and K. Jacobs. Dynamic assembly of class ii hydrophobins from t. reesei at the air-water interface. *Langmuir*, 35(28):9202–9212, June 2019. ISSN 0743-7463. doi: 10.1021/acs.langmuir.9b01078.

[21] J. Hakanpää, M. Linder, A. Popov, A. Schmidt, and J. Rouvinen. Hydrophobin hfbii in detail: ultrahigh-resolution structure at 0.75 å. *Acta Crystallographica Section D: Biological Crystallography*, 62(4):356–367, 2006.

[22] J. Hakanpaää, A. Paananen, S. Askolin, T. Nakari-Setaäläa, T. Parkkinen, M. Penttilaä, M. B. Linder, and J. Rouvinen. Atomic resolution structure of the hfbii hydrophobin, a self-assembling amphiphile. *Journal of Biological Chemistry*, 279(1): 534–539, 2004.

[23] H. Hahl, J. N. Vargas, M. Jung, A. Griffo, P. Laaksonen, M. Lienemann, K. Jacobs, R. Seemann, and J.-B. Fleury. Adhesion properties of freestanding hydrophobin bilayers. *Langmuir*, 34(29):8542–8549, 2018.

[24] K. Huang. *Statistical Mechanics*. John Wiley & Sons, 2 edition, 1987.

[25] J. J. Joensuu, A. J. Conley, M. Lienemann, J. E. Brandle, M. B. Linder, and R. Menassa. Hydrophobin fusions for high-level transient protein expression and purification in nicotiana benthamiana. *Plant physiology*, 152(2):622–633, 2010.

[26] L. P. Kadanoff. *Statistical physics: statics, dynamics and renormalization*. World Scientific, 2000.

[27] S. C. Kapfer and W. Krauth. Two-dimensional melting: From liquid-hexatic coexistence to continuous transitions. *Phys. Rev. Lett.*, 114: 035702, Jan 2015. doi: 10.1103/PhysRevLett.114.035702. URL `https://link.aps.org/doi/10.1103/PhysRevLett.114.035702`.

[28] J. Kosterlitz. The critical properties of the two-dimensional xy model. *Journal of Physics C: Solid State Physics*, 7(6):1046, 1974.

[29] J. M. Kosterlitz and D. J. Thouless. Ordering, metastability and phase transitions in two-dimensional systems. *Journal of Physics C: Solid State Physics*, 6(7):1181–1203, apr 1973. doi: 10.1088/0022-3719/6/7/010. URL `https://doi.org/10.1088/0022-3719/6/7/010`.

[30] L. P. Kozlowski. Proteome-pi: proteome isoelectric point database. *Nucleic acids research*, 45(D1):D1112–D1116, 2017.

[31] W. Krauth. *Statistical mechanics: algorithms and computations*, volume 13. OUP Oxford, 2006.

[32] R. Kumar, S. Singh, and O. V. Singh. Bioconversion of lignocellulosic biomass: biochemical and molecular perspectives. *Journal of industrial microbiology and biotechnology*, 35(5):377–391, 2008.

[33] I. Langmuir. The adsorption of gases on plane surfaces of glass, mica and platinum. *Journal of the American Chemical society*, 40(9):1361–1403, 1918.

[34] T. D. Lee and C. N. Yang. Statistical theory of equations of state and phase transitions. ii. lattice gas and ising model. *Phys. Rev.*, 87:410–419, Aug 1952. doi: 10.1103/PhysRev.87.410. URL `https://link.aps.org/doi/10.1103/PhysRev.87.410`.

[35] M. Lessel, O. Bäumchen, M. Klos, H. Hähl, R. Fetzer, M. Paulus, R. Seemann, and K. Jacobs. Self-assembled silane monolayers: an efficient step-by-step recipe for high-quality, low energy surfaces. *Surface and Interface Analysis*, 47(5):557–564, 2015. doi: https://doi.org/10.1002/sia.5729.

[36] L. Li, S. Reich, and J. Robertson. Defect energies of graphite: Density-functional calculations. *Physical Review B*, 72(18):184109, 2005.

[37] M. Lienemann, J.-A. Gandier, J. J. Joensuu, A. Iwanaga, Y. Takatsuji, T. Haruyama, E. Master, M. Tenkanen, and M. B. Linder. Structure-function relationships in hydrophobins: probing the role of charged side chains. *Applied and environmental microbiology*, 79(18):5533–5538, 2013.

[38] M. Lienemann, M. S. Grunér, A. Paananen, M. Siika-Aho, and M. B. Linder. Charge-based engineering of hydrophobin hfbi: effect on interfacial assembly and interactions. *Biomacromolecules*, 16(4):1283–1292, 2015.

[39] M. B. Linder. Hydrophobins: proteins that self assemble at interfaces. *Current Opinion in Colloid & Interface Science*, 5(14):356–363, 2009.

[40] M. B. Linder, M. Qiao, F. Laumen, K. Selber, T. Hyytiä, T. Nakari-Setälä, and M. E. Penttilä. Efficient purification of recombinant proteins using hydrophobins as tags in surfactant-based two-phase systems. *Biochemistry*, 43(37):11873–11882, 2004.

[41] M. B. Linder, G. R. Szilvay, T. Nakari-Setälä, and M. E. Penttilä. Hydrophobins: the protein-amphiphiles of filamentous fungi. *FEMS microbiology reviews*, 29(5): 877–896, 2005.

[42] Magakar. Hydrophobin film structure for hfbi and hfbii and mechanism for accelerated film formation. *PLoS Computational Biology*, 10(7):1–13, 2014. doi: 10.1371/journal.pcbi.1003745. URL `https://doi.org/10.1371/journal.pcbi.1003745`.

[43] K. Meister, A. Paananen, B. Speet, M. Lienemann, and H. Bakker. Molecular structure of hydrophobins studied with site-directed mutagenesis and vibrational sum-frequency generation spectroscopy. *The Journal of Physical Chemistry B*, 121(40):9398–9402, 2017.

[44] N. D. Mermin and H. Wagner. Absence of ferromagnetism or antiferromagnetism in one- or two-dimensional isotropic heisenberg models. *Phys. Rev. Lett.*, 17:1133–1136, Nov 1966. doi: 10.1103/PhysRevLett.17.1133. URL `https://link.aps.org/doi/10.1103/PhysRevLett.17.1133`.

[45] J. C. Meyer, C. Kisielowski, R. Erni, M. D. Rossell, M. Crommie, and A. Zettl. Direct imaging of lattice atoms and topological defects in graphene membranes. *Nano letters*, 8(11):3582–3586, 2008.

[46] D. L. Nelson, A. L. Lehninger, and M. M. Cox. *Lehninger principles of biochemistry*. Macmillan, 2008.

[47] D. R. Nelson and B. I. Halperin. Dislocation-mediated melting in two dimensions. *Phys. Rev. B*, 19:2457–2484, Mar 1979. doi: 10.1103/PhysRevB.19.2457. URL `https://link.aps.org/doi/10.1103/PhysRevB.19.2457`.

[48] M. Nitta, Y. Shida, H. Okada, M. Osumi, and W. Ogasawara. Hyphal surface architecture and cell morphology of trichoderma reesei. *Journal of electron microscopy*, 61(3):187–192, 2012.

[49] W.-K. Qi, Z. Wang, Y. Han, and Y. Chen. Melting in two-dimensional yukawa systems: A brownian dynamics simulation. *The Journal of Chemical Physics*, 133(23):234508, 2010. doi: 10.1063/1.3506875. URL `https://doi.org/10.1063/1.3506875`.

[50] L. Reichl. *A Modern Course in Statistical Physics*. E. Arnold, 1980. ISBN 9780713127775. URL `https://books.google.de/books?id=ue1ScAAACAAJ`.

[51] T. Sarlin, T. Kivioja, N. Kalkkinen, M. B. Linder, and T. Nakari-Setälä. Identification and characterization of gushing-active hydrophobins from fusarium graminearum and related species. *Journal of basic microbiology*, 52(2):184–194, 2012.

[52] Y. Shida, A. Morikawa, R. Tamochi, N. Nango, H. Okada, M. Osumi, and W. Ogasawara. Ultrastructure of the cellulolytic fungus trichoderma reesei. *PLANT MORPHOLOGY*, 27(1):15–20, 2015.

[53] R. D. Stanimirova, T. D. Gurkov, P. A. Kralchevsky, K. T. Balashev, S. D. Stoyanov, and E. G. Pelan. Surface pressure and elasticity of hydrophobin hfbii layers on the air–water interface: rheology versus structure detected by afm imaging. *Langmuir*, 29(20):6053–6067, 2013.

[54] L. STRYER, J. BERG, and J. TYMOCZKO. Bioquimica/jeremy m. berg. john l. tymoczko. lubert stryer. Technical report.

[55] T. Susi, M. Kaukonen, P. Havu, M. P. Ljungberg, P. Ayala, and E. I. Kauppinen. Core level binding energies of functionalized and defective graphene. *Beilstein journal of nanotechnology*, 5(1):121–132, 2014.

[56] G. R. Szilvay, T. Nakari-Setälä, and M. B. Linder. Behavior of trichoderma reesei hydrophobins in solution: interactions, dynamics, and multimer formation. *Biochemistry*, 45(28):8590–8598, 2006.

[57] G. R. Szilvay, A. Paananen, K. Laurikainen, E. Vuorimaa, H. Lemmetyinen, J. Peltonen, and M. B. Linder. Self-assembled hydrophobin protein films at the air–water interface: structural analysis and molecular engineering. *Biochemistry*, 46(9):2345–2354, 2007. doi: 10.1021/bi602358h. URL `https://doi.org/10.1021/bi602358h`. PMID: 17297923.

[58] M. Torkkeli, R. Serimaa, O. Ikkala, and M. Linder. Aggregation and self-assembly of hydrophobins from trichoderma reesei: low-resolution structural models. *Biophysical journal*, 83(4):2240–2247, 2002.

[59] I. M. Tucker, J. T. Petkov, J. Penfold, R. K. Thomas, A. R. Cox, and N. Hedges. Adsorption of hydrophobin–protein mixtures at the air–water interface: the impact of ph and electrolyte. *Langmuir*, 31(36):10008–10016, 2015.

[60] N. G. Van Kampen. *Stochastic processes in physics and chemistry*, volume 1. Elsevier, 1992.

[61] S. Varjonen, P. Laaksonen, A. Paananen, H. Valo, H. Hähl, T. Laaksonen, and M. B. Linder. Self-assembly of cellulose nanofibrils by genetically engineered fusion proteins. *Soft Matter*, 7(6):2402–2411, 2011.

[62] J. Wessels. Developmental regulation of fungal cell wall formation. *Annual review of phytopathology*, 32(1):413–437, 1994.

[63] T. A. Witten Jr and L. M. Sander. Diffusion-limited aggregation, a kinetic critical phenomenon. *Physical review letters*, 47(19):1400, 1981.

[64] F. Y. Wu. The potts model. *Rev. Mod. Phys.*, 54:235–268, Jan 1982. doi: 10.1103/RevModPhys.54.235. URL https://link.aps.org/doi/10.1103/RevModPhys.54.235.

[65] R. Yamasaki, Y. Takatsuji, H. Asakawa, T. Fukuma, and T. Haruyama. Flattened-top domical water drops formed through self-organization of hydrophobin membranes: A structural and mechanistic study using atomic force microscopy. *ACS Nano*, 10(1):81–87, 2016. doi: 10.1021/acsnano.5b04049. URL https://doi.org/10.1021/acsnano.5b04049. PMID: 26595357.

[66] C. N. Yang and T. D. Lee. Statistical theory of equations of state and phase transitions. i. theory of condensation. *Phys. Rev.*, 87:404–409, Aug 1952. doi: 10.1103/PhysRev.87.404. URL https://link.aps.org/doi/10.1103/PhysRev.87.404.

[67] F. Zampieri, H. A. Wösten, and K. Scholtmeijer. Creating surface properties using a palette of hydrophobins. *Materials*, 3(9):4607–4625, 2010.

[68] W. Zhang, W.-C. Lu, H.-X. Zhang, K. Ho, and C. Wang. Tight-binding calculation studies of vacancy and adatom defects in graphene. *Journal of Physics: Condensed Matter*, 28(11):115001, 2016.

[69] X. L. Zhang, J. Penfold, R. K. Thomas, I. M. Tucker, J. T. Petkov, J. Bent, A. Cox, and R. A. Campbell. Adsorption behavior of hydrophobin and hydrophobin/surfactant mixtures at the air–water interface. *Langmuir*, 27(18): 11316–11323, 2011.

Appendix A

KT-transition

To calculate the mean magnetization, here, I follow [6].

$$\frac{< S_x >}{S_0} =< \cos \vartheta(\vec{x}) >= \operatorname{Re} < e^{i\vartheta} >= \operatorname{Re} e^{\frac{(\vartheta(\vec{x}))^2}{2}} \equiv e^{-W},$$

where S_0 is the magnitude of spin, and $\vartheta(\vec{x})$ denotes the fluctuations of spin about its local minimum at position $\vec{x}$. e^{-2W} is called the Debye-Waller factor. The Debye-Waller factor determines how thermal fluctuations influenced the measurements from its maximum value at zero temperature [6].

$$W = \frac{(\vartheta(\vec{x}))^2}{2} = \frac{T}{2} \int \frac{d^d q}{(2\pi)^d} \frac{1}{\rho_s q^2}$$
$$= K_d \frac{T\Lambda^{d-2}}{2\rho_s},$$

where Λ is the wave number cutoff, and ρ_s is simply the rigidity in the continuum limit. We see that in two dimensions, $d \to 2$, the Debye-Waller factor tends to ∞. Thus, $< S_x >$ is zero and, a two-dimensional XY-system does not show a long-range order phase.

Appendix B

KT-transition

$$H - E_0 = \frac{J}{2} \int d^2r \left(\nabla \theta(r) \cdot \nabla \theta(r) \right)$$

$$\int d^2r \left(\nabla \theta(r) \cdot \nabla \theta(r) \right) = - \int d^2r \theta(r) \nabla^2 \theta(r)$$

$$= - \int d^2r \left(2\pi \int d^2r' \rho(r') g(r - r') \right) (2\pi \rho(r)) = -4\pi^2 A.$$

Substituting $\rho(r) = \sum_i q_i \delta(r - r_i)$ gives,

$$A = \int d^2r \sum_i \sum_j q_i q_j \left(\int d^2r' g(r - r') \delta(r' - r_j') \right) \delta(r - r_i)$$

$$= \int d^2r \sum_i \sum_j q_i q_j \delta(r - r_i) g(r - r_j')$$

$$= \sum_i \sum_j q_i q_j g(r_i - r_j) = \sum_{i \neq j} q_i q_j g(r_i - r_j) + \sum_i q_i^2 g(0),$$

where I used r_j instead of r_j', and since $g(0) = 0$, the last term vanishes. In the limit of $|r| \gg a$, $g(r) = \frac{1}{2\pi} \ln \left(\frac{r}{r_0} \right)$. Thus,

$$H - E_0 = -J\pi \sum_{i \neq j} q_i q_j \ln \left(\frac{|r_i - r_j|}{r_0} \right)$$

$$= -J\pi \sum_{i \neq j} q_i q_j \ln \left(\frac{|r_i - r_j|}{a} \right) - J\pi \sum_{i \neq j} q_i q_j \ln \left(\frac{a}{r_0} \right).$$

$$\sum_{i \neq j} q_i q_j = \left(\sum_i q_i \right) \left(\sum_j q_j \right) - \sum_i q_i^2 = - \sum_i q_i^2,$$

where in the last step, the electrically neutral condition, $\sum_i q_i = 0$, was used. Introducing a chemical potential $\mu = J\pi \ln \frac{a}{r_0}$ will complete the calculation.

$$H - E_0 = -J\pi \sum_{i \neq j} q_i q_j \ln \left(\frac{|r_i - r_j|}{a} \right) + \mu \sum_i q_i^2$$

Acknowledgments