MODELLING COLLAGEN FIBRIL STRUCTURE USING CONTINUUM THEORY

by

Samuel Cameron

Submitted in partial fulfillment of the requirements
for the degree of Master of Science

at

Dalhousie University
Halifax, Nova Scotia
June 2019

Table of Contents

List of Tables

List of Figures

Abstract

Collagen fibrils are microscopic, rope-like, biological structures that provide mechanical support to tissues within humans and other animals. It is still not fully understood how these fibrils self-assemble from their molecular constituents. In this thesis, we address questions of fibril formation and structure by applying continuum theories of soft matter physics to the collagen fibril. We introduce a theoretical model which allows us to predict structural and mechanical properties of collagen fibrils, and demonstrate that these properties are consistent with experimental observation. Our model also predicts coexistence between two different fibril phases, something which has been alluded to in the literature for many years.

List of Abbreviations and Symbols Used

BFGS Broyden-Fletcher-Goldfarb-Shanno.

CG Conjugate Gradient Descent.

ODE Ordinary Differential Equation.

PBLG poly-γ-benzyl-L-glutamate.

PFCF Phase-Field Collagen Fibril.

Acknowledgements

I would like to thank the people in my life for supporting me throughout my MSc: family, friends, co-supervisors, committee members, fellow group members, and discs made out of plastic.

First and foremost, I want to thank my co-supervisors, Andrew Rutenberg and Laurent Kreplak. Their guidance and encouragement has been a major part of my success in completing this thesis. Their support has given me the confidence to continue pursuing a career in physics research, while simultaneously providing me with the practical tools I need for this pursuit to be remotely plausible. I would also like to thank them for their patience in explaining an idea multiple times that I just can't seem to wrap my head around, in guiding me through the experimental literature and techniques as if they had just read the relevant paper that morning, and in reminding my math-loving brain that what matters most is that our work can describe reality (and not that we've only expanded to second order).

I would like to thank my parents for supporting my decision to continue studying physics, and actively showing interest in what it is I do every day (even when I am being particularly inarticulate). I would also like to thank them for reminding me how important it is to be able to connect my work to the rest of the world.

Finally, I would like to thank my wife, Molly, who happily[1] agreed to marry me half-way through the process of writing this thesis, and who has supported my work even through it renders me unable to multi-task and keep track of time for days on end. I would also like to thank Molly for her patience and always reminding me that we have frisbee practice in 10 minutes, and that maybe I should finish plotting that figure later.

---

[1] I think?

Chapter 1

Introduction

1.1 Motivation

Collagen fibrils, which are microscopic, biological rope-like assemblies found in the extra-cellular matrix of animals, have interested biologists for over 100 years[1]. They play an integral role in providing mechanical and structural support to tissues throughout the body of most organisms in the animal kingdom.

In contrast to the large experimental literature on collagen fibril structure, discussion on the theory of fibril structure is much more limited. Descriptive geometrical models have been put forth in attempt to ascribe observed fibril structure to the underlying molecular details. However, few of these models have the ability to make quantitative predictions, while at the same time providing concrete mechanisms for how such structure emerges. This gap provides a general motivation for the work I will present in this thesis: I will construct a collagen fibril model which can provide an equilibrium thermodynamic mechanism and quantitative predictions of fibril structure.

My personal motivation in studying collagen fibrils is two-fold. First and foremost, I find the task of understanding how complex, hierarchical systems arise from simple physical concepts to be both fascinating and satisfying. The collagen fibril is a wonderful demonstration of how micro scale structure can self-assemble from nano scale constituents. When I first became interested in physics, my impression was that everything outside of textbook problems was impossible to describe accurately with physical theory. Over the past few years, my research and modelling of collagen fibrils has shown me that with slightly more advanced theory, artful approximations, and a powerful computer, you can even model complex biological structure and capture (at least qualitatively, and in many cases quantitatively) the important physical

---

[1] To the author's knowledge, the first documented discussion that connective tissue may be composed of smaller, fibrillar units, is from 1903 [1].

phenomena you are interested in. My second motivation is much less deep, in that collagen fibrils are an important part of our daily lives. They are the main component of cornea, tendon, skin, bone, and other tissues found throughout the body. Changes in fibril structure and mechanical properties are therefore linked to certain diseases and the associated health problems that come with them. An extensive and thorough review on fibril structure, mechanical properties, and their relation to disease can be found in ref. [2].

1.2 Outline

Chapter 2 of this thesis will provide an extensive background to the collagen fibril literature. It will outline the current understanding of fibril structure and provide context for the remaining chapters. In this chapter, I will introduce the basic building block of collagen fibrils, the tropocollagen molecule, and discuss the hierarchy of collagen structures found in humans and other animals. I provide a self-contained overview of different features of fibril structure observed in experiment, including the fibril radius, molecular orientation, molecular packing, and mechanical properties. I will also briefly outline several theoretical models which have been applied to or built specifically for predicting fibril structure and mechanical properties.

Chapters 3 and 4 have been copied verbatim from published and in-preparation journal articles, respectively, and so are themselves more or less brief but self-contained in their discussions and literature reviews. I have tried to keep the repetition between these two chapters and Chapter 2 to a minimum where possible. In addition, I have added extra sections to the beginnings and ends of these chapters to connect them to the overarching research goal of constructing a predictive model of collagen fibril structure.

In Chapter 3, we present a simple, liquid crystal model of collagen fibril structure. This model considers only the molecular orientation within the fibril, and so ignores the periodic D-band structure along the fibril axis. The mathematical framework of this model was presented in previous work[3], and so we here expand on it by introducing a more efficient numerical scheme to solve the underlying equations. We also reduce the number of tune-able parameters through simple dimensional analysis. With these new additions to the model, we explore parameter space and reconcile our model with limiting cases of experimental structure. This main content of this chapter was published last year [4]. In the preface I outline my contributions as well as differences between this chapter and the published work.

In Chapter 4, we derive a new model of collagen fibril structure. We combine the liquid crystal model of Chapter 3 with techniques from the phase field crystal literature [5], and the resulting model allows calculation of both molecular order and the periodic axial D-band that is integral to collagen fibril structure. We combine the existing numerical framework utilized in Chapter 3 with a global optimization scheme to determine the equilibrium structure of this new model. We compare the predictions of our model with experimental observations. A welcome enhancement of this model over the previous model is the ability to "pull" on the fibril by straining the D-band period. In the final piece of this chapter we compare the mechanical response of our model to available experiment results.

Chapter 2

Background

2.1 Collagen fibril structure

2.1.1 Collagen molecules and their hierarchy

Collagen fibrils are at an intermediate level in the collagen structure hierarchy, which spans scales from nanometer (single molecule) to micrometer (tissue) range. An example of this hierarchy for tendon tissue is demonstrated in Figure 2.1. Typically, different tissues have similar collagen building blocks on the nanometer scale (i.e. molecules and fibrils in Figure 2.1), and variations in structure arise only on larger ($\geq$ µm) scales[1]. For example, in contrast to the fascicles and fibres in the tendon hierarchy, collagen fibrils in cornea tissue form 2d sheets which stack on top of each other (see e.g. [6]).

The smallest collagen member of tissue in all cases, the collagen molecule, consists of three left-handed helical poly-peptide chains (commonly referred to as α chains[9]). These three α chains are kept together via hydrogen bonding, and wind around each other to form a right-handed triple helix [10, 11]. It is this final, right-handed triple helix structure that is referred to as the collagen molecule. There are at least 28 different types[2] of collagen molecules[14], with differences between types arising from differences in amino acid sequence, the presence of non-helical (globular) domains at the ends of the molecule, and small molecular interruptions at points within the triple-helix structure[15]. Certain collagen types are more abundant in animals (with type I being considered the most abundant), and not all of the 28 types of collagen molecules aggregate together to form fibrils and the subsequent hierarchical structure shown in Figure 2.1. Non-fibrillar structures include collagen networks (where molecules are

---

[1] When looking more carefully, there are actually differences on the nanometer scale as well (e.g. collagen molecule types, which we discuss below), but for our purposes and in true physics fashion, we end up coarse-graining these details away in our models

[2] Collagen types are labelled using roman numerals, and these roman numerals are chronological with the date of each types discovery [12, 13].

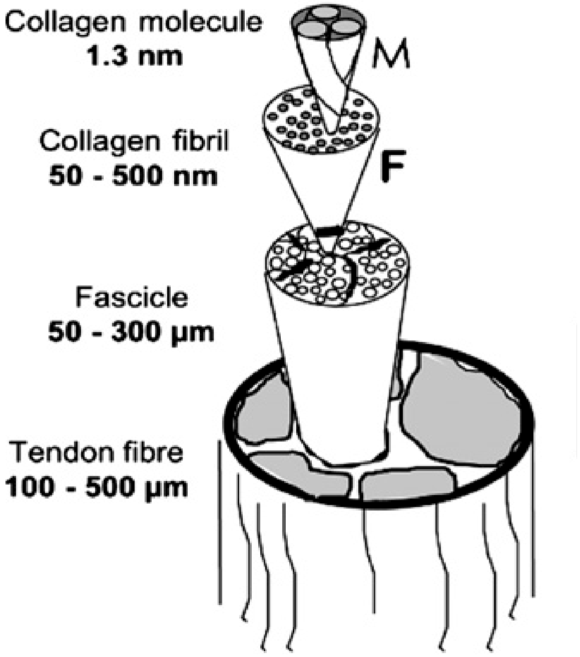

Figure 2.1: The hierarchical structure of collagen in tendon fibrils. The smallest collagen unit is the collagen molecule, which is $\sim 1.3\,\mathrm{nm}$ in diameter and $\sim 300\,\mathrm{nm}$ in length. The next level of the hierarchy is that of the collagen fibril, with diameter $\leq 500\,\mathrm{nm}$. Higher level ordering of collagen (e.g. fascicles and fibres) vary between different tissues. In this thesis, I focus only on the fibril level of the collagen hierarchy. Figure reprinted from [7] under CC BY-NC-ND [8].

held together by inter-molecular chemical cross-links) and trans-membrane collagens which adhere to different cell types and play a role in cell signalling [16].

The fibril forming collagen molecules (mainly types I, II, III, and V), also known as tropocollagen[9], arise from the cleaving of globular domains off of the ends of the three α chains, which allows the formation of the tropocollagen triple helical (chiral) structure. The triple helix is $\sim 300\,\mathrm{nm}$ long and $1.5\,\mathrm{nm}$ in diameter. This rod-like geometry forces the tropocollagen molecules to align nearly parallel to each other[3], analogous to lyotropic[4] liquid crystal systems [17]. The tropocollagen chirality is thought to provide a mechanism for the cylindrical (*vs* space filling) shape of collagen fibrils [3], and we shall explore this concept of chirality-driven radial control further

---

[3] As I will demonstrate in this thesis, the specific orientation of tropocollagen molecules within collagen fibrils is an interesting and partially unsolved question.

[4] Lyotropic liquid crystals transition from a disordered, liquid phase to liquid crystalline phases through increased density. This is in contrast to thermotropic liquid crystals, in which the phase transition is controlled by temperature.

in this thesis.

As alluded to in the previous paragraph, collagen fibrils are composed of tightly packed tropocollagen molecules. These fibrils have a cylindrical shape, though they are flexible and so have mechanical properties similar to that of a rope [18]. The radius of a collagen fibril is typically in the range of 15 nm to 200 nm, though fibrils outside of this range have been reported (see section 2.1.5 below). Almost all fibrils in the human body are predominantly made from type I tropocollagen molecules with small amounts of types III and V present[5]. These fibrils are found in many tissues throughout the human body including tendon, skin, and ligaments. Only fibrils found in cartilage are found to be predominantly composed of type II tropocollagen. These differences in molecular composition have been shown to control fibril ultra-structure [19].

Although this thesis focuses on the structure and mechanical properties of single collagen fibrils, it is worth mentioning that length scales greater than a micron in the collagen hierarchy consist of different ways to bundle together or stack collagen fibrils. These super-fibrillar bundles and networks provide a structural backbone to many tissues including skin, tendon, ligament, and bone. The arrangement of fibrils can vary widely even within tissues that are adjacent within the body. A clear example of this is in the layers of the cornea, where fibrils may be either bundled parallel to each other with a quasi-hexagonal lattice, or arranged in a random network (the former is found in the stroma, the latter in Bowman's layer) [20]. In contrast to cornea and other soft tissues, fibrils in bone have small minerals interwoven within the gaps of tropocollagen spacing to increase rigidity [21, 22, 23]. Super-fibrillar (tissue) mechanical properties are important for understanding injuries such as tendon rupture and more severe impact injuries [24], whereas both super-fibrillar structure and the control of fibril mechanical properties and structure is important in understanding collagen related diseases [25].

---

[5] The differences between these three tropocollagen molecules amount to different types and combinations of their constituent polypeptide chains which wind to form the triple-helix [14]

2.1.2 Mechanical properties of single collagen fibrils

The importance of collagen fibrils in the human body is largely due to its role in the mechanical properties of tissues. A vast literature exists on the study of mechanical properties of collagen from molecular to macroscopic length scales, and these mechanical properties vary between the different levels of hierarchy (see ref. [26] and the references therein). Reasons for these differences include the straightening and relative sliding of fibrils within fibril bundles and tissues [6]. However, since the focus of this thesis is on the properties of collagen fibrils, I will focus only on the mechanical properties at the fibril level in this discussion.

There are several procedures which allow the measurement of different single fibril mechanical properties, but for simplicity here I focus on the mechanics of small axial extension of the fibril, as partially characterized by a Young's modulus which linearly relates the axial strain to the fibril stress. Experimental stress-strain curves measure Young's modulus of the fibril at small strains, and in some cases a second elastic modulus is reported at higher strain [26] due to nonlinearities in the stress strain curve. The value of Young's modulus, Y, is dependent on the fibril environment. Fibrils immersed in water (wet fibrils) have $Y = 0.2 - 0.8\mathrm{GPa}$; "dry" (partially dehydrated, for ease of experimental study) fibrils have Y that are an order of magnitude larger than wet fibrils [27]. The models of fibril structure and mechanical properties that I present in this thesis are based on a liquid crystal framework, and so we expect our predictions to align well with wet fibrils.

One important consideration to take into account is that all fibrils measured from *in vivo*[6] samples contain intra-fibrillar chemical cross-links[7] [28] (typically the older the tissue, the more cross-links present [29]) which are expected to affect mechanical properties [30]. The presence of these cross-links is not considered in our model, making it difficult to quantitatively compare the mechanical properties we predict with those observed in studies of (cross-linked) *in vivo* fibrils. In the future, it would be interesting to apply the theoretical framework of liquid crystal elastomers (i.e.

---

[6] *In vivo* collagen fibrils are those which have been extracted from an organism and studied without modification. *In vitro* collagen fibrils are those which are self-assembled in e.g. test tubes from tropocollagen.

[7] Cross-links, which are chemical bonds facilitated through either enzymatic or non-enzymatic reactions [28], can be formed between molecules within the fibril, and between fibrils. In this thesis, we are always considering cross-links within fibrils (intra-fibrillar).

cross-linked liquid crystalline materials) [31] to enhance our model and so allow for more quantitative comparison with *ex vivo* studies, though the non-ideal nature of collagen cross-links may complicate this [32]. Alternatively, the ideal experiment in which we could compare with the mechanical predictions of our model would be measurements of single *in vitro* (i.e. non-cross-linked), hydrated fibrils. No such studies exist that I am aware of, with the closest study being an *in vitro* study of dry fibrils [33] which reported a Young's modulus of 32 MPa (though see ref. [27]).

2.1.3 Self-assembly of collagen fibrils: Fibrillogenesis

Models of collagen fibril structure that I will present in this thesis will involve several parameters, all of which are controlled by the environment in which the fibril is immersed in. In this section I provide an overview of the different environments where fibrillogenesis (i.e. assembly of collagen fibrils through aggregation of tropocollagen molecules) occurs. The two broad differences in fibril environment are within the living organisms *in vivo*, or through extraction and solution processing of tropocollagen *in vitro*. I outline the current understanding of both processes briefly, beginning with *in vitro* fibril assembly. For a thorough overview of fibrillogenesis both *in vivo* and *in vitro*, see the recent review by Holmes et al. [34].

I will not go into great detail on the many experimental procedures of producing *in vitro* fibrils; instead I will just overview the basic processing required to reconstruct fibrils[35, 36]. *In vitro* fibrillogenesis begins by obtaining fibrils/fibres of collagen taken from an animal sample, removing non-collageneous proteins by treating the fibrils with trypsin, and solubilizing the fibrils with an acidic solution. After this procedure has taken place, the resulting solution contains only tropocollagen molecules and solvent (this is often checked by imaging samples of the tropocollagen solution with electron microscopy) which has a pH $\lesssim 3.8$. To induce precipitation of fibrils (i.e. to "reconstitute" the fibrils), slides coated with the tropocollagen solution can either be dipped in NaCl solution or buffered solution (with pH $\gtrsim 4.8$). The fibrils that form from these methods generally depend on all experimental parameters including ionicity, pH, and temperature of solution during the precipitation stage. A thorough study of solution condition effects on *in vitro* fibrillogenesis can be found in a series of papers by Wood[36, 37, 38].

Possible mechanisms of *in vivo* fibril assembly are much more complicated than those *in vitro*. For quite some time (roughly between 1930-1980), it was debated whether fibrils were constructed within specialized cells known as fibroblasts [39], or if aggregation of tropocollagen molecules into fibrils occurred in the extra-cellular matrix analogous to the *in vitro* processes described above [40, 41]. The difficulty in determining which mechanism is dominant *in vivo* stems from the two dimensional nature of electron microscopy imaging, which makes it difficult to determine whether fibrils are within or above the cell [41]. The debate was more or less settled when careful sectioning and imaging of tendon fibroblasts revealed that fibrils are initially formed within cavities on the cell surface[42]. It has since been confirmed that these cavities (now known as "fibropositors" [43]) are also present in cornea tissue [44]. Once fibrils of approximately 15 nm radius[43] have been formed within these fibropositors, they are secreted into the extra-cellular matrix. Once they are in the extra-cellular matrix, the fibrils may continue growing in radius and length through accretion of extracellular tropocollagen molecules as in e.g. tendon fibrils, or remain nearly the same size as in cornea fibrils (see section 2.1.5 for discussion of fibril radii seen *in vivo* and *in vitro*). End-to-end fusion of fibrils can also occur in early stage fibril growth [45].

2.1.4 D-band and the packing of tropocollagen molecules in fibrils

A defining feature of collagen fibril ultra-structure is the periodic banding pattern that occurs along the long axis of the fibril, known as the fibril "D-band"[8]. This D-band, shown in Figure 2.2, is a striking feature of collagen fibrils when viewed using high resolution (i.e. sub-micron) imaging or scattering technique. It is so named due to its periodically alternating bands of light and dark shading along the long axis of the fibril. Measurements of the D-band period, $D \sim 67\,\text{nm}$, were not made until the advent of the electron microscope [47]. Since this time, the D-band has been characterized using multiple techniques including x-ray diffraction, electron microscopy, and atomic force microscopy. Variations in the D-band period between tissues within humans and other vertebrates are small, with tendon fibrils typically

---

[8] The name "D-band" (also known as the d-period) was first put forth in 1974 [46]. The "d" does not appear to refer to anything deep, but instead is due to the fact that the letter D was used as the variable referring to the length of the periodic banding, i.e. $D = 67\,\text{nm}$.

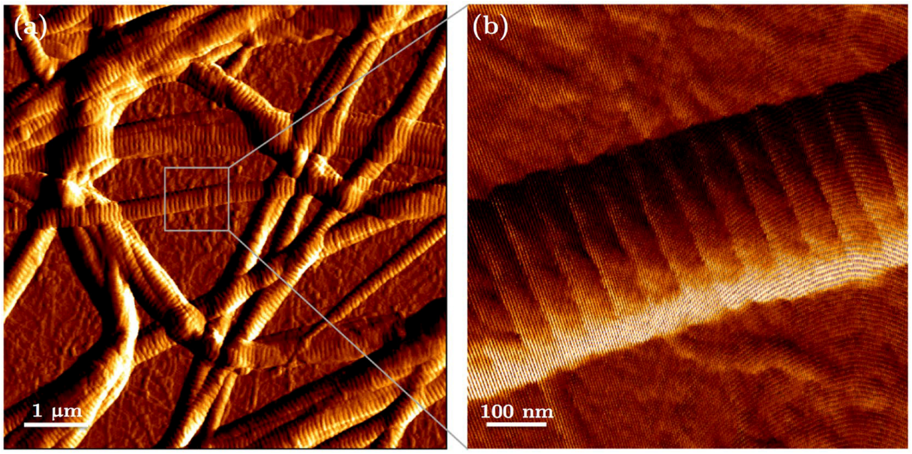

Figure 2.2: (a) Atomic force microscopy image of collagen fibril network synthesized *in vitro* from type I collagen. (b) Higher resolution image of a single *in vitro* fibril showcasing the characteristic D-band striations with period of 67 nm. Figure reprinted from [19] under license CC BY [51].

having $D = 67\,\text{nm}$ and other fibrils (most notably from cornea and skin) having slightly smaller values of $\sim 64\,\text{nm}$ [48, 49, 50]. The D-band of *in vitro* fibrils is similar to that seen *in vivo* with exceptions arising only due to highly altered molecular composition within fibrils [19]. The currently accepted molecular model for the D-band was proposed by Hodge and Petruska in 1964 [11], and a sketch of this model is shown in Figure 2.3. I will present here the conclusions of this model to illustrate the current understanding of how D-band emerges from the packing of tropocollagen molecules.

The key insight of Hodge and Petruska in constructing their model was to recognize that the tropocollagen molecule (length $\sim 300\,\text{nm}$) is approximately $4.4D$. This allowed them to deduce that the centre of mass separation between molecules in a straight line (analogous to lining up pencils by connecting the lead tip of one pencil to the eraser of the other pencil) must be at least $5D$. If there is a separation of $5D$, then the distance between two molecules in a straight line must be $0.6D$. They then chose a unit cell containing five tropocollagen molecules lined up parallel and stacked on top of each other. Molecules perpendicularly adjacent to each other with respect to the long axis of each molecule are longitudinally offset by a factor of D, forming a staircase structure where each stair is D in length. Two periodically repeating units emerge from the connecting a large number of unit cells along the long axis of the

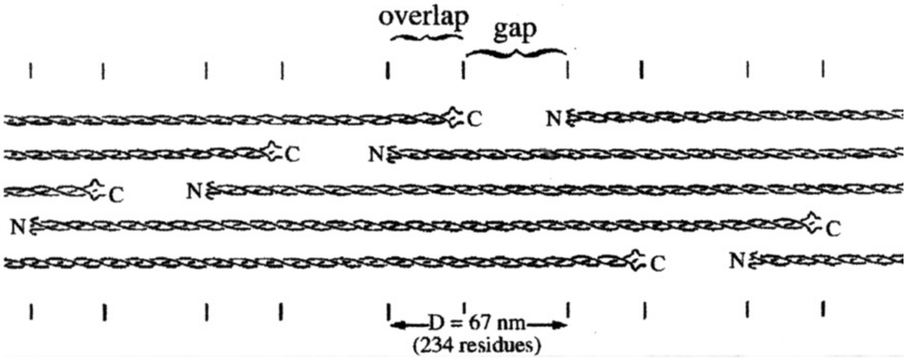

Figure 2.3: Schematic diagram of the Hodge-Petruska [11] model of the collagen fibril. The fibril cylindrical axis is in the same direction as the schematic's long axis (horizontal across the page). Tropocollagen molecules (long triple helices) stagger with respect to each other to form alternating gap and overlap regions, with gaps being slightly wider than overlaps. In gap regions, only a four out of five molecules are present in a cross-section perpendicular to the long axis of the fibril. In the overlap region, all five molecules are present. The period of this alternating gap and overlap is 67 nm (or 234 amino-acid *residues*). C and N labels at the ends of each tropocollagen molecule indicate very small, non-helical end regions ("telopeptides") present in the tropocollagen molecule. Figure reprinted from [52] (Copyright 2008) with permission from Elsevier.

five molecules indefinitely. In the first unit of length $0.4D$, known as the "overlap region", all five tropocollagen molecules are present in a cross-section perpendicular to the molecular orientation. The other unit of length $0.6D$, known as the "gap" region, has only four out of a possible five molecules present in the perpendicular cross-section. Thus, the repeat of this gap-overlap structure provides an explanation for the alternating striations viewed in Figure 2.2b.

The weakness of the Hodge-Petruska model is that it does not address how the five molecules within the unit cell are arranged three dimensionally, and how the repeating unit cells are pieced together to form higher order fibrillar structure. It is possible that they are exactly as depicted in Figure 2.3, i.e. two dimensional sheets with weak curvature. These sheets could then wrap around the fibril axis into a quasi-hexagonal packing, with a defect at the fibril centre due to the 5 molecule unit cell. Another proposition dating back to 1968 is that each unit cell is actually structurally independent, forming a collagen "micro-fibril" [53]. These five molecule micro-fibrils would then aggregate to form larger fibrils. Electron microscopy of tendon collagen fibrils has indicated that a cross-sectional length scale of approximately 4 nm is present in the mature fibril along with the 1.5 nm length scale arising from the diameter of individual tropocollagen molecules [54]. This 4 nm length scale has lead researchers to conclude that the micro-fibril picture of collagen fibril structure is likely correct. A more recent study of cornea collagen fibrils imaged using automated electron tomography has also indicated a 4 nm length scale is present [55], demonstrating that micro-fibrils are present in fibrils extracted from different anatomical locations. To the author's knowledge, no quantitative data on the presence of micro-fibrils *in vitro* exists; studies have indicated that small micro-fibrils of diameter $3 - 15$nm are later formed into mature fibrils [56, 57] but it was not investigated whether these smaller length scales were still present in mature fibrils (in addition, 15 nm appears to be too large to be reasonably called a single micro-fibril). A crystallographic description of the tendon micro-fibril unit cell using x-ray fibre diffraction was first obtained in 2006 [58], providing evidence that micro-fibrils of radius 2 nm are indeed present in within tendon fibrils, and also that these micro-fibrils have a (radially) quasi-hexagonal arrangement within the fibril as suggested in earlier work [54]. However, further details on the packing of micro-fibrils (e.g. whether they are parallel or tilted with respect to the long fibril axis) was neglected.

In this thesis, no importance is placed on the fibrils being composed of micro-fibrils or individual tropocollagen molecules, as we are using continuum methods which coarse grain the fibril at length scales larger than both molecules and micro-fibrils. We only require that the building blocks of the fibril are chiral, which is true for both molecules and micro-fibrils [58]. Any other molecular details will be encoded into our coarse-grained parameters. For simplicity in the remainder of this thesis, I will discuss fibril structure as if micro-fibrils do not exist (i.e. tropocollagen immediately aggregates to form mature fibrils) as it will not affect subsequent discussion on fibril structure unless otherwise noted.

2.1.5 Fibril radius: Experimental results and *in vitro* control

In this section, I review the literature on experimental measures of collagen fibril radius, R. This review will be useful when comparing our model predictions of R to those measured in experiment. This section also serves to illustrate the wide range of fibril radii that have been measured experimentally. I will begin by discussing measurements of fibrils taken *in vivo*, and then discuss some *in vitro* measurements of R along with some experimental findings on how to control the size of R.

In vivo collagen fibrils

Motivation to determine the radius of collagen fibrils began around 1910, when it was hypothesized that sub-micrometer radii of collagen fibrils would be necessary to explain the transparency of cornea tissue[9]. However, it was not until the advent of the electron microscope that the first quantitative measurements of fibril radius were reported. These initial measurements were all performed on corneal stroma tissue stained with osmium tetroxide (to improve contrast), and indicate that stroma fibrils have radii between the range of 15 nm to 25 nm depending on which animal the tissue was extracted from[10]. More recent studies using electron microscopy[20, 61] and x-ray scattering [62, 55] support this range of fibril radii. One important finding of that is shared amongst these measurements is that qualitatively, the distribution of radii within the corneal stroma is sharply peaked around its average value[11].

From 1940 to the mid 1970s, researchers began studying fibril radii in tissues other than corneal stroma. In 1977, Parry et al. produced an extensive review of the topic, summarizing both their own and other prominent studies of collagen fibril diameter distributions for over 20 different connective tissues [65]. This review illustrated that different tissues may have vastly different distributions of collagen fibril radii, which they broadly classified as either unimodal or bimodal. More recent studies confirm the bimodality of collagen fibrils radial distributions in tendon and ligament tissue [66, 67, 68]. In contrast to cornea, other connective tissues encompass much wider ranges of fibril radii, from 15 nm to 250 nm [69]. In particular, tendon fibrils appear to have the widest distribution of fibril radii.

In this thesis, I will address corneal fibrils and tendon fibrils as representative extremes of collagen fibril ultra-structure Tendon fibrils have a wide range of fibril radii, $R \in [15\,\text{nm}, 250\,\text{nm}]$, whereas cornea fibrils have a very small range of fibril radii, $R \sim 15 - 25\text{nm}$. Other fibrils found *in vivo* generally have ranges of R which are in between cornea and tendon. Thus, cornea and tendon provide useful upper

---

[9] see reference to H. Virchow in [59]

[10] for a brief review of these early results, see ref. [60]

[11] for quantitative proof see e.g. [63, 64]

and lower bounds on R, which we may compare to our theoretical models of collagen structure.

In vitro **collagen fibrils and methods of size control**

In vitro studies on collagen fibrils allows for systematic investigation of how fibril ultra-structure changes with solution conditions. For example, understanding how fibril radius changes with experimental parameters like pH and temperature will be useful later when we look to reconcile our model predictions with experimental results. The first quantitative measures of in vitro fibril radii distributions were performed by Wood and Keech[36], demonstrating the large variations in fibril radius possible through tuning different experimental parameters, and revealed that ionicity of solution strongly controls both the R distribution's average and width, with R increasing with ionicity (see also ref. [70] for a more recent study on pH and ionicity). Aside from external solution conditions, it has also been shown that ratios of tropocollagen types within the fibril serve to control R. In particular, inducing fibrillogenesis from mixtures of types I and III or types I and V tropocollagen tends to decrease the average fibril radius [19, 71]. These *in vitro* studies of fibril radius will provide insight into the connection between our model parameters and experimental conditions.

2.1.6 Twist

D-band and fibril radius are two "obvious" features of fibril structure, in that they are very easily visible via standard imaging/scattering experiments. A third, less obvious component to the fibril structure is that of the molecular (or micro-fibrillar) orientation within fibrils. In section 2.1.4 above, I discussed the Hodge-Petruska model which assumes that molecular orientation within the fibril is parallel to the long (z) axis of the fibril. However, less than a decade after this model was proposed, it was discovered that, at least on the surface of fibrils, tropocollagen orientation was at a non-zero angle with respect to the fibril z axis (see ref. [72] and the references therein). In this thesis, I refer to this angle at the fibril surface as the "surface twist", and denote it by $\psi(R)$ (for reasons to become obvious in the next paragraph). The discovery of this non-parallel molecular orientation on the fibril surface, $\psi(R) \neq 0$, has since lead to increased interest in the functional role of molecular orientation

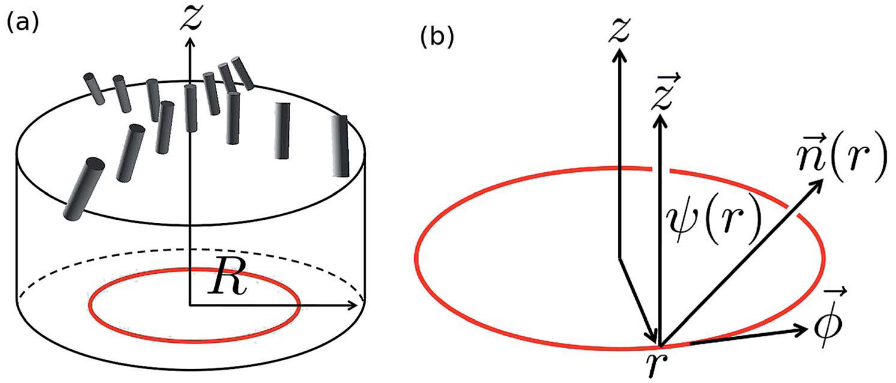

Figure 2.4: Molecular orientation of tropocollagen allowed in our double-twist model of collagen fibril structure. (a) Schematic representation of the molecular orientation of tropocollagen molecules (small cylinders) within a collagen fibril of radius R. (b) Definition of the constrained director field, $\boldsymbol{n}(r)$, corresponding to the local orientation of tropocollagen molecules. $\boldsymbol{n}$ is constrained to be azimuthally symmetric, and so its orientation with respect to the cylindrical (z) axis of the fibril is defined by the local twist angle, $\psi(r)$. Figure reprinted from [3] with permission from Royal Society of Chemistry.

within fibrils [73].

To simplify visualization of the ensuing discussion, I will define a unit vector field, $\boldsymbol{n}$, which represents the local (average) orientation of molecules within a fibril of radius R. In the theory of liquid crystals, this vector field is known as the director field. I will further constrain $\boldsymbol{n}$ to have a "double-twist" form [3], which is parameterized by a single function $\psi(r)$ (see Figure 2.4)[12]. $\psi(r)$ is referred to as the twist angle at a given distance r from the centre of the fibril, with $0 \leq r \leq R$. Note that $\psi(R)$ is the molecular orientation on the fibril surface, i.e. the surface twist. The double-twist form of this director field and the resulting angle $\psi(r)$ is an integral part of this thesis, and will be revisited many times.

Initially, as mentioned above, interest in the molecular orientation was restricted to that of the surface twist $\psi(R)$ due to difficulties in probing the nano scale structure within the fibril. Table 2.1 summarizes measurements of $\psi(R)$ determined in different tissues (see mainly [74] and the references therein), along with typical values of R. We also note whether the radial distribution of the fibrils is unimodal or bimodal. Notably, as with fibril radius, the two extremes of surface twist are those of tendon,

---

[12] This double-twist form assumes azimuthal symmetry in the fibril.

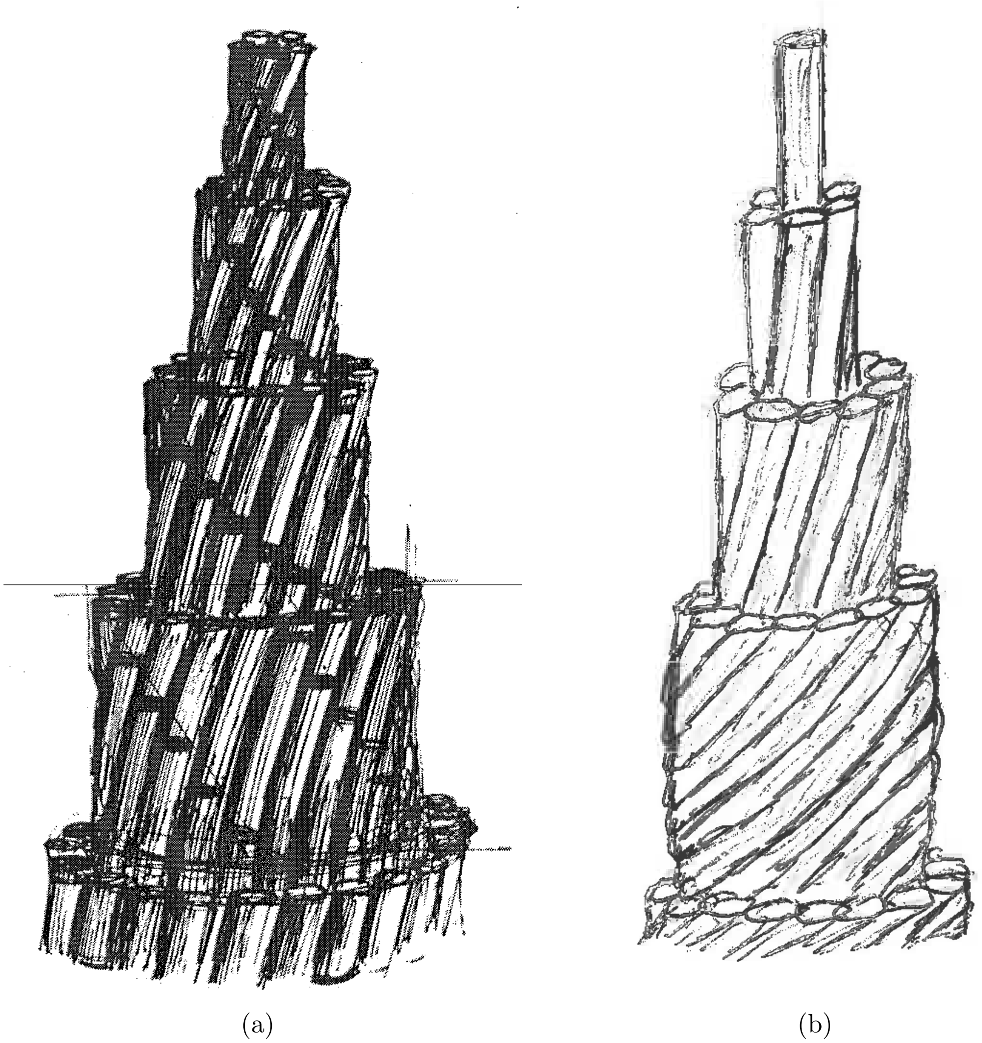

Figure 2.5: Two descriptive models of collagen fibril structure. (a) Constant angle model, $\psi(r) = \psi_0$, proposed by Galloway [77]. (b) Constant angle gradient model, $\psi(r) = \psi'_0 r$, proposed by Folkhard et al. [78]. See text for definition of $\psi(r)$. (a) reprinted from [77] with permission from Springer Nature.

with $\psi(R) \sim 5°$, and cornea, with $\psi(R) \sim 17°$.

Table 2.1: Radius and surface twist for several different *in vivo* fibrils.

Fibril type/location	**Radius** R	**Surface twist** $\psi(R)$
Tendon [74, 48]	$10 - 200$nm (unimodal or bimodal)	5°
Skin [74, 48, 75]	30 nm (unimodal)	18°
Cornea [74, 55]	15 nm (unimodal)	$15 - 18°$
Cartilage [76]	10 nm and > 20 nm(bimodal)	3°

With the increasing amount of experimental data came an increased number of speculations on what $\psi(r)$ might be within the fibril. In Figure 2.5, I show two distinct models of what the inner fibril structure might be, along with their $\psi(r)$ curves (conveniently, both models obey the double-twist symmetry required to be parameterized by $\psi(r)$). The first model, proposed initially by Galloway [77] and shown in Figure 2.5a, asserts that the inner molecular orientation is such that $\psi(r) = \psi_0$ is a constant. In particular, Galloway states that tropocollagen molecules (not microfibrils) form concentric cylindrical shells in multiples of 5, such that the innermost cylindrical shell has only five molecules (i.e. is a Smith micro-fibril [53]), the second shell has 10 molecules, and this pattern continues on such that shell m has $5m$ molecules. Galloway claimed that this model is consistent with the x-ray diffraction data proposed by ref. [54] through semi-quantitative arguments (which were later refuted from more quantitative calculations [79]). A further shortcoming of this model is that there is no clear mechanism for the constant twist of molecules. Given the chiral nature of the tropocollagen, one would expect that molecules would twist with respect to each other (analogous to chiral nematic phases in liquid crystals). This is not the case for $\psi(r) = \psi_0$ as the radial neighbours of each molecule are all aligned at the same angle. Furthermore, it is not clear how such a model would limit the growth of collagen fibrils radially to a single value (as is observed in cornea collagen fibrils[63, 64]).

The second model (Figure 2.5b), proposed initially by Folkhard et al. to explain observed x-ray diffraction patterns of different fibrils[78] (and first visualized by Raspanti et al. [80]), is that of a linearly increasing angle $\psi(r) = \psi_0' r$. This model therefore proposes molecular orientation analogous to the linear double-twist orientation found within liquid crystal blue phases [81]. Unlike the model proposed by Galloway, this model has the advantage of naturally accounting for the chirality of the tropocollagen molecules. Furthermore, this linear twist provides a simple mechanism for limiting radial growth [82, 3] (more discussion on this in section 2.2.2 below). However, this model has its own downfalls. With molecules holding different angles at different distances from the centre of the fibril, r, it would appear that a constant D-band is unattainable, as the effective axially-projected length of molecules at each r would be different by a factor of $\cos\psi(r)$. The only possible reconciliation of this model with the constant D-band observed in experiment is if the individual tropocollagen molecules undergo axial strain such that molecules near the fibril centre (surface) are under compression (tension) [48].

The two models discussed above both have their limitations in terms of reconciling experimental observation with the underlying molecular of collagen fibrils. This leaves two options, either devise an experiment to better elucidate the inner molecular orientation of molecules within the fibril, or build a theory which predicts structure from more fundamental physical concepts. The former was (valiantly) attempted by Holmes et al. in 2001 [55]. In this study, automated electron tomography was used to reconstruct images of sawn cornea fibrils, as shown in Figure 2.6. The resulting images indicated that molecules were twisting at angles up to $\sim 15^\circ$ within the cross-section, as is shown in the "slice" and "inverse transform" images in Figure 2.6. The authors interpreted these images as being consistent with Galloway's constant twist $\psi(r) = 15^\circ$ model. However, from my perspective, the resulting slices, while informative, provide inconclusive information at best. It appears that a range of twists are observed in all three slices, and most dominantly in the middle slice. I would go as far as saying that $\psi(0) = 0$ in the middle slice. In order to truly resolve $\psi(r)$, a much higher imaging resolution would be required with more than three slices.

The above discussion highlights an intrinsic difficulty of determining nano scale structure within complex, hierarchical materials. Typically, as is the case for the experiment performed by Holmes et al. [55], the data measured must be interpreted via fitting to a structural model, making any conclusions model dependent. If, on the other hand, a predictive, theoretical model of collagen fibril structure was available which matched reasonably well with experiment, the uncertainty in interpretation of experimental data would be greatly reduced. This motivates the need for a predictive model of collagen fibril structure derived from fundamental physical concepts. With this in mind, I will now discuss some candidate models of fibril structure, before moving on to the main results of my thesis.

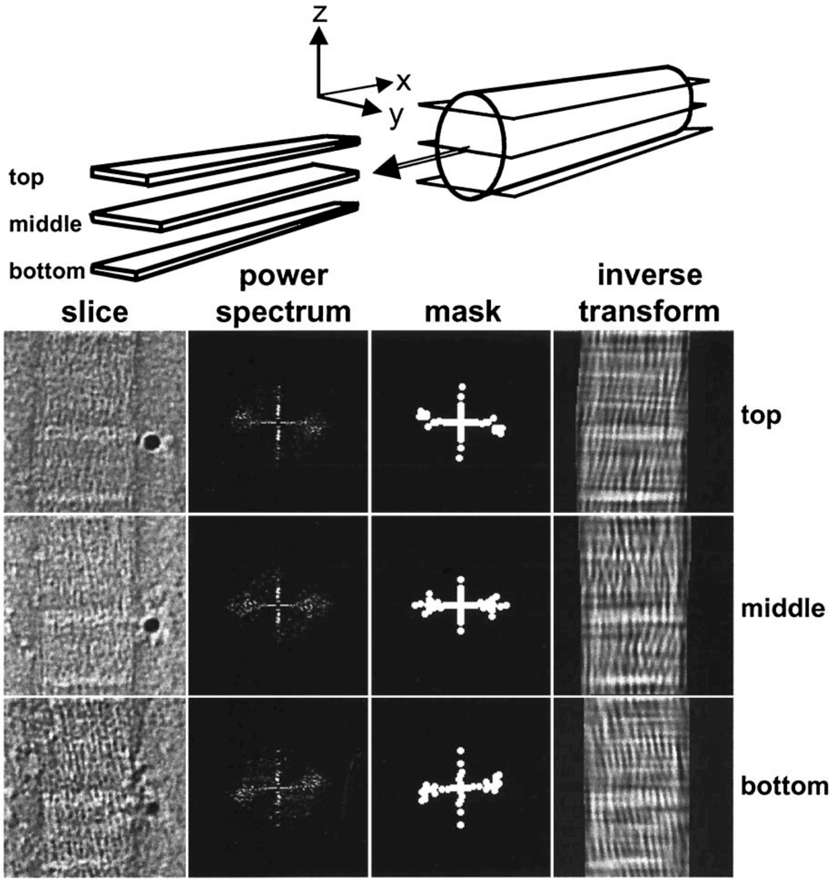

Figure 2.6: Three sawn slices of corneal collagen fibril structure as viewed by transmission electron tomography. The top schematic illustrates the locations of the three slices within the collagen fibril. The three rows underneath this schematic have the longitudinal (x-y plane) images of fibril structure for the top, middle, and bottom slices, respectively. The first column shows the raw 3D reconstruction of the microfibrillar orientation within the fibril. The second and third columns are the raw and masked power spectrums of this reconstruction, respectively. The fourth column is the resulting filtered image of the fibril structure. The mask (column three) applied to the inverse transform was input by eye, and so the resulting inverse transform of column four is biased. Reprinted from [55]. Copyright 2001 National Academy of Sciences.

2.2 Previous Quantitative Models of fibril structure

In this section, I present models of fibril structure from the literature which are quantitative, and provide energetic mechanisms which predict fibril structure. I begin with two models which are molecular in nature, but fairly simplistic in that they consider only specific packing structure of individual tropocollagen molecules.

First, I will briefly mention a model utilizing Leonard-Jones type energetic relaxation, where circular cross-sections containing tropocollagen molecules are allowed to relax slightly [79]. The second model is a more sophisticated molecular dynamics approach which explicitly coarse-grains the true amino acid sequence of individual tropocollagen molecules to determine model parameters. These two models are fairly constrained in what forms the structure may take due to the initial packing configuration, and so are not entirely satisfactory to meet the above requirements of predicting larger scale fibril structure.

The third model I will present is based on minimization of a free energy for hexagonally packed chiral filaments [82], and so applies generically to collagen fibrils. This model bears similarity to the main theory I will present in Chapters 3 and 4. However, it includes an additional assumption of hexagonal packing and so introduces elastic coefficients from the theory of solid elasticity [83], which we do not introduce in our model. I conclude this background chapter with a brief discussion of how these quantitative models fit in with the main models developed and presented in this thesis.

2.2.1 Lattice models with ad-hoc energy relaxations

The first (pseudo) energetic model of collagen fibril structure was presented by Hulmes et al. in 1995 [79]. In their model, they are most concerned with reproducing x-ray diffraction data from an earlier study [54]. They approach this problem via inputting several radial packing structures of collagen fibrils (consistent with the Hodge-Petruska D-band model discussed in section 2.1.4), and subsequently relaxing the lattice by minimizing a Leonard-Jones potential with parameters tuned to match experiment (i.e. energy well depth and equilibrium distance). They concluded that the most appropriate lattice which may reproduce experiment is that of a quasi-hexagonal packing. This model is limited in its predictive power in that a very simplistic energy relaxation scheme was used (i.e. Leonard-Jones), the different radial packing arrangements were not compared energetically, the fibril radius was held fixed at 50 nm, and no consideration of axial molecular orientation (i.e. $\psi(r)$) was considered as the molecules were modelled as achiral cylinders.

More recently, researchers have employed both coarse-grained [84, 85] and atomistic [86] molecular dynamics simulations of collagen microfibril structure in attempt to elucidate microfibril mechanical properties, and also to relax the microfibril into a minimum energy structure. The coarse-grained studies involved two major steps. The first step was to measure the mechanical response of an atomistic tropocollagen molecule (see ref. [84] for the definition of atomistic) under tensile strain, bending strain, and shear strain (the latter requiring two tropocollagen molecules). The second step was then to use the measured response of these three mechanical tests to construct spring constants for a bead and spring model of microfibrils. Application of this model was limited to a two dimensional, periodic unit cell containing 10 bead and spring tropocollagen molecules (each molecule with 200 beads), and the corresponding mechanical properties measured were within two orders of magnitude [87]. The author also claimed that this system displayed the same structure as the what is observed in experiment, but no quantitative (or qualitative) evidence was provided.

A more sophisticated model of the collagen microfibril was presented in 2011 [86]. The key piece of information utilized in this model was earlier x-ray diffraction studies of the microfibril [58], which provided a precise description of the microfibril unit cell. Tropocollagen molecules with full amino acid sequences were placed on the unit cell with periodic boundary conditions, and so the resulting structure was that of an infinitely long, infinitely wide microfibril (in contrast to true collagen fibrils with finite radius). This bulk microfibril was allowed to equilibrate for 8.5 ns to determine its minimum energy configuration. With this bulk microfibril model, the authors were able to extract a Young's modulus within the range of values observed in experiment [86]. This model also suggested a hierarchical deformation mechanism, where at small strain, the microfibrillar twist reduced until the molecules were parallel with each other, at which point the molecules themselves began to straighten. Furthermore, the inclusion of the amino acid sequence in this model provides a solid framework to investigate the effects of tropocollagen mutations which are related to collagen-based diseases [88].

However, there are still limitations in this model for predicting structural properties of full sized collagen fibrils, which is what I am interested in for this thesis. One issue with the atomistic model is that the microfibril structure was only relaxed over

8.5 ns, and so it seems unlikely that it was able to reach an equilibrium configuration. This short relaxation is due to the large computational cost associated with a fully atomistic treatment of tropocollagen, and is unlikely to be addressed with anything but faster computer processing speeds.

A second issue of this atomistic model is that a microfibril with periodic boundary conditions in all directions is unable to capture the structure of collagen fibrils observed in experiment for two reasons: a) One important observed quantity in fibrils is the fibril radius, which is infinite in this model and b) it has been shown experimentally that the orientation of microfibrils within the fibril is not a simple unit cell structure [55] – a repeated unit cell of microfibrils is incompatible with fibril surface twist and azimuthal symmetry. From these two observations, it is clear that the model by Gautieri et al. [86] is fundamentally unable to reproduced the structure of collagen fibrils.

The atomistic model described above was developed mainly as a tool to elucidate the mechanical properties of collagen fibrils, and so was less concerned with structure of the fibril. However, the authors themselves showed that the mechanical properties of the microfibril were connected deeply with its structural properties[13]. Therefore, it seems reasonable to assume that if a full sized fibril was pulled on, the response would begin with a fibril-sized deformation mechanism, followed by microfibril and molecular deformations. This further motivates the need for full sized fibril studies on the relationship between mechanical response and fibril structure.

Given that the above model already requires considerable computational time to determine structure and properties on the microfibril scale, it seems unlikely that expanding this atomistic model to a full collagen fibril (or even multiple microfibrils) will be possible with current computational resources. Therefore, to capture important structure of the fibril such as the fibril radius and the molecular orientation, it is likely that lower computational cost, coarse-grained approaches will be required.

---

[13] As discussed, this response began with low strain microfibril deformation, followed by high strain molecular deformation

2.2.2 Hexagonal packing of long, chiral filaments

An alternative approach to molecular models is to use highly coarse-grained continuum theories to predict structural and mechanical properties of fibrils. In continuum theory, the microscopic details of the constituent molecules (or microfibrils) are largely ignored. Instead, only the symmetry of the constituent components of the system are considered in constructing free energies. For collagen molecules, the most important molecular symmetries are their chiral nature and their long aspect ratio. Conveniently, generic models of long, chiral molecules forming into fibrous bundles have been recently developed in the literature [82, 3]. These models are relevant to collagen fibrils as they provide a theoretical framework which allows prediction of fibril radius and molecular orientation, and build off of the continuum theory of liquid crystals (in particular, the double-twist geometry of blue phase liquid crystals) [81, 17]. Here I focus on a model presented initially by Grason and Bruinsma [82, 89]. This model considers a system of hexagonally packed, chiral filaments, and constructs a free energy which respects these symmetries.

There are four components that must be considered in constructing a free energy for chiral filaments. The first is the free energy of deforming a two dimensional hexagonal packing structure which is taken directly from the theory of solid elasticity [83], though with a re-definition of the (2D) strain tensor to account for the rod-like filaments [90]. The second term is the Frank free energy of a chiral liquid crystal phase, which penalizes any distortions of molecular orientation (represented by a director field $\boldsymbol{n}(\boldsymbol{r})$ as discussed in section 2.1.6 above [17]). The third term is the axial deformation energy associated with stretching and shearing filaments from their (bundle or fibril) long axis, again taken from the theory of solid elasticity (this term uses the standard definition of the 3D strain tensor). Notably, by varying the elastic coefficients in this term one can model a filament bundle with either three dimensional hexagonal solid symmetry, or two dimensional columnar liquid crystal (the latter having no energetic penalty for shearing molecules along the bundle axis). The final term of the free energy is that of a symmetry breaking term that is only found in chiral, hexagonal systems [91].

The minimization of these four free energy terms corresponds to determining the equilibrium displacement $\boldsymbol{u}(\boldsymbol{r})$ of filaments from their hexagonal lattice site (i.e. lattice point i having position vector $\boldsymbol{r}_i = \boldsymbol{R}_i + \boldsymbol{u}(\boldsymbol{r})$ where $\boldsymbol{R}_i$ is a hexagonal lattice vector), and the orientation of the filaments locally represented by the director field $\boldsymbol{n}(\boldsymbol{r})$. To simplify this problem, a variational guess was taken on the form of both $\boldsymbol{u}$ and $\boldsymbol{n}$, the latter having the form of a double twist with $\psi(r) = \psi_0' r$, which is analogous to the fibril structure anzatz put forth by Folkhard [78, 80] (see section 2.1.6 above). An analytic approach to the energy minimization was taken by expanding the free energy in terms of $\psi_0' R$, where R is the bundle radius. The authors found that the chiral nature of the filaments provided a mechanism in which the equilibrium bundle radius was finite in certain parameter regimes, and so equilibrium radius R and surface twist, $\psi_0' R$ were observed [89]. These two quantities are directly relevant to collagen fibril structure, the former being obvious in that it corresponds to the fibril radius R, and the latter corresponding to the surface twist $\psi(R)$ discussed in section 2.1.6.

This continuum model of hexagonally packed filaments shows great promise in being used (either in current form, or with slight modifications) to predict fibril structure, and has proposed a mechanism for radial control of fibrils due to the chiral nature of the underlying tropocollagen molecules. However, to be applied to collagen fibril structure, tweaks must be made both to the underlying free energy and the director field anzatz leading to a linear $\psi(r)$. The latter adjustment is easy to obtain in practice, as instead of constraining $\psi(r)$ to a specific functional form, one can just minimize the free energy as functional of $\psi(r)$ using standard calculus of variations techniques. This would lead to even more general predictions of molecular orientation within fibrils that could perhaps include constant angle fibril ultrastructure[14] $\psi(r) = \psi_0$ as proposed by Galloway [77] and discussed in section 2.1.6 above.

The adjustment of the underlying free energy is perhaps a more complex problem. Scattering studies of fibril structure have indicated a quasi-hexagonal packing of molecules/microfibrils within the fibril, and so removal of the 2D hexagonal packing energy included in Grason's model may provide a more accurate description of fibril ultrastructure. However, the main difficulty in modelling fibril structure is how best

---

[14] Ultrastructure refers to the details which can only be viewed under high resolution (e.g. electron microscopy) imaging. For a collagen fibril, this ultrastructure would include the fibril D-band and the molecular twist.

to describe the ubiquitous fibril D-band using continuum theory. It is not immediately clear how Grason's model can be modified to account for this peculiar molecular detail.

Comparison of Grason's model to the phase field liquid crystal model presented in this thesis

In this section, I briefly discuss the differences between the model of chiral filaments [82, 89] presented above, with the models that I will present in the remainder of this thesis, which for now I will refer to as the Phase-Field Collagen Fibril (PFCF) model (though it is not until Chapter 4 that the "phase-field" prefactor of this nomenclature is reasonable). There are three main differences between the Grason's model and the PFCF model. The first two are differences in the free energy structure, the final being a difference in how the free energy is minimized with respect to $\psi(r)$.

1. In contrast to Grason's model, the PFCF model does not impose any positional order of the molecules within a circular (2D) cross-section of the fibrils (perpendicular to the cylindrical axis of the fibril). This is reasonable for collagen fibrils, as experimental observation indicates that any sort of ordering of tropocollagen in this 2D cross-section is quasi-hexagonal at most [55].

2. We add an additional term in the free energy (Chapter 4 only) to include the D-band structure observed along the cylindrical axis of the collagen fibrils. This additional term includes additional observable quantities which must be optimized to minimize the free energy (a fibril D-band period and a fibril D-band amplitude).

3. We allow $\psi(r)$ to freely take any functional form, and chose $\psi(r)$ such that the free energy constructed is minimized.

With this distinction clearly stated, I now move on to the main work of my thesis. Chapter 3 presents numerical investigations of a simplified model of collagen fibril structure that neglects D-band (i.e. item 2 in the list above). Chapter 4 presents a complete model of collagen fibrils which includes D-band structure.

Chapter 3

Polymorphism of stable collagen fibrils

This chapter is adapted from a paper my supervisors and I have published in Soft Matter, volume 14, pages 4772 to 4783, 2018 [4], with permission from the Royal Society of Chemistry. For this paper my contributions were to write the code used in the numerical calculations, perform the analytical and numerical calculations, generate the figures, and write the first draft. I was an equal partner with my supervisors in revising the paper. The main change between this chapter and the corresponding paper is the replacement of the paper's conclusions section with a summary section which highlights the need for the more sophisticated model presented in Chapter 4.

Our motivation in writing this paper was to fully explore the predictive capabilities of a purely liquid crystalline model of collagen fibril structure (the framework of which having been developed previously, see ref. [3] for details). Previous work with this model outlined the type of calculations that could be performed and found marginal agreement between the model calculations and experimental observations of cornea (but not tendon) collagen fibril ultrastructure [3]. Furthermore, the authors did not demonstrate that the model could produce twist angles similar to those measured in tendon fibrils.Tendon fibrils are known to have values of surface twist which are smaller than 10°, but the results presented in ref. [3] did not include such small $\psi(R)$ values. However, this previous work did not fully explore the model parameter space due to computational limitations and the large number of parameters. Therefore, to fully explore the model predictions, we implemented a new numerical scheme to minimize the collagen fibril free energy which can more efficiently calculate the features of fibril ultrastructure for systematic parameter variation. This numerical scheme requires an initial guess for the minimum energy fibril configuration, for which we developed an analytical asymptotic approximation. We also reduced the number of parameters in the model from 5 to 3 using simple dimensional analysis arguments. These two improvements allowed us to completely map out the parameter space of the model, and to identify agreement with both reported tendon and cornea radial structure.

3.1 Introduction

Tropocollagen is the most abundant protein in the human body, integral to the structure of fibrous tissues such as skin, tendon, and cornea. There are at least 28 different tropocollagen molecules found in vertebrates [16], with types I, II, III, V, XI, XXIV, XXVII capable of forming the rope-like mesostructures that are collagen fibrils [92]. The assembly of tropocollagen molecules into collagen fibrils depends on the local environment. A suitable environment *in vivo* is within the extra-cellular space [93] after procollagen, a precursor to tropocollagen, is secreted from cells and cleaved by enzymes [94]. *In vitro*, ionicity, pH, and temperature of the solvent [36, 70], as well as concentration [95], have been shown to affect whether fibrillogenesis occurs.

Since fibrils are approximately cylindrical with radius R, it is convenient to separately consider their axial structure along the fibril's cylindrical axis and their radial structure within a circular cross-section. The axial D-banding has been well studied and remains close to 67 nm for both type I [96] and type II [97] collagen. Conversely, the observed radial ultrastructure of collagen fibrils depends on both the tropocollagen type and the anatomic location of the fibril *in vivo* [80, 98, 99, 100], and on solution conditions *in vitro* [36]. Factors such as temperature and ionicity of solution [36], or fibril age *in vivo* [80, 65], affect the observed fibril radii.

In this work, we focus on radial structure. We are concerned with what constrains the fibril radius, R, but also with the orientation of collagen molecules both on the fibril surface and within the fibril.

Observing the orientation of molecules on the *interior* of a circular cross-section of fibril is difficult experimentally, requiring diffraction studies [78] or electron tomography [55]. However, careful high resolution imaging can reliably characterize the molecular orientation at the *surface* of fibrils. Early work using transmission electron microscopy to image freeze-fractured fibrils found that molecules at the fibril surface were tilted with respect to the fibril axis, with the degree of tilt depending on where the fibrils were found anatomically [72, 74]. Further work demonstrated that tendon

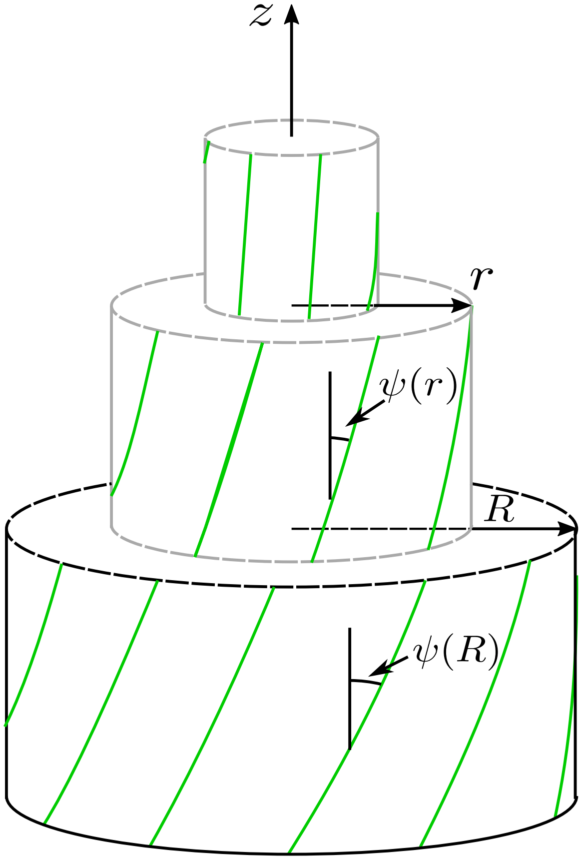

Figure 3.1: Double-twist configuration for a cylindrical fibril of radius R. The green lines represent the average orientation of collagen molecules for three different $r \in [0, R]$. The twist angle, $\psi(r)$, is the angle between the average orientation and the fibril (z) axis. The bottom cylinder illustrates the surface twist, $\psi(R)$. The smaller cylinders, outlined in grey, are cutaway views with $r < R$.

fibrils, while exhibiting a large range of R values (15 nm-200 nm), have limited molecular surface tilt $\simeq 5°$ [101, 79], while corneal fibrils, with a narrower range of R from $15\,\text{nm} - 20\,\text{nm}$, exhibit much larger surface tilt, $\simeq 18°$ [102, 103, 55]. Different hypotheses of radial molecular orientation have been proposed to fit these experimental results [80, 55, 79] – but without consideration of thermodynamic stability.

Recently, an equilibrium liquid crystal model of radial collagen fibril structure was developed to predict molecular configurations of tropocollagen within individual fibrils [3]. Consistent with the surface tilt observations mentioned above, a double-twist geometry of tropocollagen molecules was imposed within the fibril. With this double-twist geometry (see Fig. 3.1), the twist angle of molecules with respect to the fibril axis at a given radial distance, $\psi(r)$, fully describes the molecular orientation of the tropocollagen molecules. The corresponding elastic free energy functional [104], valid for arbitrary smoothly varying $\psi(r)$, is parameterized by the costs of twist distortion, K_{22}, bend distortion, K_{33}, saddle-splay distortion, k_{24}, surface tension, γ, and the preferred pitch of a cholesteric phase, $2\pi/q$. By minimizing the free energy per unit volume of fibril with respect to $\psi(r)$, the equilibrium fibril radius, R_{eq}, and the surface twist angle, $\psi(R_{\text{eq}})$ were determined for different values of the model parameters. These R_{eq} and $\psi(R_{\text{eq}})$ were then compared with experimental findings. The model [3] showed good agreement with corneal fibrils, which have small radius and large surface tilt. However, it was unclear whether it could also capture the smaller surface tilt and the broad range of radii observed for tendon fibrils.

The physical mechanism of collagen fibril formation *in vivo*, as well as the self-assembly of tropocollagen molecules into collagen fibrils *in vitro*, is poorly understood. *In vitro* studies [95] have demonstrated that uniform fibril formation will occur without cross-linking or other non-equilibrium processes. This suggests that an equilibrium description of fibrils is appropriate, at least for *in vitro* fibrillogenesis. The importance of collagen in biotechnology applications is therefore sufficient motivation for us to further explore the equilibrium picture of radial fibril structure. However, it is attractive to hypothesize that fibrillogenesis *in vivo* also exploits equilibrium self-assembly processes. Better understanding whether and how equilibrium processes could lead to observed radial collagen structures would help us identify when non-equilibrium processes may also be affecting fibril structure.

In this chapter, we use an efficient numerical relaxational method to expand on previous work with the double-twist model, which allows us to map out equilibrium values of fibril radius, surface twist angle, and energy per unit volume of fibril within the entire parameter space of stable fibrils. Using dimensional analysis, we show that just three reduced parameters fully control the experimentally observable behaviour of the system. We use this comprehensive approach to confront both corneal and tendon fibril phenomenology.

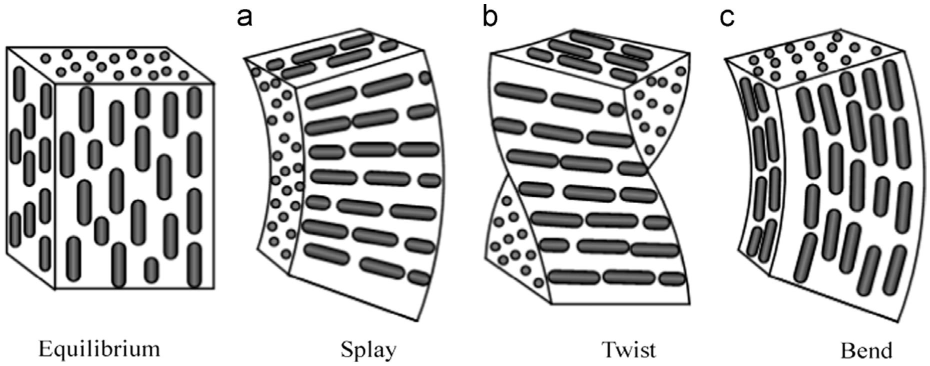

Figure 3.2: Geometric representation of (a) splay, (b) twist, and (c) bend distortions in a liquid crystal nematic. These three distortions have free energy costs proportional to the K_{11}, K_{22}, and K_{33} elastic constants discussed in the text. Figure reprinted from [105] (Copyright 2014) with permission from Elsevier.

3.2 Model

3.2.1 Elastic free energy density

Individual tropocollagen molecules within collagen fibrils are essentially rod-like, with a length of $\sim 300\,\mathrm{nm}$, and a diameter of $\sim 1.5\,\mathrm{nm}$. To describe the molecular orientation within fibrils, we use a director field, $\boldsymbol{n}(\boldsymbol{r})$, which is a unit vector that represents the local, average orientation of molecules within the fibril.

Following earlier work [3], we propose that the fibril free energy depends on elastic energy contributions from the orientation field $\boldsymbol{n}(\boldsymbol{r})$ together with an interfacial energy. We use a leading order gradient expansion for the elastic contributions. The elastic free energy density [104] of a chiral liquid crystal system with no external stress is

$$\begin{aligned} f_{\mathrm{el}} =& \frac{1}{2}K_{11}\left(\nabla\cdot\boldsymbol{n}\right)^2 + \frac{1}{2}K_{22}\left(\boldsymbol{n}\cdot\nabla\times\boldsymbol{n} + \frac{k_2}{K_{22}}\right)^2 \\ &+ \frac{1}{2}K_{33}(\boldsymbol{n}\times(\nabla\times\boldsymbol{n}))^2 + k_{13}\nabla\cdot(\nabla\cdot\boldsymbol{n})\,\boldsymbol{n} \\ &- \frac{1}{2}(K_{22}+k_{24})\nabla\cdot(\boldsymbol{n}\times(\nabla\times\boldsymbol{n}) + \boldsymbol{n}(\nabla\cdot\boldsymbol{n}))\,, \end{aligned} \tag{3.1}$$

where we have taken $f_{\mathrm{el}} = 0$ in the cholesteric phase. From the last two terms, we see that it is possible to have $f_{\mathrm{el}} < 0$ even when all elastic constants are positive. In using this free energy, we assume that any gradients in $\boldsymbol{n}$ are slowly varying compared to the molecular length scale ($\simeq 1.5\,\mathrm{nm}$). Higher order gradient terms are thereby

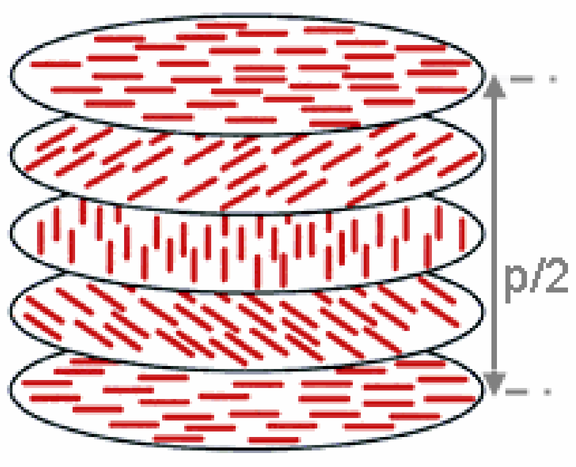

Figure 3.3: Schematic of a cholesteric liquid crystal phase. p is the pitch of the cholesteric, which is inversely related to $q = 2\pi/p$. Figure reprinted from Wikipedia under CC BY-SA [108].

ignored [106].

Each term in eqn. 3.1 corresponds to a specific distortion. The terms with K_{11}, K_{22}, and K_{33} correspond to the usual splay, twist, and bend deformations [107] shown in Figure 3.2, and are always greater than zero. k_2 is the "chiral strength" and can be of either sign. k_{13} and k_{24} are the splay-bend and saddle-splay elastic constants, respectively. The terms with k_{13} or k_{24} can be negative, and when integrated will appear as surface terms. They contribute to equilibrium phases that have a proliferation of interfaces, such as a system of collagen fibrils.

3.2.2 Cholesteric and double-twist fibril phases

Equilibrium phases of collagen molecules are determined by the form of $\boldsymbol{n}$ that minimizes the total free energy of the system. We consider two phases. The first is a bulk cholesteric phase (see Figure 3.3), which has been observed for concentrated tropocollagen solutions *in vitro* [95]. The director field is e.g. $\boldsymbol{n} = \cos(qz)\hat{\boldsymbol{x}} + \sin(qz)\hat{\boldsymbol{y}}$, where $q \equiv k_2/K_{22}$ here determines the inverse cholesteric pitch of the cholesteric phase. Inserting this into eqn. 3.1 gives $f = 0$. Since the cholesteric phase is a bulk phase, any surface effects are negligible and the total free energy per unit volume, $E_{\text{cholesteric}} = 0$, for all values of the elastic constants. Any phase with bulk average free energy density $E < 0$ is therefore thermodynamically stable with respect to the cholesteric phase.

The second phase we consider has individual fibrils with a double-twist director field [3],

$$\boldsymbol{n} = -\sin\psi(r)\hat{\boldsymbol{\phi}} + \cos\psi(r)\hat{\boldsymbol{z}}, \tag{3.2}$$

where $\psi(r)$ is the angle between the director field and the fibril axis. $\psi(R)$ is then the "surface twist" (molecular tilt) of a fibril of radius R. Since we are interested in the radial structure, we ignore contributions from axial packing (e.g. D-banding) of collagen molecules along the fibril. This amounts to an assumption that coupling between radial and axial structure is weak (see Discussion). Excluding radial/axial coupling greatly simplifies our calculations.

A surface energy term must be included to account for the cost of creating an interface between individual fibrils and the surrounding fluid. For a single fibril, the free energy per unit length is then

$$E_L \equiv 2\pi \int_0^R r f_{\text{fibril}}(r, \psi(r), \psi'(r)) dr + 2\pi\gamma R \tag{3.3}$$

where $2\pi\gamma R$ is the energetic cost of the interface between the fibril of cross-sectional circumference $2\pi R$ and its surroundings [3]. (Note that while the bulk k_{13} and k_{24} terms of eqn. 3.1 integrate mathematically into surface contributions, they are distinct from the interfacial cost γ.) The cross-sectional area of a single fibril is πR^2. Thus,

$$E(R) = \frac{E_L}{\pi R^2} = \frac{2}{R^2}\int_0^R r f_{\text{fibril}} dr + \frac{2\gamma}{R}, \tag{3.4}$$

where E is the total free energy per unit volume of fibril. We refer to the relationship between E and R as the energy landscape.

Using the double-twist structure eqn. 3.2 in the elastic free energy density eqn. 3.1 gives the free energy density [3],

$$\begin{aligned} f_{\text{fibril}} =& \frac{1}{2}K_{22}\left(q - \psi' - \frac{\sin 2\psi}{2r}\right)^2 + \frac{1}{2}K_{33}\frac{\sin^4\psi}{r^2} \\ & - \frac{1}{2}(K_{22} + k_{24})\frac{1}{r}\frac{d\sin^2\psi}{dr}. \end{aligned} \tag{3.5}$$

where here $q = k_2/K_{22}$ is the chiral wavenumber of the double-twist phase. Note that the K_{11} and k_{13} terms have dropped out since $\nabla \cdot \boldsymbol{n} = 0$ for double-twist.

Minimizing eqn. 3.4 with respect to the function $\psi(r)$ using standard calculus of variations techniques [3], we arrive at the boundary value problem

$$(r\psi')' = q + \frac{K_{33}}{K_{22}}\frac{\sin(2\psi)}{r}\sin^2\psi - \cos(2\psi)\left(q - \frac{\sin(2\psi)}{2r}\right), \tag{3.6a}$$

$$\psi(0) = 0, \tag{3.6b}$$

$$\psi'(R) = q + \frac{k_{24}}{K_{22}}\frac{\sin(2\psi(R))}{2R}, \tag{3.6c}$$

where eqn. 3.6c is a natural boundary condition which follows from the functional minimization procedure, and $\psi' \equiv d\psi/dr$. We must have $\psi(0) = 0$, as any non-zero twist at $r = 0$ would imply singular f_{fibril} and an infinite E from eqns. 3.4 and 3.5.

Dimensional Analysis

While there are five parameters which control the behaviour of our model, q, γ, K_{22}, K_{33}, and k_{24}, we can reduce this to three dimensionless variables (see Supplemental section 3.7), $\tilde{K}_{33} = K_{33}/K_{22}$, $\tilde{\gamma} = \gamma/(K_{22}q)$, and $\tilde{k}_{24} = k_{24}/K_{22}$, which we utilize for the remainder of the chapter. This lets us express quantities of interest in terms of dimensionless parameter combinations:

$$\tilde{E} = g_1\left(qR, \tilde{K}_{33}, \tilde{\gamma}, \tilde{k}_{24}\right), \tag{3.7a}$$

$$qR = g_2\left(\tilde{K}_{33}, \tilde{\gamma}, \tilde{k}_{24}\right), \tag{3.7b}$$

$$\psi(qr) = g_3\left(qr, \tilde{K}_{33}, \tilde{\gamma}, \tilde{k}_{24}\right), \tag{3.7c}$$

where the functions g_1, g_2, and g_3 are determined numerically, $\tilde{E} \equiv E/(K_{22}q^2)$, and we solve ψ as a function of dimensionless radius qr. We have reduced our parameter space from five to three dimensions, together with an inverse length q that sets the scale for R.

The elastic constants for collagen solutions are not well documented. We use values determined experimentally from liquid crystal systems with molecules similar to tropocollagen molecules. For the poly-peptide α-helical chain poly-γ-benzyl-L-glutamate (PBLG), the ratio of bend to twist elastic constant saturates at $K_{33} \simeq 30K_{22}$ for aspect ratios $L/D \gtrsim 100$, where L is the length and the diameter D of PBLG is between 1.5 nm to 2.5 nm [109]. The aspect ratio of tropocollagen, $L/D = 200$, then leads us to use $\tilde{K}_{33} = 30$ for this chapter. (In Supplemental section 3.6 we explore the effects of different $\tilde{K}_{33}$ values on our results for the surface twist.) Differences in solution conditions, molecular composition, and concentration can in principle affect $\tilde{K}_{33}$ [110, 111, 112]; however, approximately the same ratio is observed over a range of temperature and concentration in long-aggregates of lyotropic, chromonic liquid crystals [113][1].

3.2.3 Energy Minimization

Using the methods detailed in Appendix B.1, we solve eqns 3.6 numerically with a finite-difference relaxation scheme [114]. We have also derived an explicit (but unwieldy) power-series solution, see Supplemental section 3.8. We use the leading cubic terms of this power-series as an initial guess for our relaxation approach, and use higher-order solutions as occasional checks that the relaxation approach has converged. The iterated relaxation converges on the $\psi(qr)$ that minimizes the dimensionless version of eqn. 3.4 for a selected qR. We repeat this procedure for different qR to determine the energy landscape, $\tilde{E}(qR)$, for a given parameter set [3].

We are particularly interested in the dimensionless radius qR_{eq} that minimizes $\tilde{E}(qR)$. To find qR_{eq}, we used a standard golden ratio search. Our search bounds were $qR \in [10^{-5}, 1]$. If $\tilde{E}(qR_{\mathrm{eq}}) \equiv \tilde{E}_{\mathrm{eq}} < 0$ for a set of parameter values, then the bulk fibril phase is an equilibrium phase with respect to the cholesteric for those parameters. To avoid cumbersome notation, we will use the equilibrium result $\psi(qr) \equiv \psi_{\mathrm{eq}}(qr)$ unless otherwise noted.

3.3 Results

3.3.1 Narrow equilibrium regime

In Fig. 3.4, we show the *global* energy landscape for double-twist collagen fibrils as the dimensionless parameters $\tilde{\gamma}$ and $\tilde{k}_{24}$ are varied. The ratio $\tilde{K}_{33} = 30$ is held constant. The colour and contours represent the dimensionless minimum energy, $\tilde{E}_{\mathrm{eq}}$, for double-twist fibrils — green indicates equilibrium fibrils with respect to the cholesteric phase, while red indicates meta-stable fibrils. We see that there is only a small region of

---

[1] Lyotropic liquid crystals gain orientational order with increasing concentration. Chromonic liquid crystals are formed through aggregation of smaller, flat molecules which stack on top of each other to form cylindrical structures.

equilibrium fibrils, where we require $-1 \leq \tilde{k}_{24} \lesssim 1.2$ and $\tilde{\gamma} \lesssim 0.2$. The minimum fibril energy $\tilde{E}_{\text{eq}}$ increases monotonically with increased $\tilde{\gamma}$ or with decreasing $\tilde{k}_{24}$.

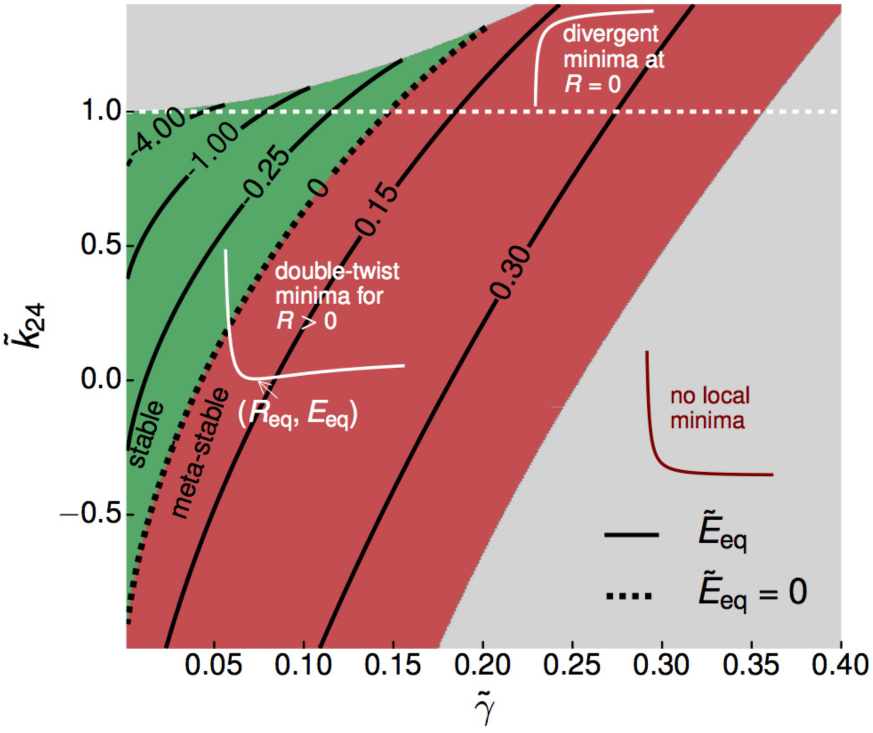

Figure 3.4: Phase diagram of stable and meta-stable double-twist phases in the $\tilde{k}_{24}$ *vs* $\tilde{\gamma}$ plane (with $\tilde{K}_{33} = 30$). The green region indicates the existence of double-twist fibrils that are stable with respect to the cholesteric phase, with $E_{\text{eq}} < 0$. The red region indicates meta-stable minima with respect to the cholesteric phase, with $E_{\text{eq}} \geq 0$. Contours indicate the values of the dimensionless free energy density $\tilde{E}_{\text{eq}} \equiv E_{\text{eq}}/(K_{22}q^2)$. The inset white curve labelled "double-twist minima for $R > 0$" demonstrates a typical relationship between E and R for values of $\tilde{\gamma}$ with $\tilde{k}_{24} \leq 1$. For $\tilde{k}_{24} > 1$ (above dashed white line), there is an additional, divergent global minimum as $R \to 0$, illustrated by the inset curve labelled "divergent minima at $R = 0$". Gray regions do not have any local minima with $0 < R < \infty$. Note that $\tilde{k}_{24} \equiv k_{24}/K_{22}$, $\tilde{\gamma} \equiv \gamma/(K_{22}q)$, and $\tilde{K}_{33} \equiv K_{33}/K_{22}$.

The energies shown in Fig. 3.4 represents the energy of double-twist fibrils that have a finite radius R. For larger values of $\tilde{\gamma}$ there is no local minimum at $R > 0$ (gray region), and we would instead expect to observe a bulk cholesteric phase. This is also what we expect in most of the metastable regime, and arises because the energy cost of the interface in a fibril phase is large due to the surface tension γ.

For $\tilde{k}_{24} > 1$, we observe a divergent minimum energy for $R \to 0$. When both a divergent minimum for $R \to 0$ and a local minimum at finite R is present, we illustrate the local minimum behaviour only (i.e. shading and contours in non-gray regions with $\tilde{k}_{24} > 1$ represent the local minima). This divergent minimum arises because sufficiently large k_{24} encourages interface proliferation in the fibril phase. This can be seen explicitly with eqns. 3.4 and 3.5 using a linearly varying ansatz for the pitch, $\psi = r\psi(R)/R$. For $\psi(R) \ll 1$ we obtain $E = \psi(R)^2(K_{22} - k_{24})/R^2 + 2\gamma/R$. For $k_{24} > K_{22}$ we obtain $E \to -\infty$ as $R \to 0$. However, this singular solution is for a continuum model where fibril radii are large with respect to the diameter of individual molecules, $d = 1.5\,\text{nm}$. We would also expect higher order gradient terms, absent in eqn. 3.1, to change (and perhaps eliminate) the singular solution at $R \approx 0$ for $\tilde{k}_{24} > 1$.

To confront our double-twist solutions with experimental measurements of collagen fibrils, we investigate our model's predictions of surface twist, $\psi_{R_{eq}} \equiv \psi(qR_{\text{eq}})$, and fibril radius, R_{eq}.

3.3.2 Experimental observables: Surface twist and fibril radius

Fig. 3.5 shows the surface twist landscape. Corresponding with Fig. 3.4, the gray regions at the upper left and to the right have no fibril phases. $\psi_{R_{\text{eq}}}$ increases with increasing $\tilde{\gamma}$ and decreasing $\tilde{k}_{24}$, with blue lines of constant $\psi_{R_{\text{eq}}}$ (in radians) shown. Double-twist phases that are stable with respect to the bulk cholesteric phase occur to the left of the black dashed line ($E_{\text{eq}} < 0$), as indicated.

Two surface twist values of particular interest are $\psi_{R_{\text{eq}}} = 0.1\,\text{rad}$ and $\psi_{R_{\text{eq}}} = 0.31\,\text{rad}$, being typical surface twist angles observed in tendon fibril and corneal fibril, respectively. We have labelled these two values of surface twist with blue dashed lines in Fig. 3.5. Furthermore, other types of fibrils in vivo tend to have smaller surface twists than corneal fibrils $\leq 0.31\,\text{rad}$, which gives the corneal dashed line in Fig. 3.5 further meaning as an upper limit of surface twist values observed in vivo [74, 72, 80]. Remarkably, this upper bound of surface twist approximately coincides with the *stable equilibrium* regime of double-twist fibrils (i.e. the region to the left of

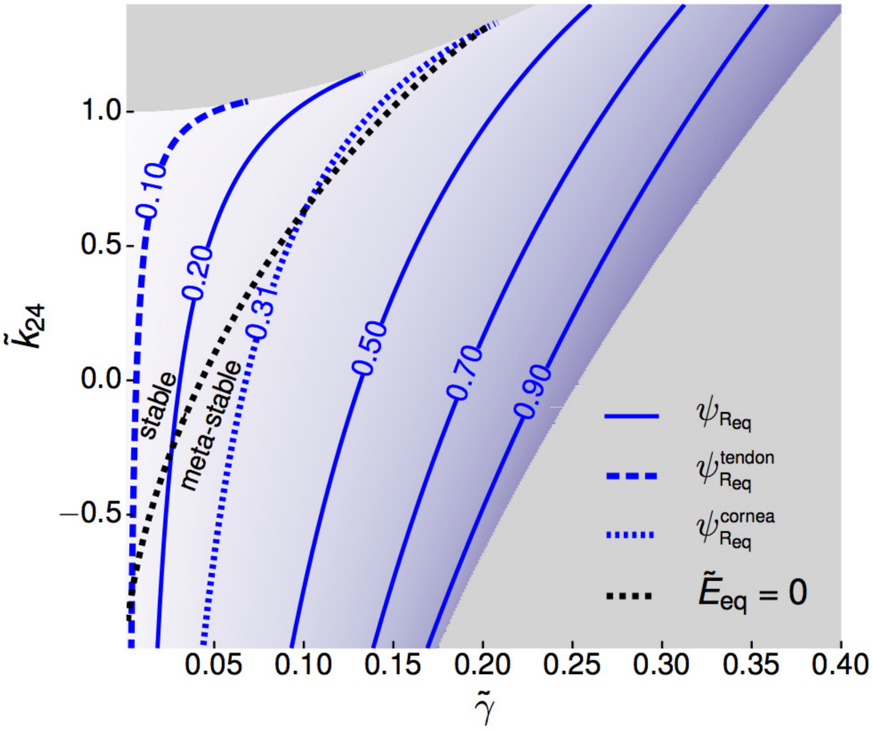

Figure 3.5: Contours of surface twist $\psi_{R_{\mathrm{eq}}} \equiv \psi(qR_{\mathrm{eq}})$ (solid and dashed blue) in radians, *vs* the reduced saddle-splay elastic constant $\tilde{k}_{24}$ and the reduced surface tension $\tilde{\gamma}$, all with $\tilde{K}_{33} = 30$. $\psi_{R_{\mathrm{eq}}} = 0.1\,\mathrm{rad}$ and $\psi_{R_{\mathrm{eq}}} = 0.31\,\mathrm{rad}$ are typical surface twists observed in tendon and cornea fibrils, respectively, and are distinguished above with dashed contour lines. Meta-stable ($E_{\mathrm{eq}} \geq 0$) and stable ($E_{\mathrm{eq}} < 0$) fibril phases with respect to the bulk cholesteric phase are separated by the black, dashed line. The gray areas correspond to parameter space regions for which no stable or meta-stable double-twist configurations are found. Note that $\tilde{k}_{24} \equiv k_{24}/K_{22}$, $\tilde{\gamma} \equiv \gamma/(K_{22}q)$, and $\tilde{K}_{33} \equiv K_{33}/K_{22}$.

the black dashed line in Fig. 3.5).

We also obtain reduced equilibrium fibril radii, qR_{eq}, as shown in Fig. 3.6. As a consequence of eqn. 3.7b, we do not obtain the radii directly. We see that qR_{eq} increases with increasing $\tilde{\gamma}$, and decreases with increasing $\tilde{k}_{24}$ — the same qualitative behaviour as $\psi_{R_{\mathrm{eq}}}$. For fixed q the behaviour of R_{eq} as other parameters are varied is immediately given: the radius decreases as k_{24} increases, or as the surface tension γ

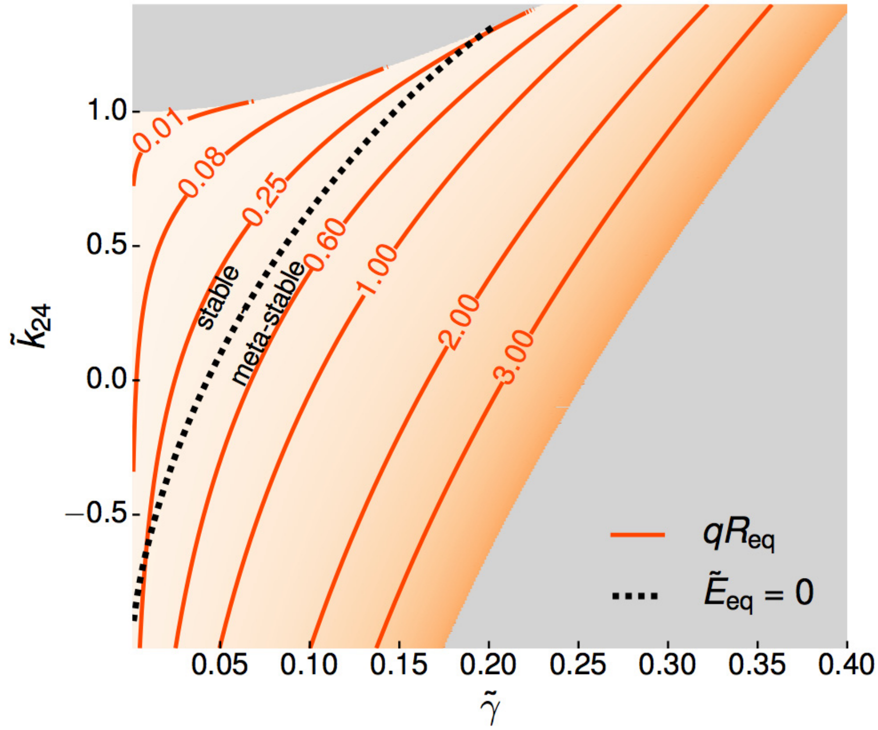

Figure 3.6: Contours of scaled equilibrium fibril radius $qR_{\rm eq}$ as a function of the reduced saddle-splay elastic constant, $\tilde{k}_{24}$, and the dimensionless surface-tension $\tilde{\gamma}$, all with $\tilde{K}_{33} = 30$. $qR_{\rm eq}$ increases with increasing $\tilde{\gamma}$, and decreases with increasing $\tilde{k}_{24}$. Values of $qR_{\rm eq}$ to the left of the black, dashed line are stable with respect to the bulk cholesteric phase ($\tilde{E}_{\rm eq} < 0$). The gray areas correspond to parameter space regions for which no stable or meta-stable double-twist configurations are found. Note that $\tilde{k}_{24} \equiv k_{24}/K_{22}$, $\tilde{\gamma} \equiv \gamma/(K_{22}q)$, and $\tilde{K}_{33} \equiv K_{33}/K_{22}$.

decreases. Increasing K_{22} rescales $\tilde{\gamma}$ and $\tilde{k}_{24}$ directly towards the origin, and can either increase $R_{\rm eq}$ (for fibrils with small $\psi_{R_{\rm eq}}$) or decrease $R_{\rm eq}$ (for fibrils with $\psi_{R_{\rm eq}} \gtrsim 0.2\,\text{rad}$, or 10°). If we increase q and leave other parameters fixed, we see that the scaled surface-tension $\tilde{\gamma}$ will decrease — leading to smaller $qR_{\rm eq}$ values. Since we have increased q, we then obtain even smaller $R_{\rm eq}$ values.

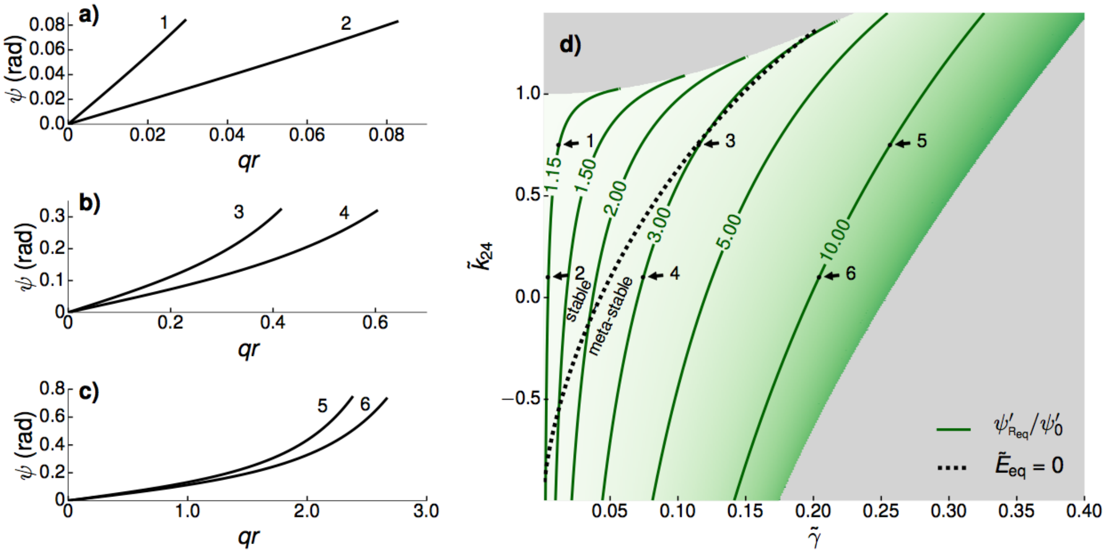

Figure 3.7: Non-linearity of the twist angle $\psi(r)$ $(0 \leq r \leq R_{\mathrm{eq}})$, where R_{eq} is the fibril radius within the double twist fibril phase, for different parameter values. In (a)-(c), six different double-twist configurations, ψ *vs* scaled radial distance qr, are illustrated for the parameter values indicated in (d) – with corresponding labels from 1-6. 1, 3 and 5 are points on the $\tilde{k}_{24} = 0.75$ line; 2, 4, and 6 are points on the $\tilde{k}_{24} = 0.1$ line. Both ψ and ψ' increase monotonically with r for all parameter values. The contours in (d) indicate the ratio of surface twist gradient at the surface to that in the fibril centre, $\psi'(qR_{\mathrm{eq}})/\psi'(0) \equiv \psi'_{R_{\mathrm{eq}}}/\psi'_0$, which captures non-linearities in the double-twist configuration. As before, the black dashed line separates fibrils that are stable with respect to the bulk cholesteric phase (left of line) from those which are only meta-stable (right of line). The gray areas of (d) correspond to parameter space regions for which no stable or meta-stable double-twist configurations are found.

3.3.3 Non-linearity of twisting within fibril

From our free energy functional, at the fibril centre collagen molecules are aligned with the fibril axis, with $\psi(0) = 0$. For $r > 0$, we illustrate $\psi(qr)$ for six parameter values in Fig. 3.7a-3.7c. All of the curves exhibit two properties: 1) $\psi(qr)$ increases monotonically with qr, and 2) the twist gradient also increases with radius, i.e. $\psi''(qr) > 0$. With these two properties in mind, we quantify the double-twist nonlinearity with the ratio of the twist angle gradient at the fibril surface, $\psi'_{R_{\mathrm{eq}}} \equiv \psi'(qR_{\mathrm{eq}})$, to the twist angle gradient at the fibril centre, $\psi'_0 \equiv \psi'(0)$, as shown in Fig. 3.7d. Nonlinearity increases with increasing $\tilde{\gamma}$, and decreases with increasing $\tilde{k}_{24}$. We see that equilibrium fibrils may have significant twist nonlinearities, up to $\psi'_{R_{\mathrm{eq}}}/\psi'_0 \approx 3$.

3.4 Discussion

We identified dimensionless parameter combinations (eqn. 3.7) that reduced the number of independent parameters in our equilibrium free energy density (eqn. 3.4) for the collagen orientation within double-twist fibrils (eqn. 3.2). We solved the dimensionless equations numerically, and identified a narrow parameter regime (green region of Fig. 3.4) that produces double-twist fibrils that are thermodynamically stable with respect to a bulk cholesteric phase.

The parameters of our model are the coarse-grained elastic constants that determine the free energy costs of spatial-gradients of the collagen orientation (K_{22}, K_{33}, K_{22}, k_{24}) together with a surface energy γ and a chiral wavenumber q. One dimensionless parameter combination is relatively well determined by the long semi-flexible configuration of individual collagen molecules ($K_{33}/K_{22} = 30$). Remarkably, we find that only two dimensionless parameter combinations (k_{24}/K_{22} and $\gamma/(K_{22}q)$) are then required to determine both the surface twist $\psi_{R_{\mathrm{eq}}}$ (Fig. 3.5) and the dimensionless radius qR_{eq} (Fig. 3.6) of equilibrium collagen fibrils.

We find that equilibrium surface twists should all satisfy an upper bound: $\psi_{R_{\mathrm{eq}}} \leq 0.33\,\mathrm{rad}$ (19°), which approximately coincides with the maximum surface twist reported in the *in vivo* literature [55].

3.4.1 Polymorphism of collagen fibrils

A surprise in considering Figs. 3.4, 3.5, and 3.6 is the wide range of equilibrium configurations available to collagen fibrils over a relatively narrow parameter regime. This polymorphism allows different aspects of fibril structure to be emphasized for different parameterizations.

Collagen fibril stability

The thermodynamic stability with respect to the cholesteric phase is assessed by the free energy per unit volume, as illustrated in Fig. 3.4. We see that the most stable (lowest energy) fibrils are in the upper-left corner with a combination of small γ and large K_{22} and q — above point "1" in Fig. 3.7, with $k_{24} \simeq K_{22}$. Note that what is presented is $E_{\mathrm{eq}}/(K_{22}q^2)$, so that with large K_{22} and q the cohesion energy is even larger.

One consequence of selecting for more stable fibrils is that the expected surface twist values would be quite small, according to Fig. 3.5. Interestingly, we would expect a uniform twist gradient (Fig. 3.7) in this regime as well. In contrast, to allow for fibrils with larger surface twist, $\tilde{\gamma}$ must be fine-tuned to values near the stability boundary — close to point “3” in Fig. 3.7 — making fibrils with large surface twist (and nonlinear $\psi(qr)$) less thermodynamically stable than their small twist, linear counterparts. We note that all fibrils, which are stable with respect to the cholesteric phase, have $\psi_{R_{\text{eq}}} \leq 0.33\,\text{rad}$.

The relationship between thermal stability and fibril radius is complicated by the scaling of R_{eq} with q, as the contours in Fig. 3.6 depend on q as well as $\tilde{\gamma}$ and $\tilde{k}_{24}$. Thus, to investigate the relationship between thermal stability and fibril size, we look at the two ways in which large (small) radius equilibrium fibrils can be generated from our model. The first is to maximize (minimize) qR_{eq} at a constant q. From Fig. 3.6, this would be achieved by fine-tuning $\tilde{\gamma}$ close to (far from) the stability boundary. This approach would indicate that smaller fibrils are more thermodynamically stable than large fibrils.

The second approach to generate large (small) fibrils is to decrease (increase) the chiral wavenumber q at a constant qR_{eq}, while also keeping $\tilde{\gamma}$ and $\tilde{k}_{24}$ constant. In this approach, you would stay at the same point in Fig. 3.4 and 3.6, and so $E_{\text{eq}}/(K_{22}q^2)$ and qR_{eq} would remain constant. As you decrease (increase) q, fibril radius increases (decreases), but thermodynamic stability decreases (increases) as well. Thus, both approaches to increasing fibril radius tend to decrease thermal stability. Given this prediction, it is unclear what functional role large fibrils might have, if it is not to increase stability. While large fibrils are expected to be individually stronger than small ones, the packing fraction of large or small fibrils would be the same and so would bulk moduli of closely packed fibrils.

Influence of Collagen types

Collagen fibrils *in vivo* generally contain a tissue-dependent mixture of collagen types [115, 48]. For example, while well-studied tendon and corneal fibrils are predominantly composed of type-I collagen they contain an admixture of type-III collagen [116]. The best characterized heterotypic mixtures *in vitro* has been blends of types I and III collagen [117, 118, 19], though I/V [71, 119] and II/III blends [120] have also been studied.

The distribution of collagen types within individual fibrils has been qualitatively assessed from immunoassay double-labelling. Both type I and type III are seen on fibril surfaces [118, 117, 19] indicative of homogeneity (the evidence is, however, mixed[48]). Under the assumption that mixtures of collagen types are spatially homogeneous within a fibril, the elastic parameters of the mixture should be interpolations between those of the pure collagen types [111]. In which case, our equilibrium picture would apply to heterotypic fibrils — and the reduced elastic parameters of mixtures would sit on curves between those of the pure types.

Varying the composition of heterotypic I/III fibrils leads to variations of fibril radius [19] — from 0.1 µm (entirely type I) to 0.025 µm (entirely type III). Our model can reproduce that either by moving the reduced parameters, e.g. $\tilde{\gamma}$, or by changing q. Changes to $\tilde{\gamma}$ would be associated with a change in the surface twist, while changes to q could be assessed in the cholesteric phase. However, neither surface twist nor cholesteric q have been systematically characterized in type I/III mixtures.

D-band spacing

While we have assumed that the radial and longitudinal structures are decoupled, a simple projective-coupling has been proposed in the literature [121, 48], corresponding to the D-band period being reduced by a factor of $\cos(\psi)$ due to non-zero twist. For our nonlinear double-twist model, the question immediately arises about how a single D-band spacing can represent a continuously varying twist, $\psi(qr)$. We hypothesize that surface measurements of the D-band period via scanning electron microscopy or atomic force microscopy would probe surface twist $\psi(qR)$ while bulk measurements of the D-band period via transmission electron microscopy or X-ray scattering would probe a volume-average twist $\langle\cos(\psi(qr))\rangle$. Combining both types of measurements on the same set of fibrils would then provide additional insight into the nature of the radial and longitudinal coupling.

Our model has a maximal surface twist of 0.33 rad, and a minimal twist of $0.002 \simeq 0$ rad, corresponding to at most a 5% difference of D-band spacing between fibrils according to the projective-coupling hypothesis. While surface twist of heterotypic I/III fibrils has not been characterized, the D-band spacing has been [19]. For 100% collagen-III (compared to pure collagen-I fibrils) there is a significant 39% decrease in the D-band spacing. This exceeds our maximal surface twist effect, but could be attributed to changes in the gap-spacing of the D-band [97] or to rope-like ultrastructure [18] rather than to molecular tilt. Experimentally relating surface twist measurements of fibrils to a more detailed assessment of longitudinal structure and ultrastructure would be desirable to untangle these effects.

3.4.2 Experimental guidance on elastic parameters

The chiral wavenumber q can be directly assessed within cholesteric phases through the cholesteric pitch $P = 2\pi/q$. Polarized light microscopy observations of rat tail tendon tropocollagen solubilized in acid show that cholesteric phases emerge at concentrations above $50\,\mathrm{mg\,mL^{-1}}$, with decreasing pitch from $P \simeq 20\,\mu\mathrm{m}$ at $\sim 50\,\mathrm{mg\,mL^{-1}}$ to $P \simeq 0.5\,\mu\mathrm{m}$ at $\sim 400\,\mathrm{mg\,mL^{-1}}$ [122]. While we might expect variation in q for fibrils due to variable solution conditions [110, 111], we expect a similar range of values $q \in [0.1\pi\mu\mathrm{m}^{-1}, 4\pi\mu\mathrm{m}^{-1}]$.

The surface tension, γ, quantifies the cost of an interface between two bulk phases. In our case, the interface is between individual fibrils and the surrounding aqueous collagen solution. No experimental measurements of γ have been reported for collagen. However, we assume surface-tensions are similar in magnitude to the nematic-isotropic interface for liquid crystal systems. A lower bound of surface tension of an isotropic-nematic interface is that of p-azoxyphenetole, for which $\gamma \gtrsim 0.5\,\mathrm{pN\,\mu m^{-1}}$ [123]. Conversely, a larger value of γ reported in this type of system is that of MBBA, with $\gamma = 24\,\mathrm{pN\,\mu m^{-1}}$ [124, 125]. Other experimental values fall within this range [126, 127, 128]. Using Onsager's theory of hard rods [110], a theoretical expression of γ has been derived for isotropic-nematic interfaces near the phase transition [129]. Applying this result to our system, we obtain $\gamma \sim 2.3\,\mathrm{pN\,\mu m^{-1}}$ which is consistent with the experimental bounds. Accordingly, we expect $\gamma \in [0.5\,\mathrm{pN\,\mu m^{-1}}, 25\,\mathrm{pN\,\mu m^{-1}}]$.

To determine the value of the twist elastic constant, K_{22}, for collagen fibrils, we again use typical values of liquid crystal systems. For PBLG, a range of K_{22} values from $0.6\,\mathrm{pN}$ to $6.2\,\mathrm{pN}$ [130, 131, 132] have been measured depending on the solvent used. In these measurements, no significant concentration [131] or molecular weight [132] dependence has been observed. We therefore expect $K_{22} \in [0.6\,\text{pN}, 6\,\text{pN}]$.

Experimentally determining the saddle-splay elastic constant, k_{24}, is difficult due to the surface-like nature that it represents in the free energy. The saddle-splay to twist ratio has been estimated to be $k_{24}/K_{22} \simeq 2$ for nematic systems using deuterium nuclear-magnetic-resonance [133] and polarization microscopy [134]. No measurements of k_{24} for long, chiral molecules similar to tropocollagen have been reported. Theoretical calculations predict that $k_{24} = \frac{1}{2}(K_{11} - K_{22})$ [106], which with $K_{11} > K_{22}$ [135, 112] implies $k_{24} \geq 0$. However, this result was derived through an interaction energy, and thus is likely valid only for thermotropic systems in which energy favours orientational order [2].

3.4.3 Comparison with *in vivo* fibril ultrastructure

Our theoretical equilibrium treatment highlights the importance of surface twist, since it significantly constrains our model parameterization. (The comparison between experiment and our model is not as definitive when looking at R_{eq}, because we can only constrain the product qR_{eq}.) The surface twist angle measured *in vivo* is correlated to the anatomical location of the fibril, as well as the type of tropocollagen found within the fibril [74, 73, 98, 80, 48]. Two well-studied fibril types *in vivo* are corneal fibrils, which have large surface twists $\simeq 0.31\,\text{rad}$ [55], and tendon fibrils, which have fairly small surface twists $\simeq 0.1\,\text{rad}$ [54].

Corneal and other helicoidal fibrils

For the high surface twist of corneal collagen fibrils, with $\psi_{R_{\text{eq}}} = 0.31\,\text{rad}$, we show in Fig. 3.8 the values of $\tilde{E}_{\text{eq}}$, $\tilde{k}_{24}$, and qR as a function of $\tilde{\gamma}$. These are determined by calculating the $\psi_{R_{\text{eq}}} = 0.31\,\text{rad}$ contour line (i.e. $\tilde{k}_{24}$ *vs* $\tilde{\gamma}$ line) in Fig. 3.5, and mapping this relationship onto Figs. 3.4 and 3.6, to determine $\tilde{E}_{\text{eq}}$ and $q\tilde{R}_{\text{eq}}$, respectively. Restricting ourselves to thermodynamically stable parameterizations, with $\tilde{E}_{\text{eq}} < 0$, from Fig. 3.8 we expect that $\tilde{\gamma} \in [0.1, 0.2]$, $\tilde{k}_{24} \in [0.6, 1.25]$, and $qR_{\text{eq}} \in [0.2, 0.4]$.

---

[2] This is in contrast to lyotropic systems, in which free energy and entropy drive phase transitions.

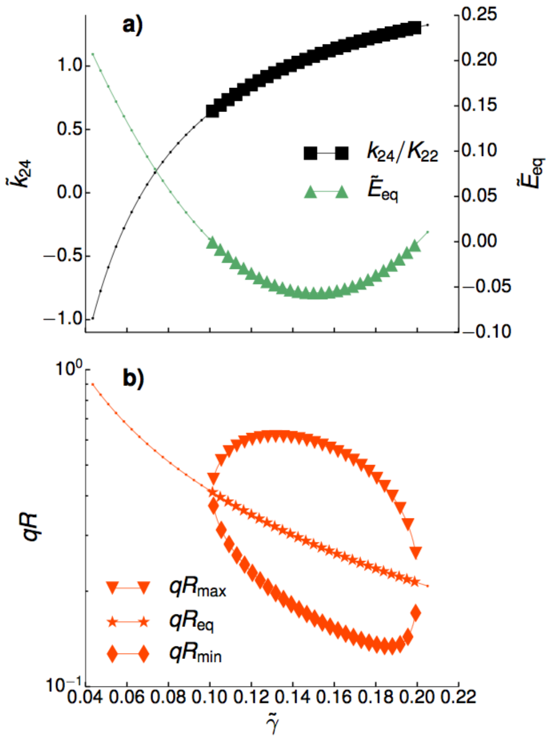

Figure 3.8: Radial fibril structure and stability of corneal fibrils. Surface twist is restricted to $\psi_{R_{\rm eq}} = 0.31$ (i.e. along the 0.31 contour in Fig. 3.5) — the experimentally measured surface twist of corneal fibrils. (a) Reduced saddle-splay $\tilde{k}_{24}$ *vs* reduced surface-tension $\tilde{\gamma}$ is indicated in black squares and dots, while reduced minimum energy-density $\tilde{E}_{\rm eq}$ *vs* $\tilde{\gamma}$ is indicated by green triangles and dots. Dots indicate where fibrils are only meta-stable with respect to the cholesteric phase, and shapes indicate where fibrils are stable with respect to the cholesteric phase, $\tilde{E}_{\rm eq} < 0$. (b) The dimensionless fibril radius qR *vs* $\tilde{\gamma}$. The equilibrium radius that minimizes $\tilde{E}$, qR_{eq}, is indicated by stars (when $\tilde{E}_{\rm eq} < 0$) and dots (when $E_{\rm eq} \geq 0$). The minimum and maximum fibril radii that are stable with respect to the cholesteric (i.e. qR values such that $\tilde{E}(qR_{\rm min}) = 0$, $\tilde{E}(qR_{\rm max}) = 0$ and $qR_{\rm min} < qR_{\rm eq} < qR_{\rm max}$), are indicated by diamonds and triangles, respectively.

Human corneal fibrils have a typical diameter of $30 - 35$nm [136, 103, 55]. We consider a radius of $R \simeq 0.015\,\mu$m for convenience. This then implies an approximate range of expected chiral wavenumber $q \in [13, 27]\mu\text{m}^{-1}$. This range abuts the expected range from Sec. 3.4.2 at larger $qR_{\rm eq}$, when $\tilde{\gamma} \simeq 0.1$ and $\tilde{k}_{24} \simeq 0.75$ – this is near point "3" of Fig. 3.7.

Using $\tilde{\gamma} \simeq 0.1$ and $q \simeq 13\mu\text{m}^{-1}$, our expected range of $K_{22} \in [0.6, 6]\,$pN from Sec. 3.4.2 implies $\gamma \in [1.6, 16]\,\text{pN}\,\mu\text{m}^{-1}$. This is entirely within the expected range of

$\gamma \in [0.5, 25]\,\mathrm{pN\,\mu m^{-1}}$. As mentioned, k_{24}, is not well constrained — but nevertheless $\tilde{k}_{24} \simeq 0.75$ is close to the expected scale [133, 134].

Corneal fibrils are very close to the stability boundary between fibrils and the cholesteric phase due to their large surface twists. This implies that only a very narrow range of fibril radii are stable with respect to the cholesteric phase, with $\tilde{E} < 0$. In Fig. 3.8 (b), in addition to qR_{eq}, we indicate the minimum and maximum values for stable fibrils, qR_{min} and qR_{max}, respectively. For a given $\tilde{\gamma}$ and $\tilde{k}_{24}$, qR_{min} and qR_{max} are defined such that $\tilde{E}(qR_{\mathrm{min}}) = 0$, $\tilde{E}(qR_{\mathrm{max}}) = 0$ and $qR_{\mathrm{min}} < qR_{\mathrm{eq}} < qR_{\mathrm{max}}$. We see that precisely at $\tilde{\gamma} \simeq 0.1$, there is only a very narrow range of stable fibril radii available for corneal fibrils. Furthermore, a narrow range of corneal fibril radii is observed [136] and is *required* for corneal transparency [137, 59].

In Fig. 3.10 of Supplemental section 3.6, we examine different values of $\tilde{K}_{33}$ to determine whether the correlation between narrow stability and large surface twist is sensitive to our parameter choices. We find that this behaviour persists in a wide range of $\tilde{K}_{33} \in [10, 40]$, for $\psi_{R_{\mathrm{eq}}} \simeq 0.31\,\mathrm{rad}$. From this, we hypothesize that the large surface twist of corneal fibrils may be a result of being at the stability boundary, which in turn is required to narrow the range of accessible fibril radii. Cross-linking after fibrillogenesis could then mechanically stabilize corneal fibrils.

Other “helicoidal” or “C”-type[73] collagen fibrils also exhibit large surface twists with $\psi_{R_{\mathrm{eq}}} \simeq 0.3\,\mathrm{rad}$ and a narrow unimodal distribution of fibril radii [48, 73, 80, 74]. These helicoidal fibrils are found in e.g. skin, interstitial stroma, and nerve and tendon sheaths. They have a slightly shorter D-period, consistent with the projective coupling hypothesis [48, 121]. Despite their similarity of surface twist, in each tissue helicoidal fibrils exhibit a different unimodal radius – from $0.015\,\mathrm{\mu m}$ to $0.050\,\mathrm{\mu m}$ [48]. Larger radii than seen in corneal fibrils could be accommodated in our model by smaller q, or by different points along the stability boundary of Fig. 3.5.

Interestingly, some originally helicoidal fibrils from skin that have been disassociated and reconstituted are no longer helicoidal [138, 50] — though see ref. [139]. This implies that fibrillogenesis conditions are important in determining their reduced parameterization; parameters are not simply determined by the molecular type, but also by the environment. While our approach can constrain reduced parameterization with observations of fibril surface twist and radius, a direct assessment of elastic constants within the context of individual fibrils would require different approaches.

Tendon fibrils

For the low surface twist of tendon collagen fibrils, with $\psi_{R_{\text{eq}}} \simeq 0.1\,\text{rad}$, we show in Fig. 3.9 the values of $\tilde{E}_{\text{eq}}$, $\tilde{k}_{24}$, and qR as a function of $\tilde{\gamma}$. These correspond to mapping the $\psi_{R_{\text{eq}}} = 0.1\,\text{rad}$ contour line from Fig. 3.5 to Figs. 3.4 and 3.6, respectively. Restricting ourselves to thermodynamically stable parameterizations, with $\tilde{E}_{\text{eq}} < 0$, we expect that $\tilde{\gamma} \in [0, 0.07]$, $\tilde{k}_{24} \in [-1, 1.1]$, and $qR_{\text{eq}} \in [0.01, 0.2]$. While most of these ranges are larger than those of corneal fibrils, the values of qR_{eq} for tendon fibrils are significantly smaller.

Tendon fibrils *in vivo* have a large range of radii, from 0.02 µm to 0.2 µm [68, 140], and the distribution varies with age and tissue type. Significantly, fibrils within the same tissue exhibit a broad range of radii. Nevertheless, the average fibril tendon radius $R = 0.08$µm from older mouse tails [140] is much larger than typical corneal fibrils. This implies expected values of $q \in [0.13, 2.5]\mu\text{m}^{-1}$. These chiral wavenumbers are significantly smaller than for corneal fibrils, but are entirely within the expected range from Sec. 3.4.2. Combining possible ranges, we then expect the surface tension $\gamma \in [0, 1]\text{pN}\,\mu\text{m}^{-1}$. This is in the lower end of, but largely within, the range expected from Sec. 3.4.2.

However, to have a broad distribution of equilibrium tendon fibril radii within the same section of tissue [68, 140] would imply a broad range of reduced parameters, and hence of conditions during fibrillogenesis. Tendon fibrils in particular are almost entirely comprised of type-I collagen, and so this variation cannot be attributed to variations of composition. Rather, we believe that non-equilibrium processes are involved in the determination of tendon fibril radii — as proposed by Kalson et al. [140].

Fibrils with a small surface twist are expected to be quite stable with respect to the cholesteric phase (see Fig. 3.4). This implies that fibrils at a broad range of different radii around the equilibrium will also be stable with respect to the cholesteric, as shown by the difference in magnitude of qR_{min} and qR_{max} in Fig. 3.9b. We note that there is at least a 100-fold range of stable radii available between R_{min} and R_{max}, with a narrower 5-fold range range between R_{eq} and R_{min}. The observed 10-fold range

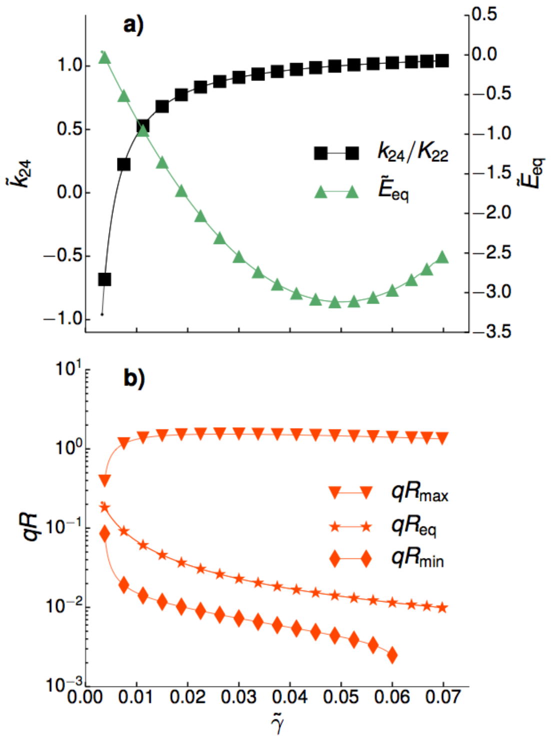

Figure 3.9: Radial fibril structure and stability of tendon fibrils. Surface twist is restricted to $\psi_{R_{\mathrm{eq}}} = 0.1$ (i.e. along the 0.1 contour in Fig. 3.5) — consistent with experimentally observed surface twist of tendon fibrils. (a)) Reduced saddle-splay $\tilde{k}_{24}$ *vs* reduced surface-tension $\tilde{\gamma}$ is indicated in black squares and dots, while reduced minimum energy-density $\tilde{E}_{\mathrm{eq}}$ *vs* $\tilde{\gamma}$ is indicated by green triangles and dots. Dots indicate where fibrils are only meta-stable with respect to the cholesteric phase, and shapes indicate where fibrils are stable with respect to the cholesteric, $\tilde{E}_{\mathrm{eq}} < 0$. (b)) The dimensionless fibril radius qR *vs* $\tilde{\gamma}$. The equilibrium radius that minimizes $\tilde{E}$, qR_{eq}, is indicated by stars (when $\tilde{E}_{\mathrm{eq}} < 0$) and dots (when $E_{\mathrm{eq}} \geq 0$). The minimum and maximum fibril radii that are stable with respect to the cholesteric (i.e. qR values such that $\tilde{E}(qR_{\mathrm{min}}) = 0$, $\tilde{E}(qR_{\mathrm{max}}) = 0$ and $qR_{\mathrm{min}} < qR_{\mathrm{eq}} < qR_{\mathrm{max}}$), are indicated by diamonds and triangles, respectively. Due to the divergent behaviour of the double-twist for $\tilde{k}_{24} \gtrsim 1$ (see Fig. 3.4), only a small range of $0 < \tilde{\gamma} < 0.07$ is accessible for $\psi_{R_{\mathrm{eq}}} = 0.1\,\mathrm{rad}$.

of tendon fibril radii fits within the larger range of stable fibrils with respect to the cholesteric.

Our hypothesis then is that non-equilibrium cross-linking works to stabilize fibril radii that are away from R_{eq}, but only have the opportunity to act on fibrils that are stable with respect to the cholesteric (between R_{min} and R_{max}). Essentially we propose that fibrillogenesis only takes place when fibrils are thermodynamically stable, while cross-linking can freeze (and so prevent) the subsequent slow relaxation of fibril radii towards the minimal energy radius R_{eq}. We note that this thermodynamic stability may also be of use during remodelling after damage for these load-bearing fibrils [141].

3.5 Summary and Future Work

We model collagen fibrils with a double-twist director field of molecular tilt, and identify where a fibril phase is more stable than a cholesteric phase. The stability, dimensionless radius qR, and surface twist of the fibrils $\psi(qR)$ are controlled by two dimensionless parameters, the ratio of surface tension to the chiral strength $(\gamma/K_{22}q)$ and the ratio of saddle-splay to twist elastic constants (k_{24}/K_{22}). The fibril phase is the equilibrium state with respect to the cholesteric phase only when the surface tension is small compared to the chiral strength. Within this limit, the fibril phase can access a wide range of equilibrium configurations $(R_{\mathrm{eq}}, \psi(R_{\mathrm{eq}}))$. Current experimental observations are consistent with our equilibrium picture, and indicate that controlled equilibrium polymorphism of collagen fibrils may be significant biologically. We suggest that corneal collagen fibrils are formed close to the fibril-cholesteric stability boundary, with large surface twists, in order to achieve a narrow range of fibril radii and to ensure corneal transparency. Conversely, tendon collagen fibrils are formed away from the stability boundary, with small surface twists, but non-equilibrium effects are needed to explain the polydispersity of tendon fibril radii within individual tissues. A key conclusion is that experimental characterization of a collagen fibril population should always include both radius and surface twist measurements.

Though this model is able to capture the radial structure of collagen fibrils, it does not explicitly consider the axial D-band structure, which is a dominant feature of fibrils. The Hodge-Petruska model of D-band (discussed in Chapter 2) [11] indicates that the D-band arises from molecules within fibrils having a specific arrangement parallel to the cylindrical fibril axis, which suggests that orientation and D-band cannot be considered separately, but must instead be coupled together. This coupling has been proposed as the reason why larger twist fibrils have shorter D-band periodicity [48, 121]. In this chapter, we have assumed that the D-band is completely malleable, and follows the double-twist with no cost to the overall free energy of the fibril. However, it is not obvious that this is true; perhaps interaction of D-band

and molecular twist is important in the self-assembly of collagen fibrils. To provide insight into this question, and to fully understand the relationship between D-band and twist, a more general model of fibril structure which couples the two together must be constructed. In the next chapter of this thesis, we address these speculations with a new model of collagen fibril structure which includes D-band.

3.6 Supplemental: Other K_{33} values

In Fig. 3.10 we show the surface twist *vs* reduced parameters $\tilde{k}_{24}$ and $\tilde{\gamma}$ for a range of $\tilde{K}_{33} \equiv K_{33}/K_{22}$ values: 10, 20, 30, and 40 for subfigures a)-d) respectively. We note that $\tilde{K}_{33} = 30$ corresponds to Fig. 3.5 but is included for ease of reference.

3.7 Supplemental: Dimensional reduction

The free energy per unit volume of fibril is

$$\begin{aligned} E =& \frac{1}{R^2}\int_0^R dr \left[K_{22} r \left(q - \psi' - \frac{\sin 2\psi}{2r} \right)^2 + K_{33} \frac{\sin^4 \psi}{r} \right] \\ & - (K_{22} + k_{24}) \sin^2 \psi(R) + \frac{2\gamma}{R}. \end{aligned} \tag{3.8}$$

Multiplying eqn. 3.8 by $1/(K_{22}q^2)$ gives the dimensionless free energy per unit volume of fibril,

$$\begin{aligned} \tilde{E} =& \frac{1}{\tilde{R}^2}\int_0^{\tilde{R}} d\tilde{r} \left[\tilde{r} \left(1 - \tilde{\psi}' - \frac{\sin 2\tilde{\psi}}{2\tilde{r}} \right)^2 + \tilde{K}_{33} \frac{\sin^4 \tilde{\psi}}{\tilde{r}} \right] \\ & - (1 + \tilde{k}_{24}) \sin^2 \tilde{\psi}(\tilde{R}) + \frac{2\tilde{\gamma}}{\tilde{R}}, \end{aligned} \tag{3.9}$$

where we have defined the dimensionless quantities $\tilde{K}_{33} = K_{33}/K_{22}$, $\tilde{k}_{24} = k_{24}/K_{22}$, $\tilde{\gamma} = \gamma/(K_{22}q)$, $\tilde{r} = qr$, $\tilde{R} = qR$, $\tilde{\psi}(\tilde{r}) = \psi(r)$, $\tilde{E} = E/(K_{22}q^2)$.

3.8 Supplemental: Power Series Solution

We first assume that a convergent power-series expansion of $\psi(r)$ in powers of the radius r exists. We then analytically continue this solution to negative r, in order to

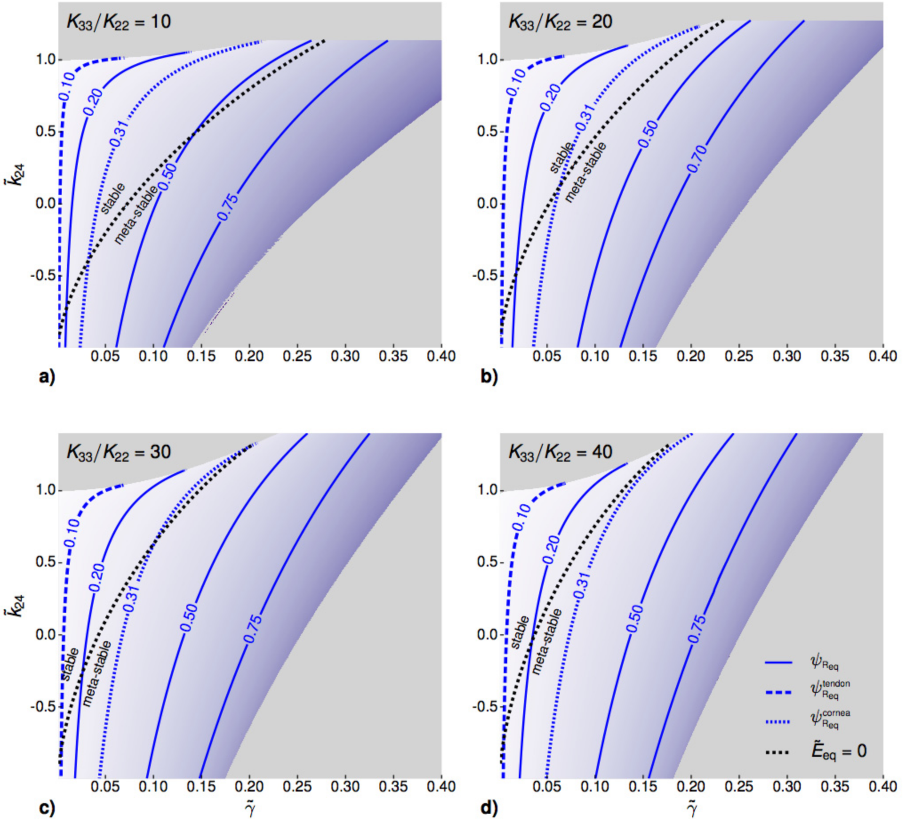

Figure 3.10: Calculated fibril surface twist $\psi(qR_{\rm eq}) \equiv \psi_{R_{\rm eq}}$ for different values of K_{33}/K_{22}. As K_{33}/K_{22} increases, the surface twist values tend to decrease in size for a given $\tilde{\gamma}$ and $\tilde{k}_{24}$. In (a)) the double-twist model predicts the existence of fibrils with very large surface twist, $\simeq 0.52$ rad, which are stable with respect to the cholesteric phase. The surface twist values shown in (b) and (c) predict a wide range of equilibrium surface twist values (dependent on parameter values), consistent with experimental observations. The 0.1 rad contour, labelled $\psi_{R_{\rm eq}}^{\rm tendon}$, is a typical surface twist value of in vivo tendon fibrils. Similarly, the 0.31 rad contour, labelled $\psi_{R_{\rm eq}}^{\rm cornea}$, is a typical surface twist value of in vivo corneal fibrils. The surface twist of both fibril types is captured for each K_{33}/K_{22} value shown, but as seen in (d), as $K_{33}/K_{22} \gtrsim 40$, the 0.31 rad (cornea) surface twist line transitions completely into the metastable regime. The gray areas in (a)-(d) correspond to parameter space regions for which no stable or meta-stable double-twist configurations are found. Note that $\tilde{k}_{24} \equiv k_{24}/K_{22}$, $\tilde{\gamma} \equiv \gamma/(K_{22}q)$, and $\tilde{K}_{33} \equiv K_{33}/K_{22}$.

simply note that if $\psi(r)$ is a solution to eqns. 3.6 then so is $-\psi(-r)$ — i.e. ψ is an odd function and will only have odd terms in its power-series expansion. (This result is independently verified by the numerically relaxed solutions.)

To simplify our derivation we will use the dimensionless formulation from the previous Supplemental section 3.7 but will drop the tildes. Then the power series solution for $\psi(r)$ with only odd terms is of the form

$$\psi(r) = \sum_{n=0}^{\infty} a_n r^{2n+1}, \tag{3.10}$$

and satisfies

$$(r\psi')' = 1 + \frac{K_{33}}{2r}\sin 2\psi + \frac{(1-K_{33})}{4r}\sin 4\psi - \cos 2\psi, \tag{3.11}$$

where the trigonometric identities $\sin^2 x = 1/2(1-\cos 2x)$ and $\sin 2x = 2\sin x\cos x$ have been used. Taylor expanding the trigonometric functions yields

$$\begin{aligned}(r\psi')' =& 1 + \frac{K_{33}}{2r}\sum_{n=0}^{\infty}\frac{(-1)^n 2^{2n+1}}{(2n+1)!}(\psi)^{2n+1} \\ &+ \frac{(1-K_{33})}{4r}\sum_{n=0}^{\infty}\frac{(-1)^n 4^{2n+1}}{(2n+1)!}(\psi)^{2n+1} \\ &- \left(1 + \sum_{l=1}^{\infty}\frac{(-1)^l 2^{2l}}{(2l)!}(\psi)^{2l}\right).\end{aligned} \tag{3.12}$$

The general form of ψ^n in terms of a_k is

$$\begin{aligned}\psi^n &= \left(\sum_{k=0}^{\infty} a_k r^{2k+1}\right)^n \\ &= r^n \sum_{k=0}^{\infty}\left(\sum_{j_1+j_2+\cdots+j_n=k} a_{j_1}a_{j_2}\cdots a_{j_n}\right) r^{2k},\end{aligned} \tag{3.13}$$

where $j_1, j_2, \cdots, j_n \geq 0$ are integer indices and we have used the Cauchy product

$$\sum_{n=0}^{\infty} a_n x^n \sum_{m=0}^{\infty} b_m x^m = \sum_{k=0}^{\infty}\sum_{l=0}^{k} a_l b_{k-l} x^k. \tag{3.14}$$

Using this we obtain

$$\begin{aligned}\sum_{n=0}^{\infty}(2n+1)^2 a_n r^{2n} =& \sum_{n=0}^{\infty} c_n r^{2n} \sum_{k=0}^{\infty} p_{2n+1,k} r^{2k} \\ &+ \sum_{l=1}^{\infty} d_l r^{2l} \sum_{k=0}^{\infty} p_{2n,k} r^{2k},\end{aligned} \tag{3.15}$$

where we have defined

$$c_n = \frac{(-1)^n 2^{2n}}{(2n+1)!}\left[K_{33} + 2^{2n}(1-K_{33})\right], \tag{3.16}$$

$$d_l = \frac{(-1)^l 2^{2l}}{(2l)!}, \tag{3.17}$$

$$p_{n,k} = \sum_{j_1+j_2+\cdots+j_n=k} a_{j_1} a_{j_2} \cdots a_{j_n}. \tag{3.18}$$

Using eqn. 3.14, we re-write eqn. 3.15

$$\begin{aligned}\sum_{n=0}^{\infty}(2n+1)^2 a_n r^{2n} =& \sum_{n=0}^{\infty}\sum_{j=0}^{n} c_{n-j}\, p_{2(n-j)+1,j} r^{2n} \\ &+ \sum_{n=0}^{\infty}\sum_{k=0}^{n-1} d_{n-k}\, p_{2(n-k),k} r^{2n}.\end{aligned} \tag{3.19}$$

We can determine each a_n recursively from the eqn. 3.19. We find that $a_0 = \psi_0'$ is arbitrary. For $n \geq 1$, eqn. 3.19 can be rearranged to give

$$a_n = \frac{\sum_{j=0}^{\infty} c_{n-j}\, p_{2(n-j)+1,j} + \sum_{k=0}^{n-1} d_{n-k}\, p_{2(n-k),k}}{[(2n+1)^2 - 1]}, \quad n \geq 1 \tag{3.20}$$

Since $p_{2(n-j)+1,j},\ p_{2(n-k),k}$ depend on all lower coefficients $a_0, \cdots, a_{n-1}$, calculating $\psi(r)$ to high order in r becomes increasingly difficult and is impractical for a broad range of parameters. Nevertheless, we can use the leading cubic term as a starting point for our numerical relaxation approach:

$$\psi(r) = \psi_0' r + \frac{(3K_{33}-4)\psi_0'^3 - 3\psi_0'^2}{12} r^3 + \mathcal{O}(r^5) \tag{3.21}$$

Chapter 4

Phase field collagen fibrils: Coupling D-band and twist

This chapter is closely based on a first draft for a scientific paper that is being developed, and so is fairly self-contained in its discussion of collagen fibril structure and how this work may fit into the literature. The work presented in this chapter is my own, with input on what figures and discussion might be interesting coming from my supervisors. A new model of collagen fibril structure is presented which builds on the model of the previous chapter. This new model not only considers the orientation of molecules within the collagen fibril, but also their periodic D-band striations along the cylindrical axis of the fibril, and so fills a major void in previous quantitative models of radial structure which neglected this important feature.

4.1 Introduction

Collagen fibrils have an important place in the study and characterization of biological materials, given their ubiquity in mammals as well as other vertebrates. These fibrils are long and cylindrical in shape, with a wide range of possible radii, $R \in 10-200$nm depending on anatomical location in vivo[65, 48] or experimental conditions in vitro [70, 142, 19]. The lengths of these fibrils, L, are much larger than their radii, and so a fibril can be locally thought of as a very long cylinder. The cylindrical structure of fibrils arises from the dense packing of 300nm long, 1.5nm wide chiral tropocollagen molecules [16], which are themselves composed of smaller peptides of a similar helical nature and aspect ratio [10]. These molecules are aligned somewhat parallel with the long axis of the fibril, though the precise orientation within the fibril is dependent on anatomical location [72, 73].

The most well characterized feature of collagen fibrils is the presence of periodic banding along the long axis of the fibril, known as the fibril D-band. The mechanism of this D-banding structure is due to specific inter-molecular interactions which force adjacent molecules to be offset vertically from each other, driving a staggering pattern (i.e. Hodge-Petruska model [11]). This staggering drives alternating regions of high density and low density along the cylindrical axis. Bulk (e.g. scattering) measurements of D-band suggest that a single D-band period is found throughout the radial cross-section (i.e. D-band is constant in a circular cross-section of fibril) [143, 49]. In vivo, the period of the D-band remains fairly consistent between fibrils found in different regions of the body, with period length between $64-67$nm [96]. In contrast, the D-band period of in vitro fibrils have much more variability [144], and can even be manipulated by the ratio of collagen I to collagen III molecules in the formation of the fibrils [19]. When pulled on at both the single fibril or tissue (fibril bundle) scale, the D-band period typically increases with applied strain [6] and can serve as measure of how much strain is being felt by the individual fibrils. In contrast to this axial D-banding, the spacing of molecules in a circular (radial) cross-section of fibrils is much more disordered [145].

Another experimentally accessible/interesting feature of collagen fibrils, the average orientation of molecules on the surface of the fibril, has received considerable attention in the literature over the past two decades [80, 73, 3, 79, 4]. We refer to this feature as the fibril surface twist, $\psi(R)$. This surface twist generally forms a small angle with respect to the long axis of the fibril, and likely arises from the chiral nature of the fibril's constituent tropocollagen molecules [3, 4, 82]. Although first measured many years ago using freeze-fractured electron microscopy [72, 74], surface twist has been widely neglected in the collagen literature, perhaps due to the difficulty of measuring such a small scale feature compared to something as prevalent as the fibril D-band. However, large differences in surface twist are observed in different anatomical locations [48], and may be associated with radial size control of fibrils [4]. The largest surface twist values are measured in fibrils extracted from cornea, and the smallest surface twist values are measured in fibrils extracted from tendon. These two tissues then serve as the two extreme examples of this feature. Corneal fibrils have surface twist around 17°. Tendon fibrils, in contrast, have surface twist closer to 5°. To the author's knowledge, no measurements of surface twist have been carried out on in vitro assembled fibrils. However, it has been inferred from D-band measurements that the surface twist of in vitro fibrils is relatively small [138, 48].

Two descriptive models have been put forth in attempt to understand observed surface twist: the constant pitch model and the constant twist model [48]. Although not explicitly said in either case, both models consider the local, average orientation of tropocollagen molecules constrained to be in a double-twist orientation (see Figure 1 of [4]). Therefore, the tropocollagen orientation can generally be described by a local twist field, $\psi(r)$, which depends only on the distance r from the fibril centre. The surface twist $\psi(R)$ is the twist at the fibril radius, $r = R$.

The constant pitch model, first proposed by Raspanti [80], is that of a constant, non-zero twist gradient starting with zero twist at the fibril centre, i.e. the local twist is linear in r, $\psi(r) = \psi_0' r$. This model draws inspiration from chiral nematic and blue phases having a constant pitch [81], and is attractive from a molecular chirality perspective as it allows the tropocollagen molecules to twist with respect to each other. Recent theoretical work [3, 4] using a generalization of this constant pitch structure has indicated that a slightly non-linear (but monotonically increasing) twist angle is thermodynamically stable, i.e. has a minimal free energy. However, this twisted structure appears to be incompatible with the claim that D-band is constant throughout the fibril phase. Using the projective coupling hypothesis put forth in the literature (see e.g. page 1671 of [48]), the D-band $d \propto \cos\psi(r)$. Therefore, any local gradients in twist will give rise to local gradients in D-band. It is possible, however, that the molecules themselves are axially compressed near $r = 0$ and stretched near $r = R$ to maintain a constant D-band. As such, the constant pitch model assumes an axially soft D-band.

The second model of molecular orientation in the fibril was first put forth by Galloway [77], in which the twist angle of molecules remains constant throughout a circular cross-section of the fibrils, i.e. $\psi(r) = \psi_0$. This constant twist model has the attractive feature of accommodating the D-band, since the twist angle is constant and so $d \propto \cos\psi_0 =$ constant. However, a new incompatibility arises in this structure, as this constant twist phase suppresses the ability of the (chiral) tropocollagen molecules to twist with respect to each other. This suppression of twist can be explained through assuming the strength of the inter-molecular potential which give rise to the D-band are much larger than the strength of the chiral interaction (i.e. it assumes a soft twist field). However, no comparison of these two energy scales (D-band *vs* twist) has been examined in the literature.

These two descriptive models of collagen fibrils have motivated experimentalists to measure the orientation of molecules within the fibril as well as on the fibril surface, with the hope of understanding this frustration between the interactions driving the D-band and those driving the molecular twist. Holmes et al. [55] used automated electron tomography to measure the local orientation of the collagen molecules within the fibril. They concluded from their experiments that the local molecular orientation was consistent with a constant twist angle of roughly 15°, and so supported the constant twist angle model proposed by Galloway. In the discussion, we address this claim further.

Although descriptive models are helpful in conceptual understanding of a physical system, it would be ideal to have a model which is not only consistent with experimental observation, but also arises from more fundamental physical concepts. In recent years, quantitative model(s) of collagen fibrils have been developed using the tools of theoretical and computational physics. Since the hierarchical scale of collagen fibrils makes it difficult to build a fully molecular model of the fibril structure, approximations must be used to simplify the phenomenology while still capturing the important physics. Previous theoretical work on continuum modelling of collagen fibrils employed techniques of liquid crystal physics to examine the local average configuration of tropocollagen in a fibril, and was able to capture the upper limit of surface twist observed in cornea fibrils [4] while also proposing a stability argument for why large surface twist might lead to the more sharply peaked radial distribution of cornea fibrils [146, 59, 147]. This work was consistent with the constant pitch, $\psi(r) = \psi_0' r$, descriptive model of fibrils, but ignored D-band effects, and so was unable to provide a definitive answer on the underlying configuration of tropocollagen, which should be influenced by both the molecular chirality and D-band interactions.

Computational approaches include atomistic molecular dynamics simulations of a periodic lattice of tropocollagen molecules [86]. This approach is quite successful in reproducing the experimental mechanical response of collagen fibrils. However, due to the large computational cost of full scale atomic simulations of mesoscopic structures, only short time scales (a few nanoseconds) are accessible. To speed mechanical convergence, the underlying structure of the tropocollagen molecules is put in by hand using x-ray crystallographic data [58] of a collagen microfibril (consisting of five tropocollagen molecules). While molecular dynamics models can examine the change in molecular configuration of tropocollagen while applying strain to the microfibril, they are limited by the large computational time typical of atomistic calculations, and so it are not able to relax to find novel equilibrium configurations (e.g. double twist and/or D-band).

To the authors' knowledge, there is no theoretical model of collagen fibril formation which considers both the D-band structure and the molecular twist of the collagen molecules. In this chapter, we construct such a hybrid phase field crystal liquid crystalline model of collagen fibrils which couples the D-band and molecular twist. We aim to answer two main questions:

1. How do tropocollagen molecules pack within the fibril to allow for a robust D-band to be seen, while at the same time tilting with respect to the fibril axis?
2. How do the mechanical properties couple to the molecular orientation of tropocollagen molecules when the fibril is pulled at one end?

In section 4.2, we develop our theoretical model of collagen fibrils using techniques from liquid crystal and phase field crystal theories and reduce our model to dimensionless form. In section 4.2.2, we touch on the numerical methods we use to solve the fibril model we have constructed, and also develop a core - shelf - surface approximation that can be used to simplify our calculations. In section 4.3, we explore the predictions of our model using our parameter estimates from section 4.2. In section 4.4, we discuss how our results change the current understanding of the interplay between the D-band of the fibril and the molecular twist within the fibrils, and how tuning the strength of the coupling between the D-band and the twisting of the molecules in the fibril can give rise to two different fibril phases. Finally, we conclude in section 4.5.

4.2 Phase field collagen fibril model

4.2.1 Construction of the free energy

The free energy per unit volume of our model is the sum of four terms,

$$\tilde{E}_{\text{tot}} = \tilde{E}_{\text{Frank}} + \tilde{E}_{\text{pfc}} + \tilde{E}_{\text{dw}} + \tilde{E}_{\text{surf}}. \tag{4.1}$$

We provide here a brief derivation of these four terms. A more comprehensive derivation along with our model assumptions is included in Appendix A of this thesis.

The first term is just the volume averaged Frank free energy density within a segment of fibril of length $\tilde{L}$ and radius $\tilde{R}$ with a double-twist director field $\boldsymbol{n} = \sin\psi(\tilde{r})\hat{\phi} + \cos\psi(\tilde{r})\hat{z}$ (see e.g. the previous chapter for details of double-twist structure),

$$\begin{aligned}\tilde{E}_{\text{Frank}} =& \frac{1}{\pi\tilde{R}^2\tilde{L}}\int_{\tilde{V}}\Bigg(\frac{1}{2}\tilde{K}_{11}(\tilde{\tilde{\nabla}}\cdot\boldsymbol{n})^2 + \frac{1}{2}\tilde{K}_{22}(\boldsymbol{n}\cdot\tilde{\nabla}\times\boldsymbol{n}+q)^2 \\ &+ \frac{1}{2}\tilde{K}_{33}(\boldsymbol{n}\times(\tilde{\nabla}\times\boldsymbol{n}))^2 + \tilde{k}_{13}\tilde{\nabla}\cdot(\tilde{\nabla}\cdot\boldsymbol{n})\boldsymbol{n} \\ &- \frac{1}{2}(\tilde{K}_{22}+\tilde{k}_{24})\tilde{\nabla}\cdot(\boldsymbol{n}\times(\tilde{\nabla}\times\boldsymbol{n})+\boldsymbol{n}(\tilde{\nabla}\cdot\boldsymbol{n}))\Bigg)d^3\tilde{x} \\ =& \frac{2}{\tilde{R}^2}\int_0^{\tilde{R}}\tilde{r}d\tilde{r}\Bigg(\frac{1}{2}\tilde{K}_{22}\bigg(\tilde{q}-\frac{\partial\psi}{\partial\tilde{r}}-\frac{\sin 2\psi}{2\tilde{r}}\bigg)^2 \\ &+ \frac{1}{2}\tilde{K}_{33}\frac{\sin^4\psi}{\tilde{r}^2}\Bigg) - (\tilde{K}_{22}+\tilde{k}_{24})\frac{\sin^2\psi(\tilde{R})}{\tilde{R}^2}. \end{aligned} \tag{4.2}$$

Note that the $\tilde{K}_{11}$ and $\tilde{k}_{13}$ terms vanish as $\tilde{\nabla}\cdot\boldsymbol{n} = 0$ for a double-twist configuration.

The second term, $\tilde{E}_{\text{pfc}}$, is motivated by phase-field crystal theory [5], being the simplest coarse-grained free energy expansion to allow for periodic structure in density perturbations. In this theory, a power series expansion of the free energy in terms of gradients in an order parameter (typically density) is written out, as is usual in phenomenological (Landau-like) field theories. The distinction of phase-field theories are that they consider gradient terms that are higher than second order (at least to fourth order in order to ensure stability). This enables simulation of fields which have some preferred periodic order parameter gradient, as is the case for e.g. local density within crystals, and allows one to extract mechanical and structural properties of crystalline materials by applying a strain (through a lengthening of the preferred crystalline period) without being inhibited by the short time scales of atomistic simulations.

In our case, we know that collagen fibrils have a one-dimensional periodicity along the cylindrical fibril axis (the D-band). To model this, we assume that locally the molecules would like to pack along their long axis with period $\tilde{d}_{\parallel} = 67\,\text{nm}$. This local orientation is not in the same direction as the fibril axis due to the local twist, $\psi(r)$, which gives rise to a coupling between the periodicity of the D-band and $\psi(r)$. With this coupling, the realized period of the D-band is $2\pi/\tilde{\eta} \neq \tilde{d}_{\parallel}$. We can write the phase-field crystal term as

$$\begin{aligned}\tilde{E}_{\text{pfc}} &= \frac{1}{\pi \tilde{R}^2 \frac{2\pi}{\tilde{\eta}}} \int_{\tilde{V}} \tilde{\phi}(\tilde{\boldsymbol{r}}) \left(\frac{4\pi^2}{\tilde{d}_{\parallel}^2} + \tilde{\nabla}_{\parallel}^2 \right)^2 \tilde{\phi}(\tilde{\boldsymbol{r}}) d^3\tilde{x} + \tilde{\mathcal{F}}_{surf} \\ &= \frac{\tilde{\Lambda}\tilde{\delta}^2}{2\tilde{R}^2} \int_0^{\tilde{R}} \tilde{r} d\tilde{r} \left(\frac{4\pi^2}{\tilde{d}_{\parallel}^2} - \tilde{\eta}^2 \cos^2 \psi(\tilde{r}) \right)^2 \end{aligned} \tag{4.3}$$

where integration is taken over one period of the fibril's length, $2\pi/\eta$[1]. In going from the first line to the second line in the above equation, we have assumed that D-band variations ϕ are only occurring along the fibril (z) axis, consistent with experiment. This assumption also allows us to write $\tilde{\nabla}_{\parallel} = \cos \psi(\tilde{r}) \partial/\partial \tilde{z}$, and gives $\tilde{\mathcal{F}}_{\text{surf}} = 0$ (see section A.2 in the appendix for a detailed proof of this). Given the dominance of the lowest mode modulation in experimental measurements of axial structure, we have further approximated the D-band modulations by $\tilde{\phi}(\tilde{z}) \sim \tilde{\delta} \cos(\tilde{\eta}\tilde{z})$. This simplification allows us to integrate $\tilde{z}$ over the length of the D-band.

The third term, $\tilde{E}_{\text{dw}}$, is the standard double-well term which arises from expanding the free energy in terms of an order parameter field, which for us is the D-band modulation $\tilde{\phi}(\tilde{z}) = \tilde{\delta} \cos(\tilde{\eta}\tilde{z})$. We can write this as

$$\begin{aligned}\tilde{E}_{\text{dw}} &= \frac{1}{\pi \tilde{R}^2 \frac{2\pi}{\tilde{\eta}}} \tilde{\omega} \int_{\tilde{V}} \tilde{\phi}^2 (\tilde{\phi}^2 - \tilde{\delta}_0^2) d^3\tilde{x} \\ &= \frac{\tilde{\omega}\tilde{\delta}^2}{2} \left(\frac{3}{4}\tilde{\delta}^2 - \tilde{\delta}_0^2 \right), \end{aligned} \tag{4.4}$$

where $\tilde{\omega}$ and $\tilde{\delta}_0$ are related to the second and fourth order terms in the Landau-like power series expansion, respectively. If $\tilde{\delta}_0^2 < 0$, then $\tilde{E}_{\text{dw}} > 0$ and any energy minimization scheme would drive $\tilde{\delta} \to 0$ (since eqn 4.3 is positive definite with respect to δ as well). We are only interested in the case in which $\tilde{\delta} \neq 0$ as the no D-band modulation case has been described in previous work, so we set $\tilde{\delta}_0^2 > 0$ in this chapter.

The final term of eqn 4.1 is the surface cost of creating an interface between the collagen fibril and the surrounding medium. Assuming a constant surface tension $\tilde{\gamma}$, and taking the fibril length to be infinite (and so ignoring end effects), the surface energy per unit volume is

$$\tilde{E}_{\text{surf}} = \frac{2\tilde{\gamma}}{\tilde{R}} \tag{4.5}$$

---

[1] Note that for non-zero twist field, the fibril period does not equal the preferred period $2\pi/\eta \neq d_{\parallel}$

Combining eqns 4.2, 4.3, 4.4, and 4.5 gives the total free energy of the fibril as a function of radius $\tilde{R}$, D-band modulation amplitude $\tilde{\delta}$, D-band modulation period $2\pi/\tilde{\eta}$, and twist angle field $\psi(\tilde{r})$.

In the remainder of this thesis, variables with a tilde on them will have units, and those without will be reduced to dimensionless form. We can re-write eqn 4.1 in dimensionless form by dividing through by $\tilde{K}_{22}\tilde{q}^2$, measuring radial distance in $r = \tilde{r}\tilde{q}$ (and so $R = \tilde{R}\tilde{q}$), and inverse D-band in $\eta = \tilde{\eta}\tilde{d}_{\parallel}$, to get the main model equation of this chapter,

$$\begin{aligned} E_{\text{tot}} =& \frac{2}{R^2}\int_0^R rdr\left(\frac{1}{2}\left(1-\frac{d\psi}{dr}-\frac{\sin 2\psi}{2r}\right)^2+\frac{1}{2}K_{33}\frac{\sin^4\psi}{r^2}\right) \\ &+\frac{\Lambda\delta^2}{2R^2}\int_0^R rdr\big(4\pi^2-\eta^2\cos^2\psi(r)\big)^2 \\ &+\frac{\omega\delta^2}{2}\left(\frac{1}{2}\delta^2-1\right)-(1+k_{24})\frac{\sin^2\psi(R)}{R^2}+\frac{2\gamma}{R}. \end{aligned} \tag{4.6}$$

Examining eqn 4.6, we see that there are five free, dimensionless parameters which control the behaviour of our system (see Appendix A.3 for a list of parameters defined in terms of their dimensional counterparts). Three of these parameters, K_{33}, k_{24}, and γ, have been mapped out in a previous study for the case of a no D-band ($\delta = 0$) model of fibril structure [4]. $K_{33} = \tilde{K}_{33}/\tilde{K}_{22}$ and $k_{24} = \tilde{k}_{24}/\tilde{K}_{22}$ are the ratios of the bend and saddle-splay elastic constants to that of the twist elastic constant. Consistent with our previous work we will fix $K_{33} = 30$, a value motivated by experimental measurements on molecules with similar aspect ratios [109], but allow k_{24} to vary freely due to lack of experimental data on similar molecules. $\gamma = \tilde{\gamma}/(\tilde{K}_{22}\tilde{q})$ is the ratio between surface tension and intrinsic twist of the molecules, and will also be a free parameter of our model.

The final two (free) dimensionless parameters, $\Lambda = 2\tilde{\Lambda}\tilde{\delta}_0^4/(3\tilde{K}_{22}\tilde{q}^2\tilde{d}_{\parallel}^4)$ and $\omega = 2\tilde{\omega}\tilde{\delta}_0^4/(3\tilde{K}_{22}\tilde{q}^2)$, correspond to the coupling strength between the D-band and the molecular twist (see eqn 4.3) and the strength of the D-band double well potential (eqn 4.4), respectively. The former is proportional to the Young's modulus at zero twist. We expect the latter to be bounded by the polymerization energy of collagen fibrils, as the double well energy stabilizes the D-band structure.

If we instead consider a bulk chiral nematic phase[2] with director field $\boldsymbol{n}(\boldsymbol{r})$ =

---

[2] Also known as a cholesteric phase.

$\cos(qz)\hat{x}+\sin(qz)\hat{y}$ and no D-band modulation ($\delta = 0$), the energy per unit volume is $E_{\text{chiral}} = 0$. In principle, it is possible to extend our new model to allow for a density-modulated chiral nematic phase, which would alter $E_{\text{chiral}} \neq 0$ and allow for a more general comparison between fibril and chiral nematic phases. We do not attempt to obtain such a generalization here. Instead, we note that whatever this energy (E_{chiral}) might be, it will only depend on the parameters Λ and ω, as the Frank free energy and surface tension terms become zero in chiral nematics [3, 4].

4.2.2 Free energy minimization

We minimize eqn 4.6 with respect to $\psi(r)$ for chosen initial values of R, η, and δ using standard Euler-Langrange energy minimization [3]. This corresponds to solving the free boundary value problem described by the Ordinary Differential Equation (ODE)

$$\begin{aligned}\frac{d}{dr}\left(r\frac{d\psi}{dr}\right) =&\Lambda\delta^2\eta^2 r(4\pi^2 - \eta^2\cos^2\psi)\cos\psi\sin\psi\\ &+ 1 - \cos(2\psi)\left(1 - \frac{\sin(2\psi)}{2r}\right)\\ &+ K_{33}\frac{\sin(2\psi)\sin^2\psi}{r},\end{aligned} \tag{4.7a}$$

subject to the boundary conditions

$$\psi(0) =0, \tag{4.7b}$$

$$\left.\frac{d\psi}{dr}\right|_{r=R} =1 + k_{24}\frac{\sin(2\psi(R))}{2R}. \tag{4.7c}$$

Eqn 4.7b is required to ensure that $r = 0$ is not singular, and eqn 4.7c is a free boundary condition. To solve eqns 4.7a-4.7c, we implement a standard numerical finite-difference relaxation technique as discussed in Appendix B.1 (see also refs. [4] and [114]). We will refer to this twist angle field which minimizes eqn 4.6 at a specified R, η, and δ, as $\psi^*(r)$.

If we calculate $\psi^*(r)$ each time we take a step in R, η, δ space, we can define a cost function, $E^*(R, \eta, \delta) = E(R, \eta, \delta; \psi^*(r))$ for a given parameter set (K_{33}, k_{24}, γ, Λ, ω) which is identical to eqn 4.6, aside from requiring the twist field to be the minimizing form $\psi^*(r)$. By solving for $\psi^*(r)$ with each optimization step, we can minimize the three variable cost function $E^*(R, \eta, \delta)$ which is equivalent to minimizing eqn

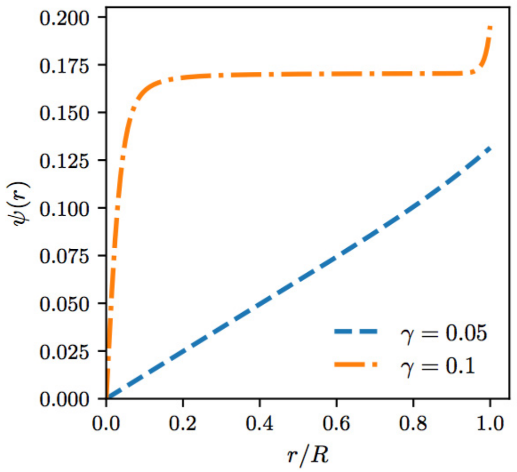

Figure 4.1: Local twist angle, $\psi(r)$, *vs* re-scaled radial distance from the fibril centre, r/R for two different values of dimensionless surface tension, γ. $\psi(r)$ corresponds physically to the local (average) angle that tropocollagen molecules would have with respect to the cylindrical (z) axis of the fibril. The twist is constrained to always be zero at the fibril centre, and the twist at surface of the fibril, $\psi(R)$, is the value of $\psi(r)$ at $r/R = 1$ in the above plot. The surface twist values are $\psi(R) = 0.131$ and $\psi(R) = 0.195$ for $\gamma = 0.05$ and $\gamma = 0.1$, respectively. We show the re-scaled radial distance for ease of comparison between the two $\psi(r)$ shapes, as the true fibril radius is $R = 0.079$ for the blue dashed ($\gamma = 0.05$) curve and $R = 1.779$ for the orange dash dotted ($\gamma = 0.1$) curve. Curves which are qualitatively similar to the blue dashed ($\gamma = 0.05$) curve will be referred to as "linear twist" fibrils. Curves which are qualitatively similar to the orange dash-dotted curve will be referred to as "frustrated twist" fibrils. All other tune-able parameters are held constant at $k_{24} = 0.75$, $\Lambda = 27$, and $\omega = 10$.

4.6. We use the GNU Scientific Library's Broyden-Fletcher-Goldfarb-Shanno (BFGS) solver (with numerical derivatives) to perform the minimization [148]. All numerical implementations are available online via GitHub [149], and in Appendix B we outline in more detail the numerical algorithms used in this work and why we have chosen these algorithms in particular.

4.3 Results

4.3.1 Linear and frustrated twist in double-twist fibrils

In Figure 4.1, we showcase the equilibrium local twist angle, $\psi(r)$, for two γ values, while holding $k_{24} = 0.9$, $\Lambda = 27$, and $\omega = 10$ constant. We have chosen the x-axis of Figure 4.1 to be the re-scaled radial distance from the fibril centre r/R (where R is the true fibril radius) for ease of comparison between the blue dashed ($\gamma = 0.05$) and orange dash-dotted ($\gamma = 0.1$) curves. For $\gamma = 0.05$, $\psi(r)$ is almost perfectly linear in radial distance from the fibril centre, whereas for $\gamma = 0.1$, $\psi(r)$ is highly non-linear. As we shall demonstrate shortly, the qualitative difference between these two curves are of great importance in our model, and so it will be helpful to assign specific names to both of them. For the remainder of this chapter, $\psi(r)$ curves which are qualitatively similar to the blue dashed curve in Figure 4.1 will be referred to as "linear twist" fibrils, while $\psi(r)$ curves qualitatively similar to the orange dash-dotted curve in Figure 4.1 will be referred to as "frustrated twist" fibrils (due to the suppression of twist between fibrils). We also note here that both linear and frustrated fibrils do in general have a non-zero D-band amplitude δ. In Figure 4.1 the linear twist fibril structure ($\gamma = 0.05$) has D-band amplitude of $\delta = 0.975$, D-band period $2\pi/\eta = 0.996$, and the frustrated twist fibril structure ($\gamma = 0.1$) has $\delta = 0.998$, $2\pi/\eta = 0.985$.

4.3.2 Characterization of Phase Transition at constant ω and Λ

In Chapter 3, we have examined the phase diagram of double-twist fibrils with no axial D-band (i.e. $\Lambda = \omega = 0$) in terms of γ and k_{24}. Thus, to begin the parameter space investigation of our new model, we first investigate the collagen phase diagram while holding our new parameters Λ and ω at constant but non-zero values. For the remainder of this section, we hold our (D-band) parameters constant, $\Lambda = 27$ and $\omega = 10$. We have selected these values of Λ and ω as they allow us to demonstrate all of the qualitatively distinct behaviour that our model predicts. We examine different values of Λ in the subsequent section and systematically explore the Λ and ω parameter space in the supplementary materials.

We start by presenting the equilibrium energy per unit volume, E, (obtained

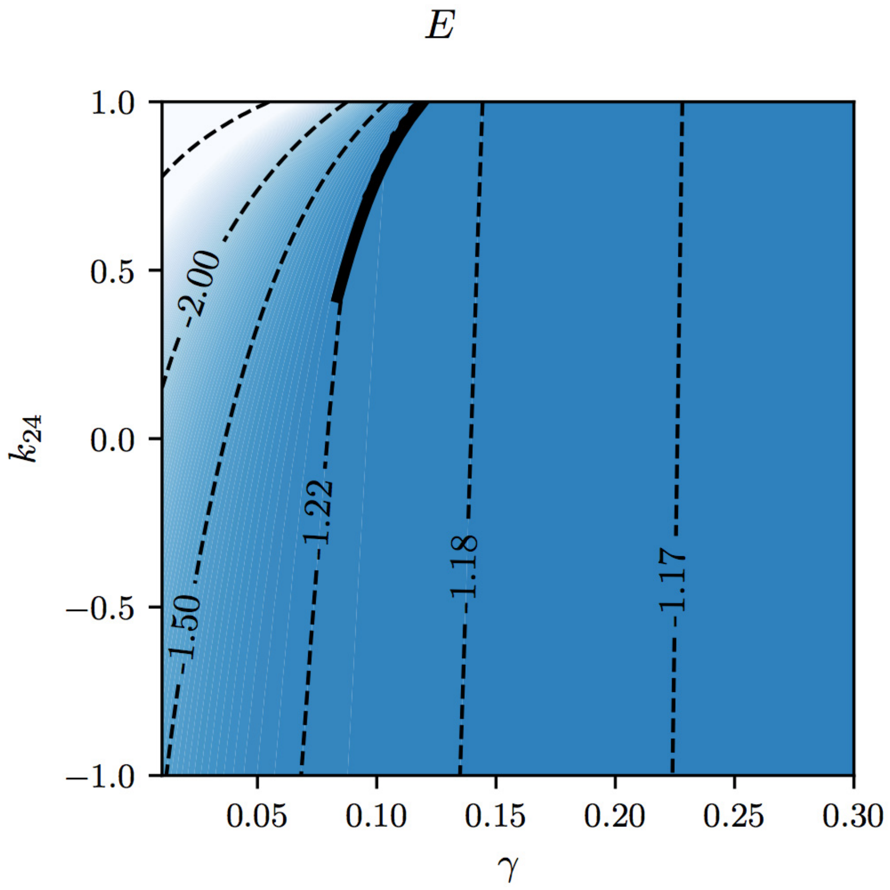

Figure 4.2: Contours of constant equilibrium energy per unit volume of fibril, E, in the k_{24} *vs* γ plane. The energy increases with increasing γ and decreasing k_{24}, but remains smaller than zero for the everywhere in the figure. At a given k_{24}, linear twist fibrils are thermodynamically stable in the upper left hand corner of the phase diagram (to the left of the thick black line). Frustrated twist fibrils are thermodynamically stable to the right of the thick black line. The thick black line beginning at the critical point $(\gamma,\, k_{24}) \simeq (0.08356, 0.4205)$ indicates coexistence between the linear twist fibril phase and the frustrated twist fibril phase (similar to the coexistence line in a liquid-gas phase diagram). A discontinuous phase transition from linear twist (left of coexistence line) to frustrated twist (right of coexistence line) fibrils occurs when this line is crossed in parameter space. $\Lambda = 27$ and $\omega = 10$ for this figure.

by minimizing eqn 4.6) in Figure 4.2[3]. The fibrils with the lowest E are in the upper left hand corner, consistent with our previous no D-band model. However, by allowing a non-zero D-band modulation, we find that a discontinuous phase transition

---

[3] In this chapter, “equilibrium” phases are those of the lowest energy which still have a double-twist molecular orientation. It is entirely possible that within our parameter space, more stable configurations such as a bulk cholesteric phase may occur.

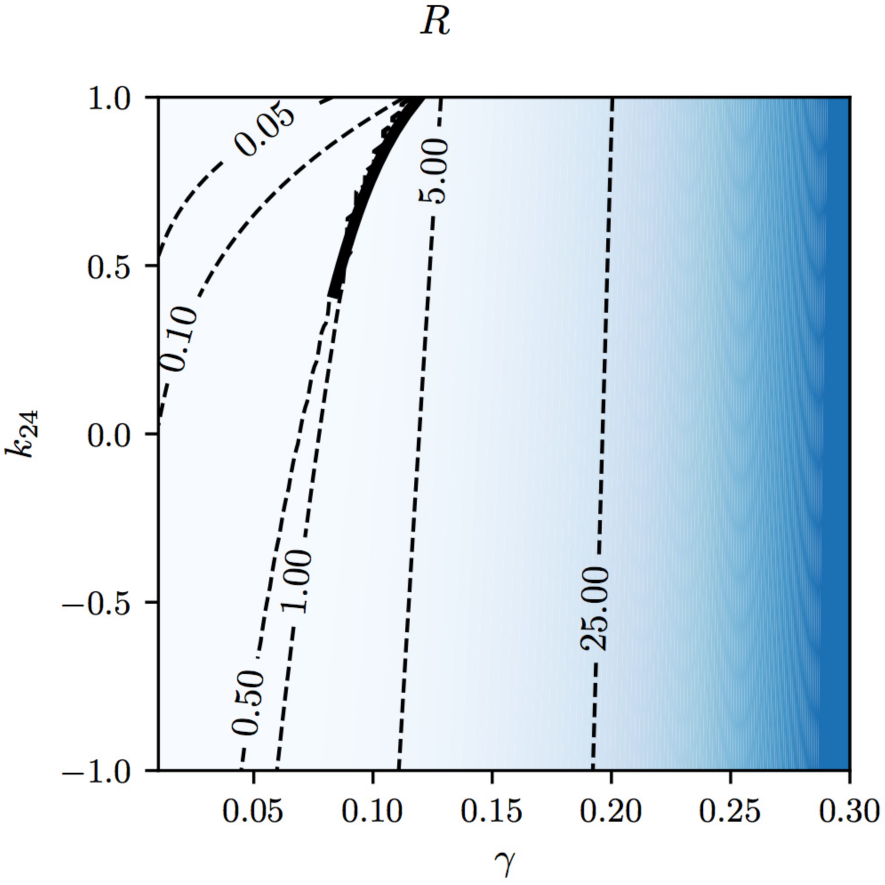

Figure 4.3: Contours of constant fibril radius R in the k_{24} *vs* γ plane. The thick black line corresponds to the coexistence line between linear and frustrated twist fibril phases. A discontinuous jump in R occurs across the coexistence line as γ is increased at constant k_{24}. γ strongly controls the size of fibrils, which range from $R = 0.001$ near $\gamma = 0.01$ at $k_{24} = 1$ to $R = 150$ at $\gamma = 0.3$. We hold $\Lambda = 27$ and $\omega = 10$ constant.

appears between linear and frustrated twist fibrils, which was not present in the previous model. The coexistence line of this phase transition is indicated by the thick black line in Figure 4.2. To the left of the coexistence line, linear twist fibrils are thermodynamically stable, while frustrated twist fibrils are thermodynamically stable to the right of the coexistence line. The lower end of the coexistence line (i.e. the critical point), is found at approximately $(\gamma, k_{24}) \simeq (0.08356, 0.4205)$. We also find that with the presence of D-band modulations, $E < 0$ throughout parameter space region shown in Figure 4.2 for the specified $\Lambda = 27$ and $\omega = 10$ values, indicating that some form of fibril (be it linear or frustrated twist) is always stable with respect to a chiral nematic phase with no cohesive D-band like interactions. As mentioned in the last paragraph of section 4.2.1, introducing D-band interactions in a chiral nematic phase will cause the chiral nematic free energy to depend on Λ and ω only. Therefore, any phase transition line between a density-modulated chiral nematic and either fibril phase will be along a constant E_{chiral} contour in Figure 4.2. Determining the value of E_{chiral}, however, would require a more complete phase field theory which allows for D-band interactions in the chiral nematic phase. We therefore take E_{chiral} as a constant but unknown value.

In Figure 4.3 we show contours of constant fibril radius, R, *vs* γ and k_{24}. The large increase in R with increasing γ indicates that a wide range of equilibrium fibril radii may be accessible if γ is easily tuned. To the left of the coexistence (thick black) line, the linear twist fibril structure is the equilibrium phase, while to the right, the frustrated twist structure is the equilibrium phase. Thus we see from this constant R contour plot that the linear twist phase has smaller fibril radius than the frustrated twist phase in all cases. Below the coexistence line ($k_{24} < 0.4205$), there is no discrete transition from linear to frustrated twist, and so there is also no discrete jump in R.

In Figure 4.4a, we examine the behaviour of R with increasing γ at a constant $k_{24} = 0.9$ slice across the coexistence line. Meta-stability in the phases occurs when two minima are present in the energy landscape ($E(R, \eta, \delta)$ in eqn 4.6). The lower energy minimum corresponding to the equilibrium phase is indicated by a thick line, and any meta-stability is shown by thin lines. So in Figure 4.4a, any thin lines to the left of the jump in R (at smaller γ) correspond to meta-stable frustrated twist fibrils. Conversely, any thin lines to the right of the jump in R (at larger γ) correspond to meta-stable linear twist fibrils. We find a small but significant region of meta-stability in the frustrated twist phase, and to a lesser extent, in the linear twist phase as we cross the coexistence line. This region of meta-stability tends to increase with k_{24} along the coexistence line, which is why we have selected a fairly large k_{24} value in Figure 4.4a. In Figure 4.4b, we follow the coexistence line from its starting (critical) point at (γ, k_{24})=(0.08356,0.4205) up to the cutoff of $k_{24} = 1$, and show the ratio of the coexisting linear and frustrated twist radii, R_ℓ and R_f, respectively. This ratio increases monotonically with increasing γ and k_{24}, and becomes as large $R_f/R_\ell = 50$

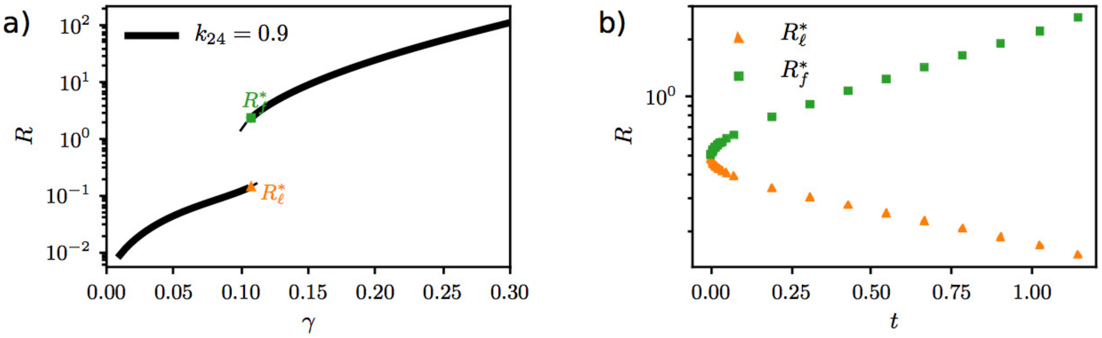

Figure 4.4: (a) Fibril radius R *vs* reduced surface tension γ at constant $k_{24} = 0.9$. The thick black line corresponds to an equilibrium fibril radius, whereas thin black lines correspond to radii of meta-stable fibrils. All radii in the bottom branch are in a linear twist fibril configuration. All radii in the upper branch are in a frustrated twist configuration. The thin meta-stable lines demonstrate that both fibril phases may exist at the same point in parameter space, and so equilibrium (i.e. equal energy E) coexistence of both the linear and frustrated phases are possible at a single point. The equilibrium coexisting radii are denoted by an orange triangle (linear fibril phase, R_ℓ^*) and a green square (frustrated fibril phase, R_f^*), near $\gamma = 0.11$. (b) Fibril radii of the linear twist fibril phase, R_ℓ^*, and frustrated twist fibril phase, R_f^*, along the coexistence line in γ, k_{24} space (see e.g. Figure 4.3 for coexistence line). The gap in radius spans nearly two orders of magnitude at large γ values (and so near $k_{24} = 1$) along the coexistence line. In both (a) and (b), $\Lambda = 27$ and $\omega = 10$. Here $t = (k_{24}(s) - k_{24}^c)/k_{24}^c$, where $k_{24}(s)$ traces k_{24} out along the coexistence line, $k_{24}^c = 0.4205$ is the critical k_{24} value.

at $k_{24} = 1$. The x-axis of Figure 4.4b, t, is defined as

$$t = \frac{(k_{24}(s) - k_{24}^c)}{k_{24}^c}, \quad s \text{ along coexistence} \tag{4.8}$$

where $k_{24}^c = 0.4205$ is the critical k_{24} value, and $(\gamma(s), k_{24}(s))$ trace out the coexistence line.

In Figure 4.5 we present contours of constant surface twist, $\psi(R)$, *vs* γ and k_{24}. $\psi(R)$ increases monotonically with γ, similar to R. Unlike fibril radius, $\psi(R)$ is not monotonic in k_{24}, particularly in the linear twist fibril phase to the left of the coexistence line. In the linear twist fibril phase, $\psi(R)$ ranges from $\psi(R) = 0.01\,\mathrm{rad}$ near $\gamma = 0.01$, up to $\psi(R) = 0.18\,\mathrm{rad}$ near the coexistence line at $k_{24} = 1$. In the frustrated twist phase (right of the coexistence line), $\psi(R) = 0.18$ near coexistence and near $\gamma = 0.3$ becomes larger than $0.5\,\mathrm{rad}$. We present the parameter space

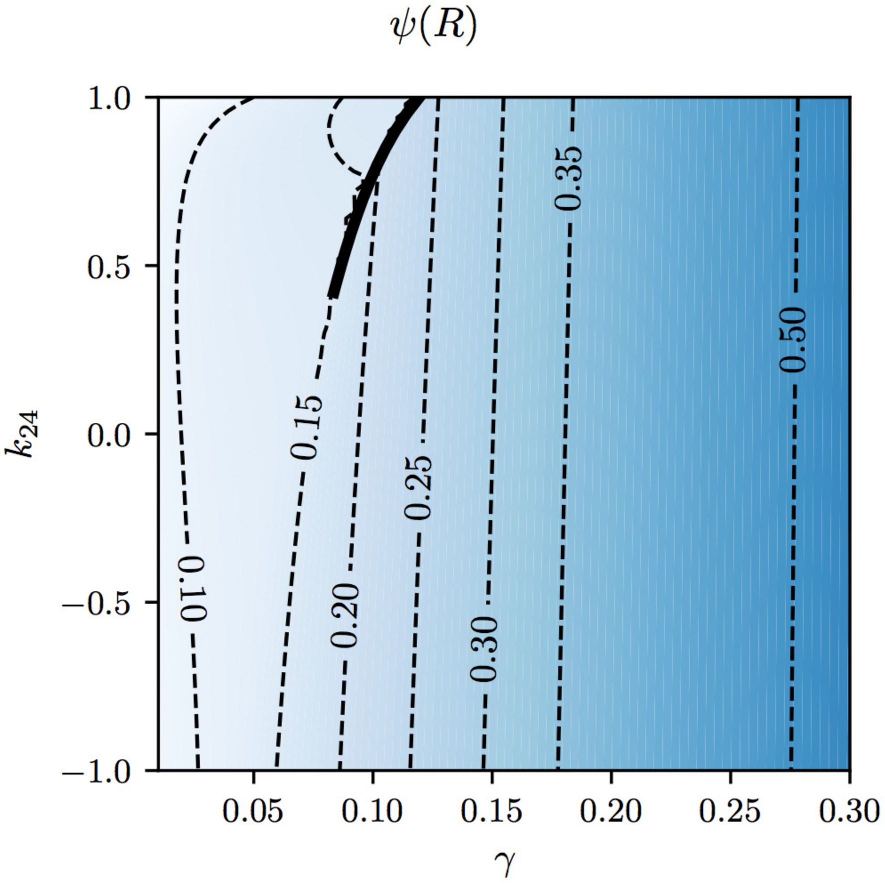

Figure 4.5: Contours of constant surface twist, $\psi(R)$, in the k_{24} *vs* γ plane. $\psi(R)$ ranges from $\psi(R) = 0.01$ near $\gamma = 0.01$, and increases up to $\psi(R) > 0.5$ at $\gamma = 0.3$. $\psi(R)$ is monotonic in γ to the left of the coexistence (thick black) line, and monotonic in both γ and k_{24} to the right of the coexistence line.

dependence of η and δ in (supplementary) Figures 4.16 and 4.17, respectively.

Figures 4.6a and 4.6b show the dependence of the fibril surface twist on γ at constant $k_{24} = 0.9$ and along the linear/frustrated twist coexistence line, respectively. The behaviour of $\psi(R)$ in Figure 4.6a is quite analogous to the fibril radius behaviour in Figure 4.4, with $\psi(R)$ increasing monotonically (albeit much more slowly than R) as γ is increased with constant $k_{24} = 0.9$. The lower $\psi(R)$ branch corresponds to the linear twist fibril phase, and the upper branch to that of the frustrated twist phase. The meta-stability of the frustrated twist phase persists longer than that of the linear twist phase, and equilibrium (i.e. equal E) coexistence between the linear and frustrated twist phase occurs near $\gamma = 0.11$. In contrast to R, the surface twist along

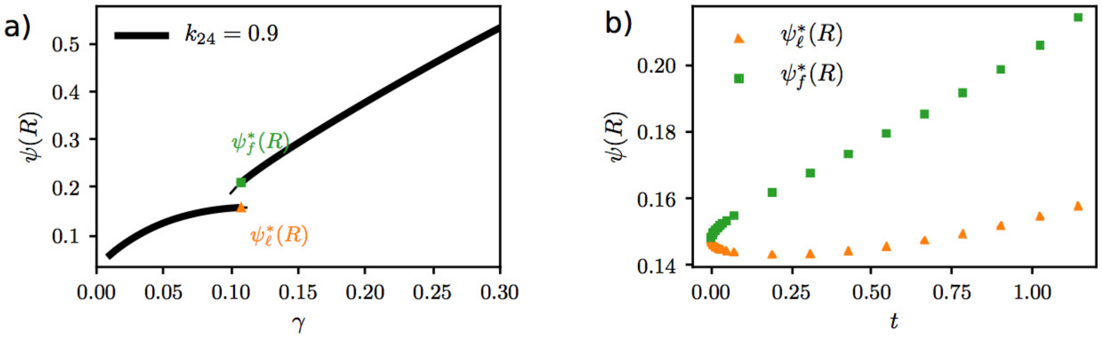

Figure 4.6: (a) Fibril surface twist $\psi(R)$ *vs* reduced surface tension γ at constant $k_{24} = 0.9$. The thick black line corresponds to an equilibrium surface twist, whereas thin black lines correspond to $\psi(R)$ of meta-stable fibrils. All data in the bottom (top) branch are in a linear (frustrated) twist fibril configuration. The thin meta-stable lines demonstrate that both fibril phases may exist at the same point in parameter space, and so equilibrium (i.e. equal energy E) coexistence of both the linear and frustrated phases are possible at a single point. The equilibrium coexisting $\psi(R)$ values are denoted by an orange triangle (linear fibril phase, $\psi_\ell^*(R)$) and a green square (frustrated fibril phase, $\psi_f^*(R)$), near $\gamma = 0.11$. (b) Surface twist of the linear twist fibril phase, $\psi_\ell^*(R)$ and frustrated twist fibril phase, $\psi_f^*(R)$, along the coexistence line in γ, k_{24} space (see e.g. Figure 4.3 for coexistence line). The difference in twist for the coexisting phases increases with γ and k_{24}. In both (a) and (b), $\Lambda = 27$ and $\omega = 10$.

the coexistence line increases monotonically with distance from the critical point near (γ, k_{24})=(0.08356,0.4205) in both the frustrated and linear twist phases.

4.3.3 Characterization of Phase Transition at varying Λ

In this section, we demonstrate that a discontinuous phase transition occurs while Λ is increased for suitable choice of γ and k_{24} values. We have chosen two representative pairs of (γ, k_{24})=(0.04,0) and (γ, k_{24})=(0.12,1) as a case study, with only the latter pair displaying a discontinuous phase transition. In Figure 4.7, we illustrate the two qualitatively different behaviours of our observables R, $\psi(R)$, and $\psi(r)$ for these two (γ, k_{24}) pairs as Λ increases, at a constant $\omega = 10$. We show a more systematic examination of our model as ω is varied in (supplementary) section 4.6.

In Figure 4.7a, the dimensionless radius, R, of the fibril initially decreases as Λ is increased for both sets of (γ, k_{24}) values. At (γ, k_{24})=(0.04,0.0), there is a smooth

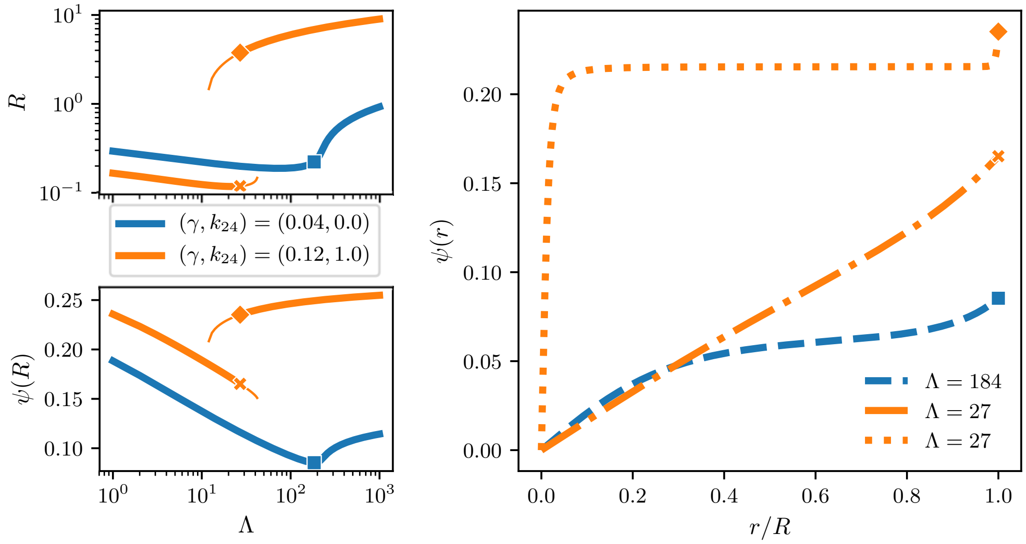

Figure 4.7: Radial structure of phase-field collagen fibrils. (a) Fibril radius, R, *vs* the coupling strength Λ between axial and radial structure at constant $\omega = 10$. The solid blue line is for surface tension $\gamma = 0.04$ and saddle-splay $k_{24} = 0$. Similarly, the solid orange line indicates (γ, k_{24})=(0.12,1). The thin line in both branches indicates a region of meta-stability. (b) Surface twist, $\psi(R)$, *vs* Λ for (γ, k_{24})=(0.04,0) (blue solid line) and (γ,k_{24})=(0.12,1) (both branches of the orange solid line). Thin lines correspond to meta-stable regions as in (a). (c) Local, average orientation of tropocollagen molecules within a collagen fibril, $\psi(r)$, *vs* radial distance from the fibril centre r, re-scaled by the fibril radius R. The dashed blue line is for (γ, k_{24})=(0.04,0) and $\Lambda = 184$. This corresponds to the turning point in (b) where surface twist $\psi(R)$ begins to increase with Λ. The dotted orange line is for (γ, k_{24})=(0.12,1) in the large radius branch of (a) when both structures have equal energy ($\Lambda = 27$). The dot-dashed orange line corresponds to $\psi(r)$ for the small radius branch at coexistence, $\Lambda = 27$.

transition from decreasing R with Lambda to monotonically increasing R at $\Lambda = 74$. In Figure 4.7b, a similar initial decrease in surface twist occurs as Λ is increased. However, the transition to increasing $\psi(R)$ occurs at $\Lambda = 184$, and so for the region $74 < \Lambda < 184$, there is a decrease in $\psi(R)$ while R is increasing.

At (γ, k_{24})=(0.12,1.0), a discontinuity in both R and $\psi(R)$ occurs at $\Lambda_c = 27$, as can be seen in Figure 4.7a and Figure 4.7b, respectively. The overlapping of the lower and upper branches of R and $\psi(R)$ indicate that two energy minima exist within the domain $11 \leq \Lambda \leq 43$ (see Figure 4.8a below). Thin lines illustrate the regions of meta-stability in both the small radius ($\Lambda < \Lambda_c$) and large radius ($\Lambda > \Lambda_c$) fibril structures. This meta-stability is indicative of a discontinuous phase transition between the small and large fibril structures. From Figure 4.7a and Figure 4.7b, we see that at the coexistence (i.e. equal energy) point $\Lambda_c = 27$, the radius and twist of the small (large) fibril structure are $R = 0.1188$ and $\psi(R) = 0.1652$ ($R = 3.766$ and $\psi(R) = 0.2352$), respectively. When there is a discontinuous transition in our observable parameters (as is the case for (γ, k_{24})=(0.12,1) in Figure 4.7 above), we refer to fibrils which are in the lower branch ($\Lambda < \Lambda_c$ for equilibrium fibrils) as being part of the "linear" twist fibril phase. Similarly, we will refer to fibrils in the upper branch as being members of the "frustrated" twist fibril phase.

In Figure 4.7c, we show average local twisting of tropocollagen molecules within the fibril, $\psi(r)$, as the radial distance r from the centre of the fibril increases. The three curves in Figure 4.7c exemplify the inner structure of fibrils in three different cases. The dashed, blue line in Figure 4.7c shows the structure of a fibril at $\Lambda = 184$ with (γ, k_{24})=(0.04,0) (i.e. when no discontinuous transition in R and $\psi(R)$ is present for variations in Λ). We have chosen $\Lambda = 184$ as it is near where R and $\psi(R)$ begin increasing with Λ, and shows intermediate behaviour between a linear twist fibril and a frustrated twist fibril. For $\Lambda < 184$, $\psi(r)$ would become more linear and so the middle, flattened regime of the blue curve in Figure 4.7c would reduce in size. Increasing values of $\Lambda > 184$ (and so stronger coupling between the D-band modulation and the twisting of the molecules) cause $\psi(r)$ to flatten out even more at intermediate distances between the fibril centre and the fibril surface when compared to the dashed blue curve in Figure 4.7c. However, no discontinuous transition occurs for when tuning Λ for the parameter set of the blue dashed curve. Conversely, when (γ, k_{24})=(0.12,1), we see a distinct transition in the structure of $\psi(r)$ right at coexistence point, Λ_c. This feature is illustrated in the two orange curves of Figure 4.7c. The dotted orange line corresponds to the linear twist phase, and the dot-dashed orange line corresponds to the frustrated twist phase, both of which are at $\Lambda_c = 27$.

Figure 4.8 shows the free energy per unit fibril volume, E, D-band period, $2\pi/\eta$, and D-band amplitude, δ, that concur with the radial configurations of Figure 4.7 (at $\omega = 10$). The thick orange line in Figure 4.8a is the equilibrium E at (γ, k_{24})=(0.12,1). The thin orange lines in the region $11 \leq \Lambda \leq 43$ which branch out at $\Lambda = 27$ correspond to the meta-stable energies of the linear twist phase ($\Lambda_c > 27$) and the frustrated twist phase ($\Lambda_c < 27$). In contrast, E *vs* Λ for (γ, k_{24})=(0.04,0) (blue

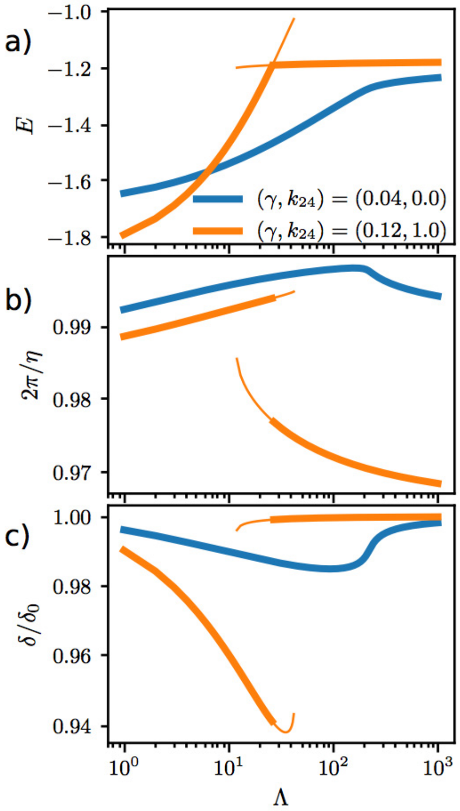

Figure 4.8: Free energy and axial structure of a phase-field collagen fibril. (a) The free energy per unit volume of fibrils, E, *vs* the coupling strength Λ between axial and radial structure for two sets of surface tension γ and saddle-splay k_{24} parameters at constant $\omega = 10$. The blue and orange curves correspond to (γ, k_{24})=(0.04,0) and (γ, k_{24})=(0.12,1), respectively. Two free energy minima occur in the region $11 \leq \Lambda \leq 43$. The thin orange lines which emerge at $\Lambda = 27$ indicate meta-stable fibril structures which arise from the presence of a second free energy minimum. (b) The axial period of the D-band (density) modulations, $2\pi/\eta$, *vs* Λ. $2\pi/\eta = 1$ would correspond to a perfect D-band (i.e. with no molecular twist). (c) D-band modulation amplitude, δ, *vs* Λ. Any non-zero value of δ indicates that axial periodicity is present in the fibril.

line) has no region of meta-stability. For both parameter sets, the energy increases rapidly for small Λ, and flattens out for large Λ. Unlike R and $\psi(R)$, E does vary with variations in ω as can be seen in (supplementary) Figure 4.18. However, this variation is essentially linear in ω, as expected from eqn 4.4. In Figure 4.8b and Figure 4.8c, we see that the dependence of D-band amplitude and period, δ and $2\pi/\eta$ also have two distinct regions above and below the critical value $\Lambda_c = 27$ for (γ, k_{24})=(0.12,1), and a single smooth transition for (γ, k_{24})=(0.04,0).

In this section, we have shown that increasing the coupling strength Λ between the twist and D-band allows for a discontinuous phase transition between the linear twist fibril phase and the frustrated twist fibril phase. However, from e.g. Figure 4.8a, it is clear that this phase transition does not always occur, but instead depends on the values of γ and k_{24} as well, which is consistent with the results of the previous section. We have now identified three parameters which strongly control the structure of collagen fibrils: γ, k_{24}, and Λ.

4.3.4 Elastic properties of the phase field collagen fibril

In order to probe the mechanical response and elastic properties of our model, we apply a strain, ϵ, to the collagen fibril. We define $2\pi/\eta_{\text{eq}}$ and R_{eq} as the equilibrium ($\epsilon = 0$) D-band period and radius, respectively. By setting $\eta = \eta_{\text{eq}}/(1+\epsilon)$ while at the same time requiring the fibril to be incompressible, $R = R_{\text{eq}}/\sqrt{1+\epsilon}$, we can minimize eqn 4.6 with respect to δ and $\psi(r)$ at these strained values of η and R. With this procedure, we are assuming a quasi-static extension of the fibril, where strain is applied slowly enough for δ and $\psi(r)$ to relax to (strained) equilibrium at each fixed strain. We expect this to apply to in vitro assembled fibrils that have not been significantly cross-linked.

The resulting stress (defined as $\sigma = dE/d\epsilon$), δ, and $\psi(R)$ for increasing ϵ are shown in Figure 4.9, with $\gamma = 0.04$, $k_{24} = 0$, and $\omega = 10$. In Figure 4.9a, the initial relationship between stress and strain is linear for all values of Λ. As ϵ increases, the fibril begins to stiffen until it reaches some maximum stress, σ_{max} at strain $\epsilon(\sigma_{\text{max}})$. In Figure 4.9b, the D-band amplitude initially increases slightly. As strain increases past σ_{max}, $\delta \to 0$ rapidly. In Figure 4.9c, the surface twist initially decreases slowly with strain until $\epsilon = \epsilon(\sigma_{\text{max}})$, at which point the twist increases slightly before $\delta \to 0$. In (supplementary) Figure 4.23 the mechanical response for a fibril with $\gamma = 0.12$, $k_{24} = 1$, and $\omega = 10$ is characterized for $\Lambda = 1.0$ and $\Lambda = 10.0$. The behaviour is qualitatively similar in between the two different (γ, k_{24}) pairs, with $\delta \to 0$ as strain increases, and an initially linear stress *vs* strain curve. For the remainder of this chapter, any figures which are measured *vs* strain will be cut off at $\epsilon(\sigma_{\text{max}})$, as decreasing stress with increasing strain is unphysical behaviour which corresponds to our D-band amplitude $\delta \to 0$, and so the D-band strain becomes undefined near this

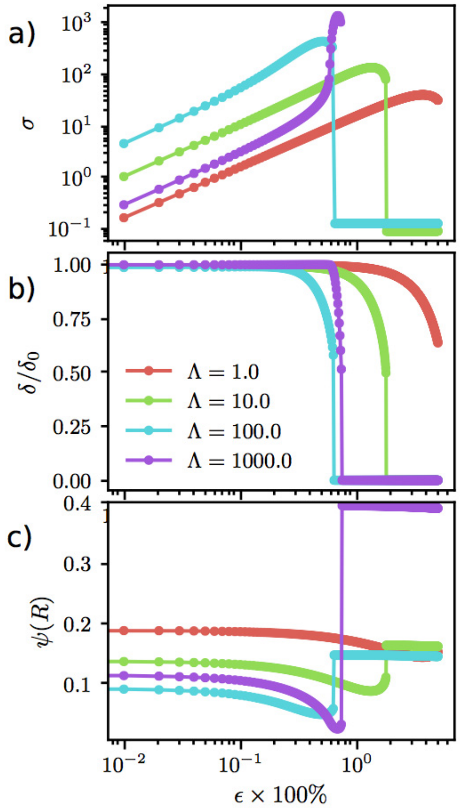

Figure 4.9: Mechanical and structural response of straining the collagen fibril D-band for an incompressible fibril. (a) Stress, $\sigma = dE/d\epsilon$, *vs* strain (ϵ) percentage for $\Lambda = 1$ (red curve), $\Lambda = 10$ (green curve), $\Lambda = 100$ (light blue curve), and $\Lambda = 1000.0$ (purple curve), all at $\gamma = 0.04$, $k_{24} = 0$, and $\omega = 10$. The stress increases linearly in all cases for small strain. As Λ increases, the stress becomes non-linear at large strain. Notably, the initial slope in the stress (i.e. the Young's modulus) does not monotonically increase with strain. (b) D-band (density) amplitude δ *vs* strain percentage. At small strain, the amplitude remains close to its equilibrium ($\epsilon = 0$ value), while at large strain δ decreases rapidly to zero ($\delta = 0$ corresponds to a fibril with no D-band). (c) Surface twist $\psi(R)$ *vs* strain percentage. $\psi(R)$ decreases with increasing strain initially, and increases slightly before the D-band disappears ($\delta = 0$) at the discontinuity in $\psi(R)$.

point[4]. Notably, the value of δ, as well as $\sigma_{\max}$ and the corresponding $\epsilon(\sigma_{\max})$ are the only quantities in our model which vary significantly with ω (see supplementary Figure 4.24).

The linear relationship between σ and ϵ at small strain in Figure 4.9a and in

---

[4] Physically, this would represent mechanical instability of our model, and such large strains that our model assumptions would not apply in detail.

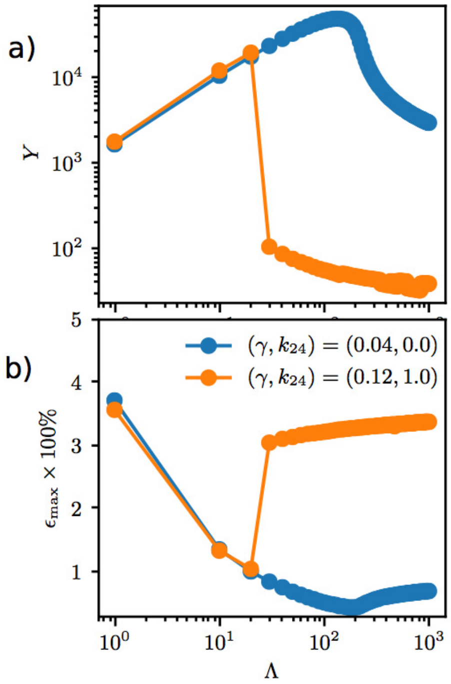

Figure 4.10: (a) Young's modulus, Y, *vs* Λ for the parameter values (γ, k_{24})=(0.04,0) in blue and (γ, k_{24})=(0.12,1) in orange. Both curves have $\omega = 10$. The maximum value of $Y = 62000$ occurs at $\Lambda = 170$ for the blue curve, and at the discontinuous transition point ($\Lambda = 27$) for the orange curve. (b) The maximum strain obtained before the stress begins to decrease, ϵ_{max}, *vs* Λ. The minimum $\epsilon_{\mathrm{max}} \times 100\% \simeq 0.39\%$ occurs at $\Lambda \simeq 220$ for the blue curve, and is bounded below by $\epsilon_{\mathrm{max}}\% > 1$ at the discontinuous transition point $\Lambda = 27$ for the orange curve.

(supplementary) Figure 4.23 allows us to extract a Young's modulus at small strain (linear regime), $Y \equiv d\sigma/d\epsilon$, from our model. In Figure 4.10a, we present Y *vs* Λ for (γ, k_{24})=(0.04,0) (blue) and (γ, k_{24})=(0.12,1) (orange) at $\omega = 10$. Both parameter sets exhibit an initial increase in Y with Λ, which is to be expected from the proportionality of the energy to that of Λ in eqn 4.6. If we compare the behaviour of Figure 4.10a to Figure 4.7, we see that the maximum Y occurs just before the fibril $\psi(r)$ structure transitions from the linear twist phase to the frustrated twist phase. In Figure 4.10b we show the strain at which stress is maximized, ϵ_{max} to characterize the amount of strain we can apply to our model before the fibril D-band breaks.

4.4 Discussion

4.4.1 Linear twist phase *vs* Frustrated twist phase

In the collagen fibril literature, two possibilities for the inner molecular orientation of fibrils are discussed. The first structure discussed, which we will refer to as the constant pitch/variable angle model [48], corresponds to that of a linearly increasing $\psi(r)$ with r in our model framework similar to strands of a rope winding about each other. At first glance, this linear $\psi(r)$ fibril structure appears to be incompatible with the robust, constant D-band that is typically observed in experiment, as it seems any molecular twist should induce a local D-band $d(r) = d_0 \cos\psi(r)$ from geometric arguments. In previous work, we determined that an approximately linear $\psi(r)$ is compatible with experimental observations of fibril radius R and surface twist $\psi(R)$ in corneal and tendon fibrils. However, we did not consider the axial (D-band) structure of the fibrils and so could not address whether the D-band would alter our picture of fibril structure. In this chapter, we do indeed still see linear $\psi(r)$, equilibrium fibril configurations with D-band present ($\delta \neq 0$), suggesting that it may still be reasonable to have linear twist with a constant D-band. For this to occur, it implicitly requires us to assume that the tropocollagen molecules within the fibril are subject to minor axial compression (tension) at the fibril centre (surface), as was hypothesized by Raspanti et al. [48]. We show in Figure 4.11 that the molecular strain for a fibril with $\psi(R) = 0.28$ (see Figure 4.12 below) would range between -1.48% in the fibril centre, and 2.41% at the fibril surface, $r = R$, where negative (positive) strain corresponds to molecular compression (tension). The total compression or elongation that a 300 nm tropocollagen molecule would have to experience to allow for such a strain is would be 4 nm (compression) and 7.2 nm (elongation). Given the semi-flexible nature of tropocollagen (persistence length ~ 50 nm)[5], it seems reasonable that the tropocollagen molecules would not be held taut in the fibril with zero applied (external) strain, and so small strain on each molecule would be insignificant energetically. We have not explicitly accounted for the energetic cost of stretching individual tropocollagen molecules in our continuum theory of fibril structure, but it contributes implicitly

---

[5] Persistence length quantifies the distance between two points along the arc length of a polymer whose tangent vectors are still pointing, on average, in the same direction subject to thermal fluctuations. See ref. [150] for a quantitative definition.

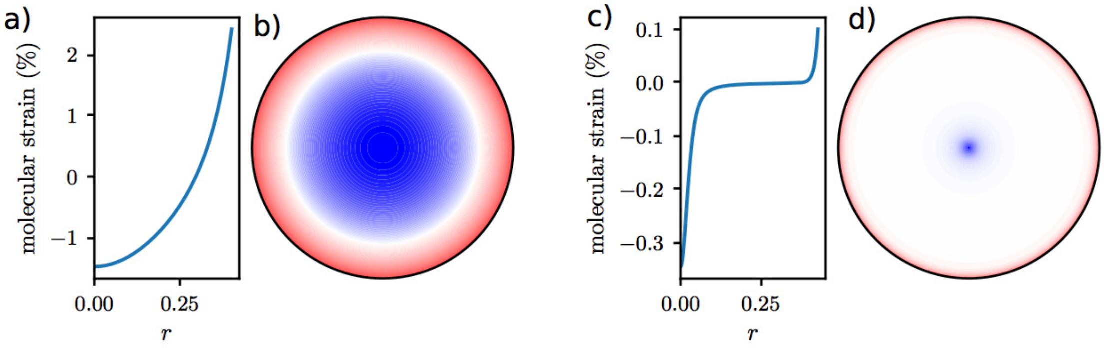

Figure 4.11: Local, longitudinal strain of tropocollagen molecules within (a),(b) linear and (c),(d) frustrated twist fibrils required to accommodate a constant D-band spacing. The molecular strain is defined as $(2\pi/\eta - \cos\psi(r))/\cos\psi(r)$ where the preferred, local D-band is $\cos\psi(r)$, but the actual D-band (constrained to be constant in our model) is $2\pi/\eta$. (b) and (d) cross-section of fibril illustrating compression of molecules (blue) at the fibril centre, $r = 0$ and the tension of molecules at the fibril surface (red). Linear $\psi(r)$ parameters used in (a) and (b): $\gamma = 0.15$, $k_{24} = 0.8$, $\Lambda = 0.9$, $\omega = 20$. Frustrated $\psi(r)$ parameters used in (c) and (d): $\gamma = 0.04$, $k_{24} = 0.5$, $\Lambda = 600$, $\omega = 20$.

through our phase field free energy (Λ) term. Another molecular mechanism that is coarse-grained into our Λ term is the energetic cost of increasing inter-molecular distance (through e.g. Van der Waals interactions). It is a combination of these two (and possibly other) mechanisms which allows the D-band to be strained. Therefore, our estimate of molecular strain in Figure 4.11 is really an upper limit, assuming (equilibrium) D-band strain is accommodated strictly by the elongation/compression of individual tropocollagen molecules.

The second fibril structure, initially proposed by Galloway [77] and subsequently by others [73, 151], is referred to as the constant angle model [48]. In our model framework, this would correspond to a constant value of $\psi(r)$ from the fibril centre to the radius R. This constant angle model is attractive for it's simplicity, as it does not require the added assumption of molecular compression and tension to accommodate a constant D-band throughout the fibril. Images of inner fibril structure taken by transmission electron micro-tomography [55], where fibrils were digitally reconstructed lengthwise into slices (analogous to flat sawn lumber), and these slices were used to construct a radial dependence of molecular orientation, similar to that of our $\psi(r)$ plots, appears to corroborate this constant angle model for corneal fibrils, with a constant twist of 17° observed. However, a constant angle structure is counter-intuitive considering the chiral nature of tropocollagen molecules; the expectation would be that molecules twist with respect to each other. The energetic cost of suppressing the twist gradient must then be outweighed by the energy gained via accommodating constant D-band. Surprisingly, we observe structure very similar to Galloway's constant angle model at $\Lambda > 27$ as seen in Figure 4.7c. In this frustrated twist phase, $\psi(r)$ has a sharply increasing twist for small distances from the fibril centre $r = 0$, but then flattens out to a constant twist angle for the mid region of the fibril before rapidly increasing in twist near $r = R$. Since this frustrated twist phase only occurs at larger Λ, a constant angle phase implies a strong coupling between the molecular orientation $\psi(r)$ and the D-band period. The longitudinal molecular strain is much smaller in the frustrated twist phase, as can be seen in Figures 4.11c and 4.11d where the maximum compression and tension are -0.3% and 0.1%.

4.4.2 Parameter estimation

In order to simplify our discussion of model predictions and how they might be consistent with experiment, it is necessary to constrain where possible what parameter values are reasonable. The most experimentally accessible parameter of our model is that of the inverse cholesteric pitch, $\tilde{q}$, which has been measured for chiral nematic collagen systems in vitro. This parameter enters our model in several places, but perhaps the most important (and implicit) of which is the definition of our dimensionless radius, $R = \tilde{R}/\tilde{q}$ (recall that all dimensional variables in this chapter will have a tilde over them). Experimental work on in vitro chiral nematic solutions of collagen estimate $\tilde{q} \in [0.1\pi\mu\mathrm{m}^{-1}, 4\pi\mu\mathrm{m}^{-1}]$ [122]. However, given that $\tilde{q}$ itself changes with e.g. pH, and fibril formation is controlled by pH, along with the understanding that $\tilde{q}$ is by definition measured in the chiral nematic state (and so its effective value in the fibril state has not been measured), it is entirely plausible to have $\tilde{q}$ values slightly outside (but probably the same order of magnitude as) the above range.

Considering a generous upper estimate of $\tilde{q} = 30\,\mu\mathrm{m}^{-1}$ for collagen systems in general, we can place upper bounds on our model R for both tendon ($R_{\mathrm{max}}^{\mathrm{tendon}}$) and cornea ($R_{\mathrm{max}}^{\mathrm{cornea}}$) fibrils. For tendon fibrils with maximum radius of $\tilde{R} \sim 200\,\mathrm{nm}$

observed in experiment, we have an upper limit on the allowed values of $R^{\text{tendon}}_{\text{max}} \sim 6$. From Figures 4.3 and 4.7a, we see that both linear and frustrated twist fibrils would still be accessible considering only this radial cut-off. For cornea fibrils with $\tilde{R} = 20\,\text{nm}$, we have a maximum cutoff of $R^{\text{cornea}}_{\text{max}} \sim 0.6$, and so in general the cornea ultrastructure is only accessible in the linear twist fibrils of our model if experimental measures of cholesteric pitch are within an order of magnitude accuracy. (We will explore values of q outside of this range to examine the ultrastructure and mechanical properties of frustrated twist cornea fibril in the next section.)

The other experimentally intuitive parameter in our model is the dimensional surface tension, $\tilde{\gamma}$. Ideally, we would use a surface tension measurement of collagen in e.g. phosphate-buffered saline solution, but unfortunately no such measurement exists to the authors' knowledge. To compensate for this in previous work, we imagined our collagen fibrils being surrounded by an isotropic collageneous solution, and so drew analogy to the surface tension between a nematic phase and its isotropic counterpart, which is on the order of $\tilde{\gamma} \sim 1 - 30\,\text{pN}\,\mu\text{m}^{-1}$ [128, 124, 127, 129]. Another reasonable substitute for the surface tension of collagen would be the surface tension of a protein droplet in a buffered solution. We draw on recent experimental work [152] which utilizes an optical trapping method to measure the surface tension of protein droplets in HEPES (buffer) solution, and finds surface tension values on the order of $\tilde{\gamma} \sim 20\,\text{pN}\,\mu\text{m}^{-1}$. Of course, our dimensionless surface tension $\gamma = \tilde{\gamma}/(\tilde{K}_{22}\tilde{q})$ also requires an estimate of $\tilde{K}_{22}$ and $\tilde{q}$. The former is $\tilde{K}_{22} \simeq 6\,\text{pN}$ as argued in previous work [4]. The upper bound on latter is mentioned above, but a single value will be determined more precisely through comparison between our model and the experimental phenomenology.

As mentioned in our methods section, we have held $K_{33} = 30$ and allowed $-1 \leq k_{24} \leq 1$. Our final two parameters, Λ and ω, will be used as fitting parameters when comparing to experiment, but their exact values for collagen systems are not clear. It seems reasonable to suspect that the coupling strength between D-band and twist, Λ, would be larger for fibrils which typically have small $\psi(r)$, as the presence of D-band reduces the energetic benefit of molecular twist. ω only has a significant effect the stability of our model at high strains (see supplementary Figure 4.24) and on the D-band amplitude δ (see supplementary Figure 4.22), and both of these effects are mainly qualitative. In particular, R, $\psi(r)$ and η are not strongly affected as long as $\omega \gtrsim 1$ as can be seen in (supplementary) section 4.6.

4.4.3 Comparison of $\psi(R)$, η, R, and mechanical properties with experiment

We began this chapter with the motivation to better understand the molecular orientation of tropocollagen molecules within fibrils. What we have shown in the results above, particularly in Figures 4.1 and 4.7c, is that there are two qualitatively different fibril ultra-structures that our model predicts, linear twist and frustrated twist fibrils. In order to narrow in on what ultrastructure best describes *in vivo* fibrils, we fit our model to the two extreme cases of well studied fibril ultrastructure: tendon and cornea fibrils. As shown in Table 2.1, the former typically have very small surface twist $\psi(R) \simeq 0.08\,\mathrm{rad}$ and have fairly wide distribution of radii, ranging from $\tilde{R} = 20\,\mathrm{nm}$ to as large as $\tilde{R} \simeq 200\,\mathrm{nm}$. The latter have $\psi(R) \simeq 0.28\,\mathrm{rad}$ and a much smaller range of radii, with a distribution that is sharply peaked near $\tilde{R} \simeq 15\,\mathrm{nm}$. In Figure 4.5, we see that surface twist in our model is strongly controlled by the dimensionless surface tension, γ. We might then expect that corneal fibrils have larger γ than their tendon counterparts, but to truly satisfy this claim we would also require smaller R. From Figure 4.3, it is clear that increasing γ also leads to much larger R, and so (as we will see below) γ alone cannot distinguish between tendon and cornea fibrils.

Cornea Fibrils

In Figure 4.12 we show the molecular twist $\psi(r)$ for two candidate cornea structures: a linear twist cornea fibril and a frustrated twist cornea fibril. We have selected parameters for these two phases (using the parameter space maps of Supplementary Figures 4.18 - 4.22 as a guide) in attempt to best recreate experimental observation of $\psi(R) \simeq 0.28\,\mathrm{rad}$ and $R \simeq 15\,\mathrm{nm}$ (see Table 2.1), though other parameter values near these values also fit well with experimentally observed ultrastructure. For the linear twist (blue dashed line), we choose parameters $\gamma = 0.15$, $k_{24} = 0.8$, $\Lambda = 0.9$, $\omega = 20$, and $\tilde{q} = 24\,\mathrm{\mu m}^{-1}$. Furthermore, with $\tilde{K}_{22} = 6\,\mathrm{pN}$ (consistent with Chapter 3) we obtain a dimensional surface tension of $\tilde{\gamma} = 22\,\mathrm{pN}\,\mathrm{\mu m}^{-1}$, which is in

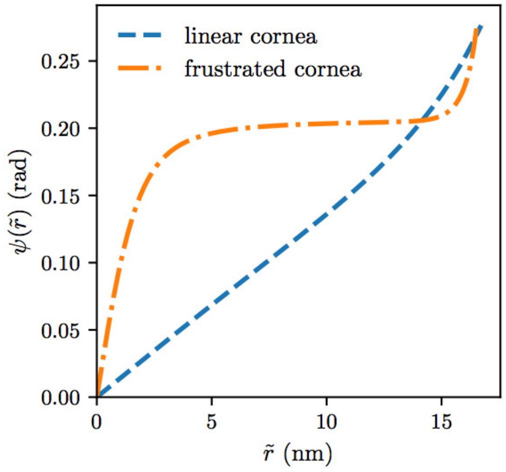

Figure 4.12: Cornea fibril local twist angle, $\psi(r)$, for a linearly twisted and frustrated twist fibril. For the linear twist (blue dashed line), the parameters chosen are $\gamma = 0.15$, $k_{24} = 0.8$, $\Lambda = 0.9$, $\omega = 20$, and $\tilde{q} = 24\,\mu\mathrm{m}^{-1}$. For the frustrated twist (orange dash dotted line), $\gamma = 0.15$, $k_{24} = 0.8$, $\Lambda = 1.5$,$\omega = 20$, and $\tilde{q} = 150\,\mu\mathrm{m}^{-1}$. (Note that $\tilde{q}$ for the frustrated cornea twist is five times larger than the expected upper bound.)

very good agreement with the expected range of surface tension for protein droplets in buffer solution [152]. For the frustrated twist (orange dash dotted line), we set $\gamma = 0.15$, $k_{24} = 0.8$, $\Lambda = 1.5$,$\omega = 20$, and $\tilde{q} = 150\,\mu\mathrm{m}^{-1}$ (note this $\tilde{q}$ value is larger than expected). With this, we obtain a surface tension of $\tilde{\gamma} = 135\,\mathrm{pN}\,\mu\mathrm{m}^{-1}$. Both structures have surface twist of $\psi(R) = 0.28\,\mathrm{rad}$ and dimensional radius $\tilde{R} = 16\,\mathrm{nm}$, well in agreement with experimental values.

In Figure 4.9, we showed that our model allows us to compute the stress strain behaviour of a single fibril, as well as the orientation of tropocollagen within and on the fibril surface. We can therefore go a step further with matching our model fibril to experiment thanks to a recent study by Bell et al. [6] which measured the twist of tropocollagen within strained fibrils using X-ray diffraction. In this study, mechanical response of cornea fibrils with twist of $\psi(R) = 16^\circ \simeq 0.28\,\mathrm{rad}$ and typical radii of $\tilde{R} \sim 17\,\mathrm{nm}$ at zero stress. In general we would begin by fitting our model to reproduce $\psi(R)$ and $\tilde{R}$, but we have already found this ultrastructure in Figure 4.12 and so use the parameters listed in the caption there. In Figure 4.13, we compare our model to their experimental results by applying a strain along the fibril axis, while

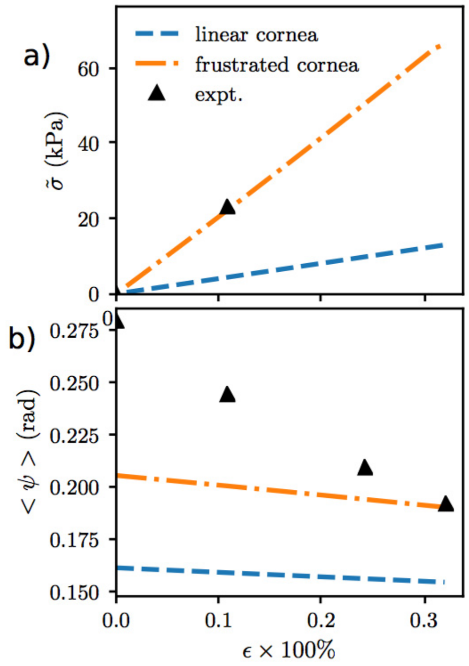

Figure 4.13: Mechanical response of straining both linear and frustrated cornea fibrils. The two curves correspond to our model predictions, under the constraint that the fibrils are incompressible. We also include data from a recent experiment [6] in which the twist and stress response of cornea tissue is measured under strain (see main text for discussion of fit). For the linear twist (blue dashed line), the parameters chosen are $\gamma = 0.15$, $k_{24} = 0.8$, $\Lambda = 0.9$, $\omega = 20$, and $\tilde{q} = 24\,\mu\text{m}^{-1}$. For the frustrated twist (orange dash dotted line), $\gamma = 0.15$, $k_{24} = -0.3$, $\Lambda = 1.5$, $\omega = 20$, and $\tilde{q} = 150\,\mu\text{m}^{-1}$. The calculated (zero strain) D-band periods for the linear and frustrated phases twist phases shown are $2\pi/\eta_\text{l} = 0.985$ and $2\pi/\eta_\text{f} = 0.979$, respectively.

holding the total fibril volume constant as stated in section 4.3.4. Note that the twist angle we present here is the volume averaged twist

$$< \psi >= \frac{2}{R^2}\int_0^R r\psi(r)dr, \tag{4.9}$$

to remain consistent with experiment.

In Figure 4.13, we measure a Young's modulus, Y, of $Y = 4.04\,\text{MPa}$ in the linear twist cornea fibril, and $Y = 20.4\,\text{MPa}$ in the frustrated twist cornea fibril.The lone study (to the authors' knowledge) of in vitro fibril mechanics reports $Y \sim 30\,\text{MPa}$ [33]; our model is within an order of magnitude of this value. However, the maximal strain we are able to apply to our model fibrils is typically in the kPa range, which is approximately 100 times smaller than what is observed in experiment on in vitro fibrils

[33]. Possible explanations of this difference between our model and experiment may be related to non-equilibrium effects[6], or slippage of fibrils under strain in experiment [27]. More studies which investigate in vitro fibril mechanics would be ideal to further support our calculations of Y.

Clearly, our model does not quantitatively agree with the experimental data of Bell et al, regardless of whether the underlying structure is linear or frustrated. There are two likely causes of this discrepancy. Either the experiment does not accurately represent the physics we are simulating in our model, or our model is too simple. The first cause is likely partly to blame, as the experimental data we compare to is collected by pulling on cornea tissue (*vs* single fibrils in our work), which means we cannot expect our stress strain curves to fit well with their results. This tissue is composed of cross-linked fibrils, which we do not account for in our model. In addition, only the first three strain points in their experiment (0%, 1.4%, and 2.8%) correspond to D-band strain as their main deformation mechanism. As strain increases, fibrils begin to slip with respect to each other, and so the dominate mechanism of mechanical deformation becomes inter-fibrillar. Our model only considers mechanisms of intra-fibrillar deformation, and so we can no longer expect our model to be compatible with experimental observation at high tissue strain. However, this does not fully explain the large difference between our model and experiment at small strain.

The second possible difference between our model and the experimental results is due to our assumption of a constant D-band amplitude within the fibril. It has been suggested that fibrils may have a region of lower density or disorder in their centre (near $r = 0$) [153, 99]. If we modified our model to include this liquid-like core, the constraint that $\psi(r) = 0$ at the core could be relaxed, as no ordered tropocollagen would be present near $r = 0$. This region of the fibril could be most simply modelled by changing the integration bounds of our free energy in eqn 4.6 to start at non-zero r. Naively, we would expect that removing twist near $r = 0$ in the $\psi(r)$ curves (see e.g. Figure 4.12) would increase the average $< \psi >$, and at least increase our model $< \psi >$ to be closer to that of the experimental data in Figure 4.13b. We will investigate this possibility in the near future.

---

[6] In this case, either experimentalists would have to decrease their strain rates, or we would have to include dynamics within our intrinsically equilibrium model.

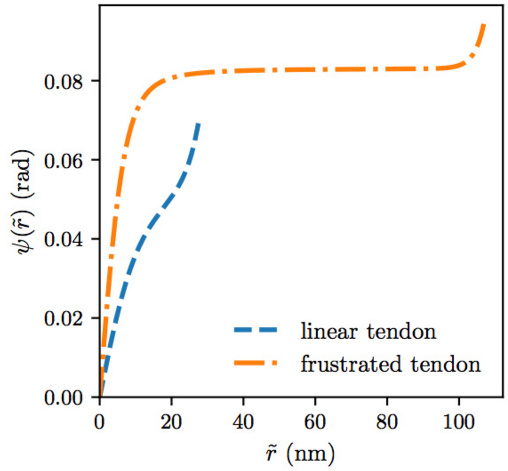

Figure 4.14: Tendon fibril local twist angle, $\psi(r)$, for a linearly twisted and frustrated twist fibril. Both linear (blue dashed curve) and frustrated (orange dash dotted curve) are at coexistence (equal energy E), with $\gamma = 0.04$, $k_{24} = 0.5$, $\Lambda = 600$, $\omega = 20$, and $\tilde{q} = 4\,\mu\text{m}^{-1}$.

Tendon Fibrils

In contrast to cornea fibrils mentioned above, fibrils with low surface twist and large radii (i.e. tendon) are accessible at a wide range of Λ values. All of these fibril types, regardless of whether they are linear or frustrated twist, appear to require relatively small γ values as can be seen in (supplementary) Figure 4.20. We show in Figure 4.14 two possibilities of tendon fibril ultrastructure, which are on the coexistence line at $\gamma = 0.04$, $k_{24} = 0.5$, $\Lambda = 600$, and $\omega = 20$. The linear twist tendon fibril (blue dashed curve) has a radius of $\tilde{R} = 28\,\text{nm}$ and surface twist of $\psi(R) = 0.069\,\text{rad}$. The frustrated twist tendon fibril (orange dash-dotted curve) has $\tilde{R} = 107\,\text{nm}$ and $\psi(R) = 0.094\,\text{rad}$. We have taken $\tilde{q} = 4\,\mu\text{m}^{-1}$ to cast r into dimensional units. With $\tilde{K}_{22} = 6\,\text{pN}$, we obtain a dimensional surface tension of $\tilde{\gamma} = 0.96\,\text{pN}\,\mu\text{m}^{-1}$, which is on the lower range of expected surface tension [7]. Since both the frustrated tendon and linear tendon are at equal E with the same parameter values (i.e. the same fibrillar environment), it is possible to in fact have coexistence of two different equilibrium

---

[7] We quote dimensional results (e.g. $\tilde{\gamma}$ here and Young's modulus below) to enough significant figures that the model results may be reproduced as shown in this thesis. In comparing our results to experiment, we would expect our model predictions to be accurate to two significant figures (at most) due to uncertainties in what the true parameter values of our model are for collagen.

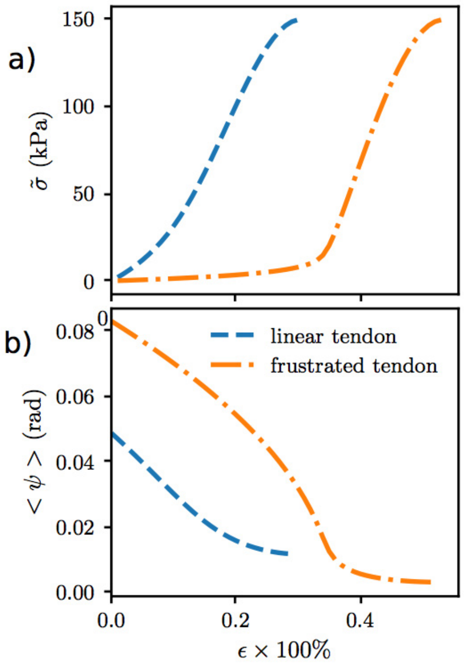

Figure 4.15: Mechanical and structural properties of linear (blue dashed) and frustrated (orange dash dotted) tendon fibrils. Both linear and frustrated fibrils are at coexistence (equal energy E), with $\gamma = 0.04$, $k_{24} = 0.5$, $\Lambda = 600$, $\omega = 20$, and $\tilde{q} = 4\,\mu\mathrm{m}^{-1}$. (a) Stress and (b) volume averaged twist of fibrils *vs* D-band strain. Note that stress strain curves are cut off when strain σ reaches a maximum (where the D-band amplitude $\delta \to 0$ and so D-band strain is no longer well defined). The calculated (zero strain) D-band periods for the linear and frustrated twist phases shown are $2\pi/\eta_{\mathrm{l}} = 0.999$ and $2\pi/\eta_{\mathrm{f}} = 0.997$

structures at in the same tendon tissue. This may lend a simple, equilibrium explanation to the bimodal distributions of radii that have been observed in tendon [68]. Thus, we would expect that the larger tendon fibrils in bimodal distributions would have frustrated twist molecular orientation, whereas the smaller fibrils would be linearly twisted. We also note that it is possible to obtain both linear and frustrated twist tendon fibrils away from the coexistence point through variation in Λ and $\tilde{q}$, and so tissue without this bimodal behaviour is still consistent with our results. In this case, we would expect the larger the fibril, the more likely it is to be in the frustrated twist phase.

In Figure 4.15, we show the result of straining the linear (blue dashed) and frustrated (orange dash dotted) tendon fibrils. The stress strain curve for both fibril types appear to have two separate regimes of strain. The first regime consists of a linear relationship between stress and strain, with Young's moduli of $Y = 24.7\,\mathrm{MPa}$ and $Y = 1.3\,\mathrm{MPa}$ for the linear and frustrated twist fibrils, respectively. The second regime is again fairly linear, with larger Young's moduli of $Y = 77.5\,\mathrm{MPa}$ and $Y = 100.1\,\mathrm{MPa}$. These values are again well within the range of experimentally observed moduli for uncross-linked fibrils [154]. We have cut off the strain curves at the region where the slope of the stress becomes negative, which can be thought of as the largest strain that can be applied to our model without melting the D-band. We note that near this breaking point, the $<\psi>$ behaviour begins to flatten out and actually begins increasing again, as in Figure 4.9c. This is likely occurring due to the rapidly decreasing D-band amplitude δ (not shown, but see Figure 4.9b), as $\delta \sim 0$ removes the coupling between D-band and twist, allowing the twist to increase rapidly with strain.

Though many studies exist which look at the large strain behaviour of tendon tissue, to the author's knowledge, no studies analogous to the studied performed by Bell et al. [6] have been published using tendon (*vs* cornea) fibrils. In particular, it would be interesting to see whether the small strain behaviour of uncross-linked, single fibril stress strain curves would reproduce the two linear regime curves shown in Figure 4.15.

Parameter control

When comparing the parameter values of cornea and tendon fibrils in our model, we see that three parameters are much different between the two: $\tilde{q}$, γ, and Λ. Many factors relating to differences in fibril environment could attribute to these differences in parameter values. For instance, differences in pH and concentration have been shown to modify the pitch of the tropocollagen chiral nematic phase in vitro [122], which would provide a mechanism for altering $\tilde{q}$ between the two phases. Further differences could be related to hydration level of the fibrils, proteoglycans and even small temperature differences in different anatomical locations. Internally, the fibrils themselves have slightly different composition of tropocollagen molecules [16], which would also provide variation in our parameters. The largest difference between cornea and tendon appears to be in the parameter Λ with $\Lambda_c = 0.9$ and $\Lambda_t = 600$, respectively. However, when considering that Λ is a dimensionless number with an inverse square proportionality to $\tilde{q}$ (see list of dimensionless parameters in the supplementary materials). It is therefore more prudent to look at how the ratio of $\tilde{\Lambda} \propto \Lambda\tilde{q}^2$ differs in the two cases. We find that $\tilde{\Lambda}_t/\tilde{\Lambda}_c = 18.5 \simeq 20$.

The deviation of the above $\tilde{\Lambda}$ ratio between cornea and tendon highlights the importance of twist in the mechanical properties of our model. If we hold our twist $\psi(r) = 0$, all radius dependence in eqn 4.6 drops out and we are left with a 1D phase field crystal model, with $R_{eq} \to \infty$ from the surface tension in eqn 4.6. This zero twist free energy becomes

$$\begin{aligned} E(\epsilon, \eta_{eq}, \delta_{eq}) =& \frac{\Lambda\delta_{eq}^2}{4}\left(4\pi^2 - \frac{\eta_{eq}^2}{(1+\epsilon)^2}\right)^2 \\ &+ \frac{\omega\delta_{eq}^2}{2}\left(\frac{3}{4}\delta_{eq}^2 - 1\right) \end{aligned} \tag{4.10}$$

where $\eta_{eq} = 2\pi$ is the equilibrium inverse D-band period of the untwisted state, and $\delta_{eq} = 1$ is the equilibrium D-band amplitude. Expanding this energy in terms of the strain ϵ,

$$E(\epsilon) = E(0) + \frac{1}{2}\frac{\partial^2 E}{\partial\epsilon^2}\epsilon^2 + \mathcal{O}(\epsilon^3), \tag{4.11}$$

it is clear that at small strains, $Y = \partial^2 E/\partial\epsilon^2$, and so $Y = 32/3\pi^4\Lambda$. If we write this in dimensional parameters, we find that the ratio of Young's moduli of cornea fibrils to tendon fibrils is expected to be $\tilde{Y}_c/\tilde{Y}_t = 18.5 \simeq 20$. However, from our discussions above, with our model we find similar values of $\tilde{Y}$ in both cornea and tendon, even though the zero twist model predicts cornea fibrils to have 18 times the modulus of tendon fibrils. Furthermore, we find that two separate $\tilde{Y}$ can coexistence for the same parameter values in tendon. This feature is not possible without the inclusion of twist in our model.

Other fibril ultrastructure

We have focused on cornea and tendon fibril ultrastructure as the two extremes of collagen fibril ultrastructure, and have showcased the phase field collagen fibril model's ability to capture both with a reasonable range of parameter values. Other *in vivo* fibrils have surface twist values within the two ranges presented in this chapter (between 0.07 rad and 0.28 rad), and so should be easily captured within the model, as a continuum of $\psi(R)$ can be attained through parameter variation. Thus our model can be used as to inform predictions on the inner molecular orientation of tropocollagen in the fibril or mechanical properties. In vitro, the projective coupling of the D-band and surface twist indicates that $\psi(R)$ is small, even if the fibrils are reconstructed with solubilized tropocollagen extracted from large surface twist fibrils *in vivo* [138]. The dominant mechanism for decreasing $\psi(R)$ in our model is through decreasing surface tension $\tilde{\gamma}$. This may provide an explanation for the reduced $\psi(R)$ if the surface tension of the in vitro solution is less than the *in vivo* environment. Unfortunately, no measurements of fibril surface tension in vitro or *in vivo* exist in the literature to validate or reject this hypothesis.

4.5 Summary

In this work, we have constructed a new model of collagen fibril structure which couples the local, radial orientation of tropocollagen molecules within the fibril, $\psi(r)$, to the D-band (density) modulations along the fibril axis. Our model predicts two distinct fibril phases: i) a linear twist phase, where the molecules twist around the fibril centre like strands in a rope, and ii) a frustrated twist phase, where the majority of the molecules in the fibril remain at a constant twist angle with respect to fibril axis. Structures similar to both of these phases have been proposed in the literature, but until now no predictive model has been constructed in attempt to confirm that either of these structures might be energetically favourable. In particular, the linear twist structure has been rejected under the assumption that any non-constant twist is incompatible with a constant D-band throughout the fibril cross-section, but we show here that only a small strain on individual molecules would be required to accommodate a constant D-band, as our linear twist phase does indeed have D-band structure. By constructing a free energy with both local twist $\psi(r)$ and D-band modulations, we are also able to calculate the radius of the fibrils, R, the period of the fibril D-band, $2\pi/\eta$, and the amplitude of the fibril D-band, δ for both linear and frustrated twist fibrils. Furthermore, we show that our model can be used to probe the mechanical response of our model fibrils by straining the D-band period and measuring the resulting stress on the fibril.

4.6 Supplemental Figures

In this section we present figures which are not integral to the understanding of the PFCF model, but still bear importance in understanding how our predictions are affected by parameter space.

4.6.1 γ ***vs*** k_{24} phase diagrams

Figures 4.16 and 4.17 demonstrate the dependence of η and δ on (γ, k_{24}) variations similar to the main text Figures 4.2, 4.3, and 4.5 for E, R, and $\psi(R)$, respectively.

4.6.2 Λ ***vs*** ω phase diagram sketches

Figures 4.18, 4.19, 4.20, 4.21, and 4.22 demonstrate the dependence of our model results on Λ and ω. These figures consist of nine contour plots, and each contour plot additionally has two sub-plots (above and to the right). The sub-plot to the right of each contour plot contains a dotted, magenta curve which shows the relevant observable (x-axis) vs ω (y-axis) at a constant $\Lambda = 10$. Notably, upon examining all sub-plots to the right of the contour plots in Figures 4.19, 4.21, and 4.19, it is clear that R, η, and $\psi(R)$ are all completely independent of ω, respectively, as their corresponding curves are constant with ω. E is linear in ω as expected from eqn 4.6 in the main text, and δ only changes near $\omega \sim 1$. The sub-plot above each contour plot contains a dotted, orange curve which shows the relevant observable (y-axis) vs Λ (x-axis) at constant ω. The value of ω is chosen to accommodate non-converged solutions present in the contour plots (white space – discussed below); for each contour plot the y-tick of the chosen ω value is magnified and coloured in orange. From both the contour plots and the constant ω sub-plots, none of the observables are independent of Λ, consistent with the main text.

There are several regions in Figures 4.18 - 4.22 with numerical artefacts (white space on the contour plots); their presence is due to the large computational cost of scanning parameter space and the difficulty of automatically setting efficient initial guesses for the optimization problem. At first glance, these artefacts may appear problematic. However, the quantitative predictions of our model at large Λ are independent of different ω values (aside from E and δ, as noted above). This independence

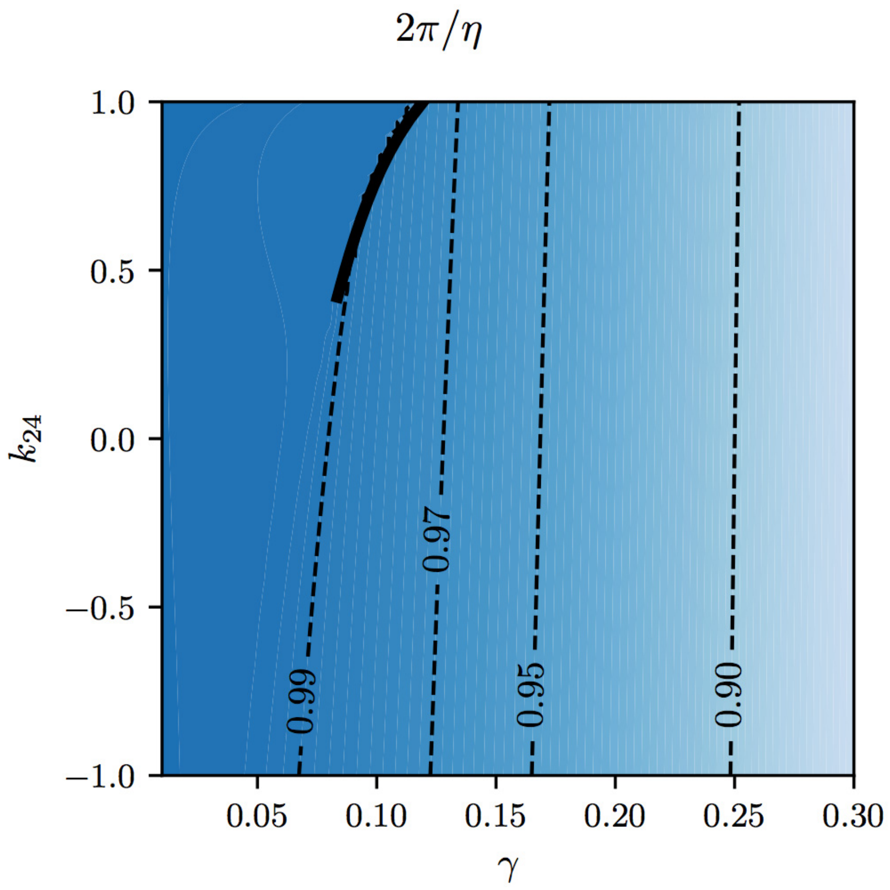

Figure 4.16: Contours of constant D-band period $2\pi/\eta$ in the k_{24} *vs* γ plane. The thick black line corresponds to the coexistence line between linear and frustrated twist fibril phases. A discontinuous jump in the period occurs across the coexistence line as γ is increased at constant k_{24}. γ strongly controls the period of the D-band, with larger γ corresponding to smaller period. The maximum value of $2\pi/\eta \sim 1$ occurs at small γ. Note that to determine the true (dimensional) period of the fibrils, one only needs to multiply the observed value of the contours indicated by $\tilde{d}_{\parallel} = 67\,\text{nm}$. We hold $\Lambda = 27$ and $\omega = 10$ constant.

of ω is easily seen in slices of our model observables at constant $\Lambda = 10$ in Figures 4.19 - 4.20 (these dotted, magenta slices are shown on the right side of each of the nine contour plots within each figure), which are constant in each observable at $\omega \gtrsim 1$. Therefore, little information is lost in these numerical artefacts at large Λ as long as at least one ω value has converged. For example, from the R data observed at (Λ,ω)=(300,25) with $(\gamma,\ k_{24})$=(0.04,0.1) in Figure 4.19, we see that $R \simeq 0.6$ from the constant $\omega = 25$ slice (dotted orange curve above the contour plot), and so we

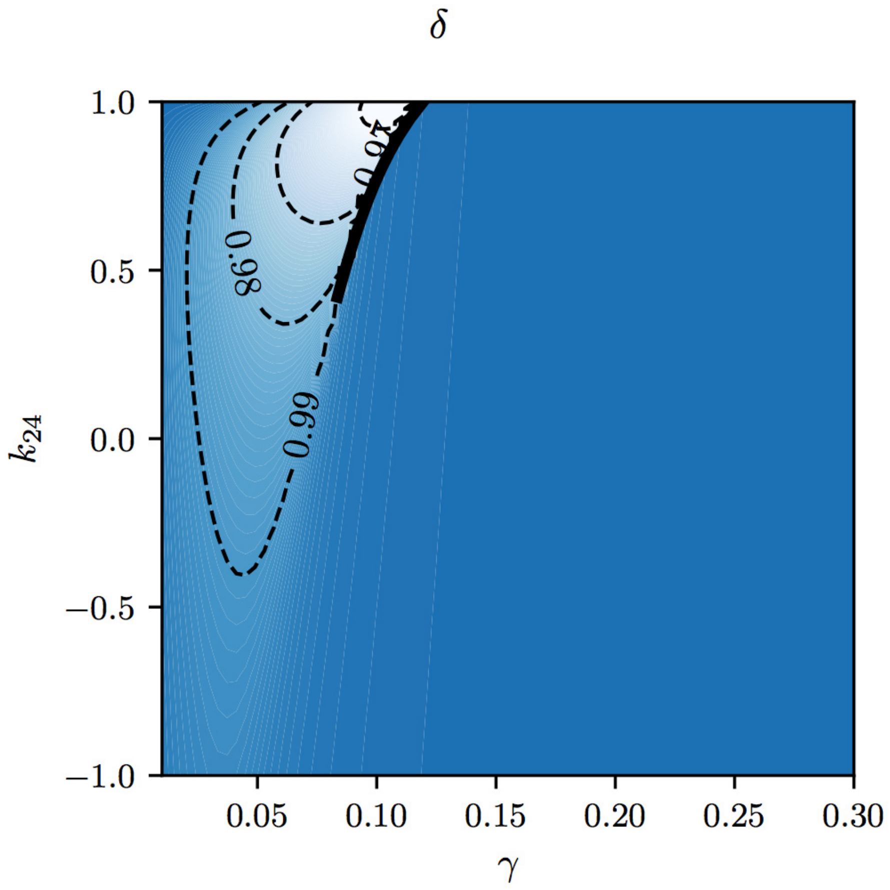

Figure 4.17: Contours of D-band amplitude δ in the k_{24} *vs* γ plane. The thick black line corresponds to the coexistence line between linear and frustrated twist fibril phases. A discontinuous jump in δ occurs across the coexistence line as γ is increased at constant k_{24}. Notably, δ does not vary much with either parameter, with its minimum value of $\delta \sim 0.94$ occurring near the left side of the coexistence line. We hold $\Lambda = 27$ and $\omega = 10$ constant.

may infer that at $\Lambda = 300$, (γ, k_{24})=(0.04,0.1), $R \simeq 0.6$ independent of ω. Therefore, have included these figures not as a main result of our model, but as a helpful guide to how our model depends on parameter space. For this reason, it is appropriate to include these in the thesis as a demonstration of how we were able to effectively select parameter values in the main sections of this chapter (e.g. in choosing parameters for cornea and tendon fibrils in section 4.4.3). To eliminate all numerical artefacts would require manually inputting initial guesses for the minimization of E for every grid point which is white in the figures. Given the computational cost of doing this for little important information, we have decided to forego this calculation.

4.6.3 Mechanical response at different parameters

Figure 4.23 shows the mechanical response of our PFCF model with (γ, k_{24})=(0.12,1.0) at different Λ (analogous to Figure 4.9 in the main text). Figure 4.24 examines the relationship between ω and the maximum stress that can be applied in our system before the D-band amplitude $\delta \to 0$.

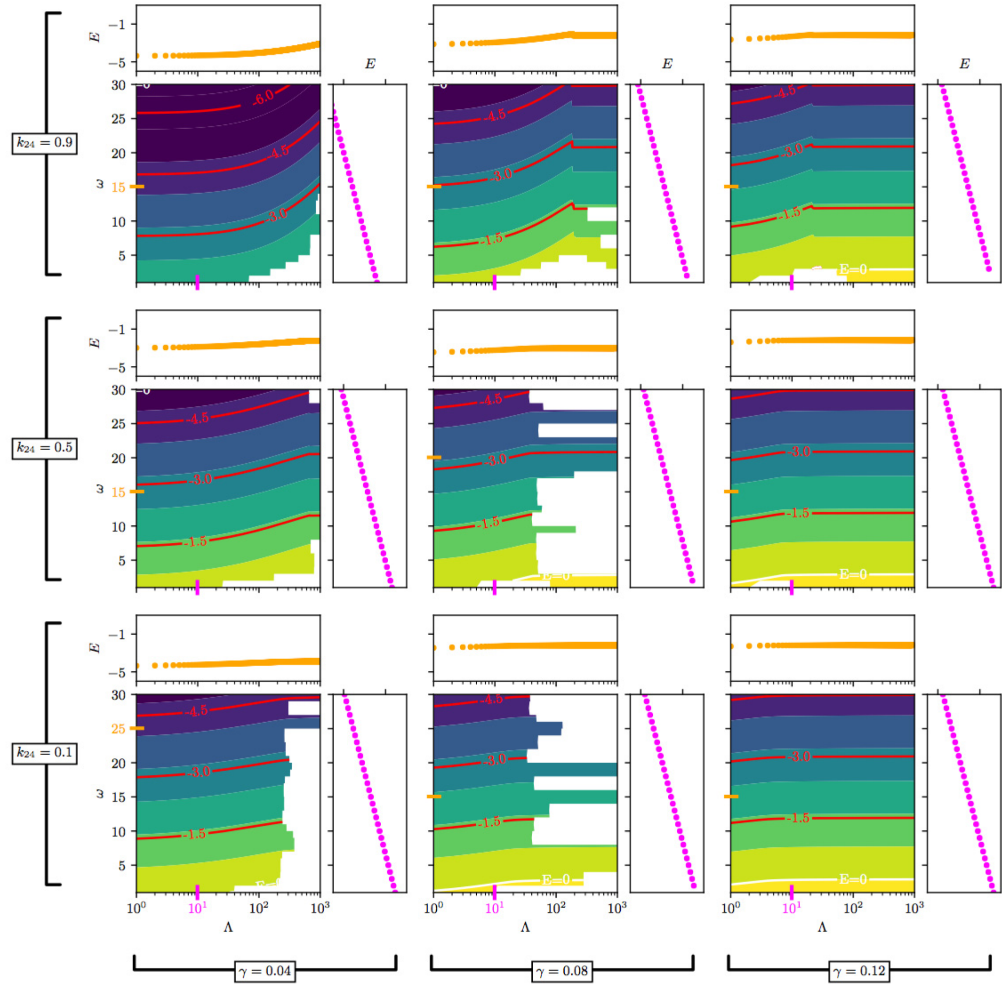

Figure 4.18: Equilibrium free energy per unit volume E contours in the ω *vs* Λ plane (main coloured plots), at several values of γ and k_{24}. Each contour plot is accompanied by a constant ω slice (above, with the chosen ω value highlighted in orange along the y-axis of the corresponding main contour plot) and a constant $\Lambda = 10$ slice (to the right). White space in the contour plots indicate non-converged solutions (in the computational time allotted) due to poor initial conditions in the ω-Λ parameter space scanning algorithm (see text for discussion).

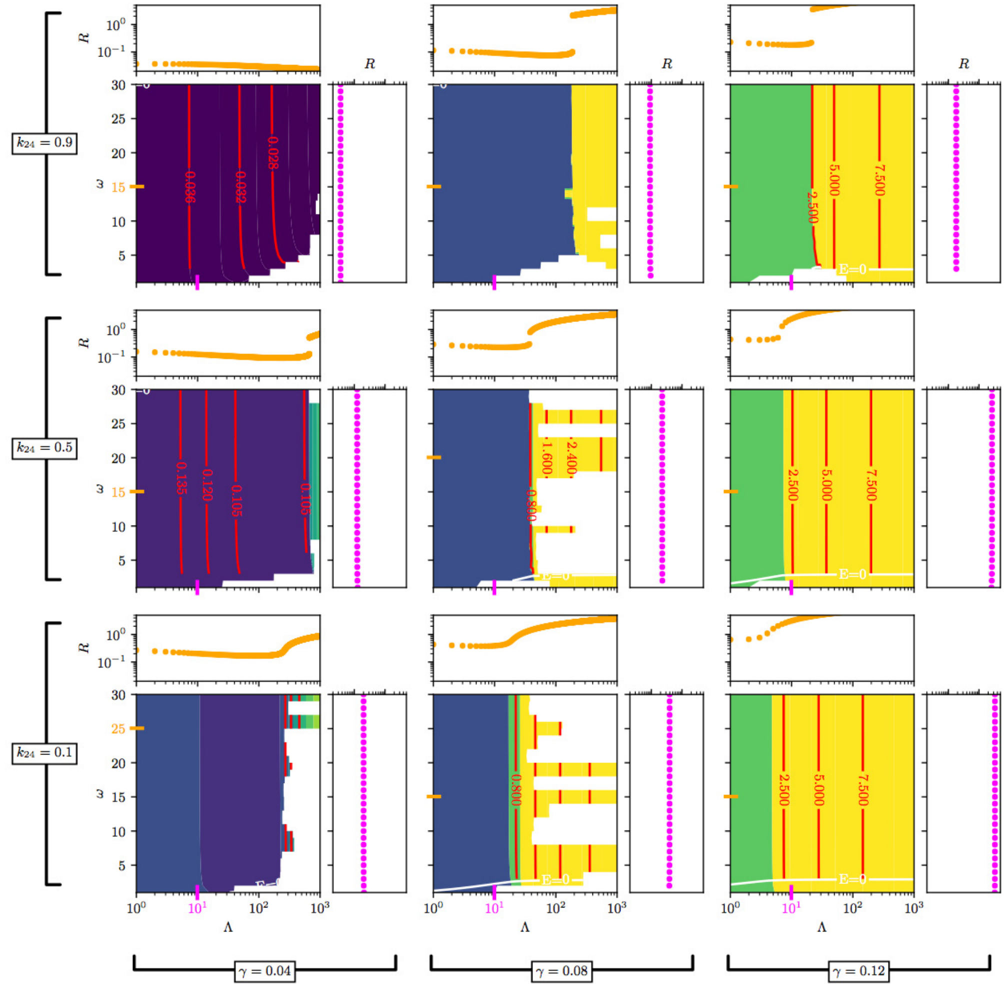

Figure 4.19: Equilibrium fibril radius R contours in the ω *vs* Λ plane, at several values of γ and k_{24}. Each contour plot is accompanied by a constant ω slice (above, with the chosen ω value highlighted in orange along the y-axis of the corresponding main contour plot) and a constant $\Lambda = 10$ slice (to the right). Consistent with the main text, at large γ and k_{24} values, a radial discontinuity emerges with increasing Λ. White space in the contour plots indicate non-converged solutions due to poor initial conditions in the ω-Λ parameter space scanning algorithm (see text for discussion).

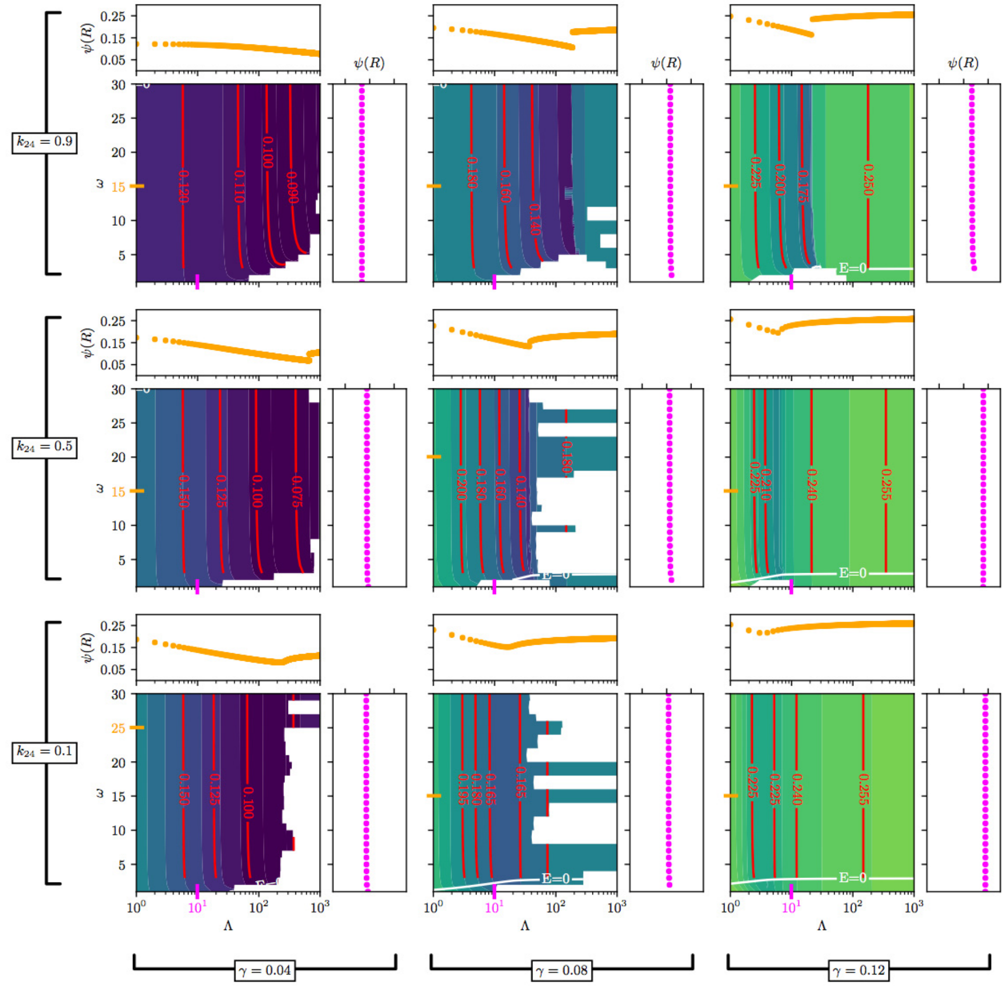

Figure 4.20: Equilibrium surface twist $\psi(R)$ contours in the ω *vs* Λ plane, at several values of γ and k_{24}. Each contour plot is accompanied by a constant $\omega = 15$ slice (above) and a constant $\Lambda = 10$ slice (to the right). Consistent with the main text, at large γ and k_{24} values, a $\psi(R)$ discontinuity emerges with increasing Λ. White space in the contour plots indicate non-converged solutions due to poor initial conditions in the ω-Λ parameter space scanning algorithm (see text for discussion).

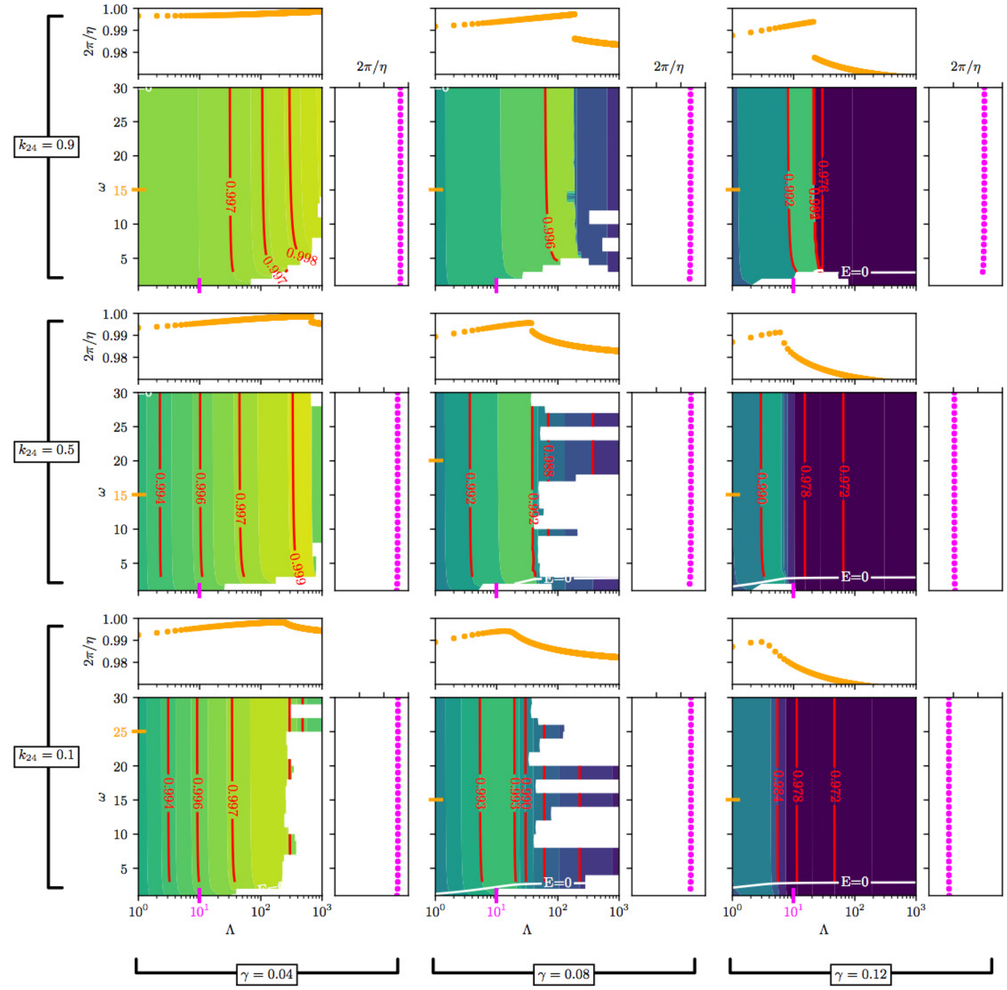

Figure 4.21: Equilibrium D-band period $2\pi/\eta$ contours in the ω *vs* Λ plane, at several values of γ and k_{24}. Each contour plot is accompanied by a constant $\omega = 15$ slice (above) and a constant $\Lambda = 10$ slice (to the right). Consistent with the main text, at large γ and k_{24} values, a discontinuity in the D-band period emerges with increasing Λ. White space in the contour plots indicate non-converged solutions due to poor initial conditions in the ω-Λ parameter space scanning algorithm (see text for discussion).

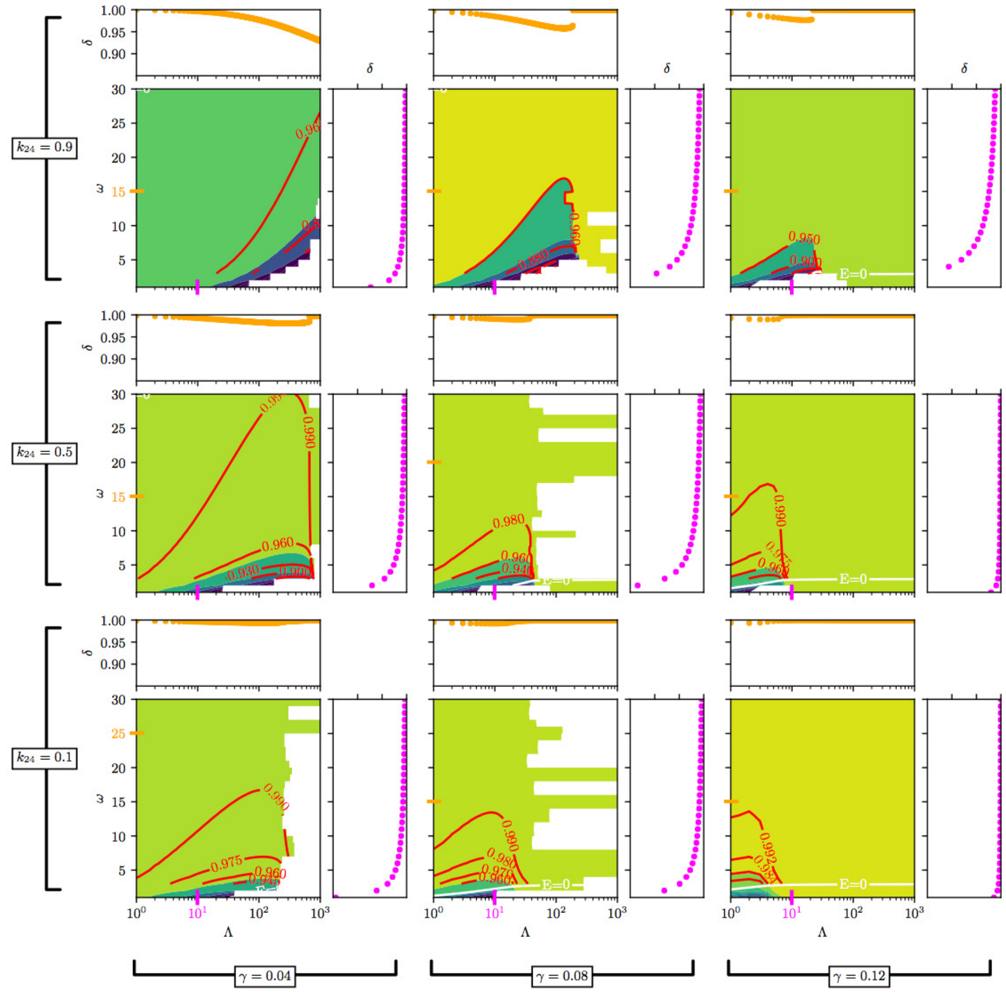

Figure 4.22: Equilibrium D-band amplitude δ contours in the ω *vs* Λ plane, at several values of γ and k_{24}. Each contour plot is accompanied by a constant $\omega = 15$ slice (above) and a constant $\Lambda = 10$ slice (to the right). Consistent with the main text, at large γ and k_{24} values, a discontinuity in the D-band amplitude emerges with increasing Λ. White space in the contour plots indicate non-converged solutions due to poor initial conditions in the ω-Λ parameter space scanning algorithm (see text for discussion).

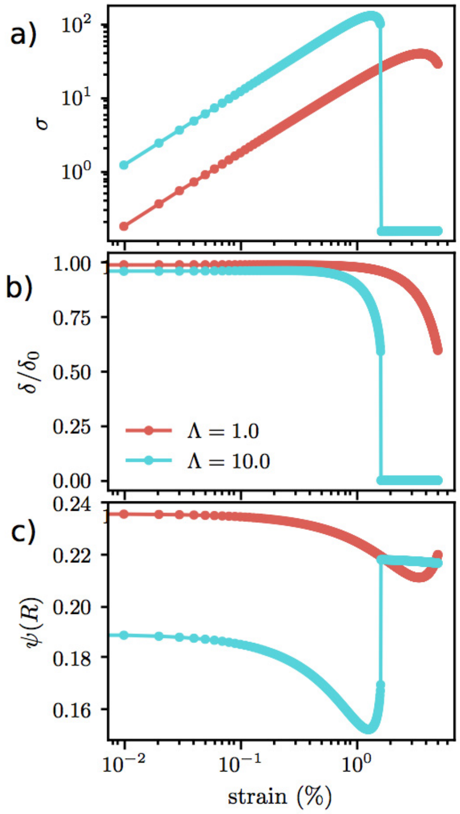

Figure 4.23: Mechanical and structural response of straining the collagen fibril D-band for an incompressible fibril. (a) Stress, $\sigma = dE/d\epsilon$, *vs* strain (ϵ) percentage for $\Lambda = 1$ (red curve) and $\Lambda = 10$ (cyan curve), all at $\gamma = 0.12$, $k_{24} = 1.0$, and $\omega = 10$. The stress increases linearly in all cases for small strain. As Λ increases, the stress becomes non-linear at large strain. Notably, the initial slope in the stress (i.e. the Young's modulus) does not monotonically increase with strain. (b) D-band (density) amplitude δ *vs* strain percentage. At small strain, the amplitude remains close to its equilibrium ($\epsilon = 0$ value), while at large strain δ decreases rapidly to zero ($\delta = 0$ corresponds to a fibril with no D-band). (c) Surface twist $\psi(R)$ *vs* strain percentage. $\psi(R)$ decreases with increasing strain initially, and increases slightly before the D-band disappears ($\delta = 0$) at the discontinuity in $\psi(R)$.

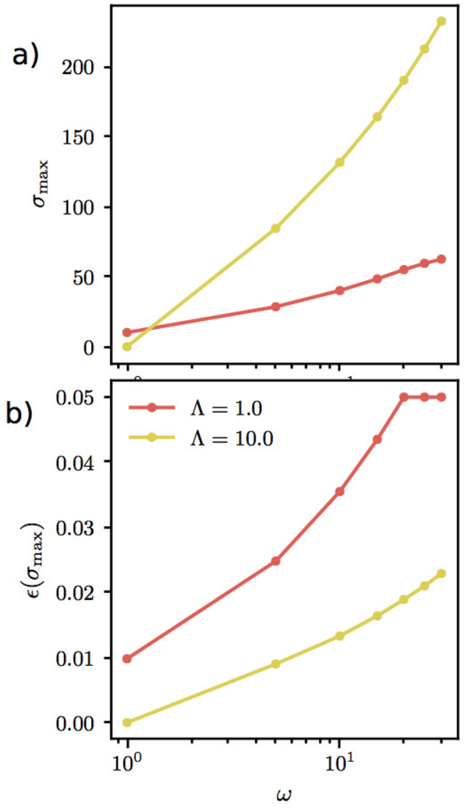

Figure 4.24: Maximum stress and corresponding strain values *vs* ω for $\Lambda = 1$ (red curve) and $\Lambda = 10$ (tan curve). Dots are data points, and the lines connecting the dots are a guide to the eye. Maximum stress is defined as the maximum value of stress attained at any strain ϵ. This maximum stress is not necessarily at the largest strain, indicating non-monotonic behaviour in the stress strain curves. This is not physical and is due to the fibril D-band "breaking" (i.e. $\delta \to 0$, and so we use the maximum stress as the upper value of strain we can apply to our model without breaking the fibril. Note that the upper bound of strain allowed is $\epsilon < 0.05$, which is why the $\Lambda = 1.0$ curve flattens out at this value in (b). Other parameters in this calculation are $\gamma = 0.12$ and $k_{24} = 1$.

Chapter 5

Conclusion and Future Outlook

5.1 Summary of results

In this thesis we have developed quantitative, predictive models of collagen fibril structure and its mechanical properties by constructing fibril free energies using the framework of statistical mechanics and soft matter physics. We minimize these free energies using numerical techniques, and determine the resulting collagen fibril structure.

In Chapter 3, we used a purely liquid crystalline model of collagen fibril structure to generate fibril structure which is consistent with those observed in experiment. We found that both radius, R, and tropocollagen twist, $\psi(r)$, generated from this model were consistent with experiment. Model parameters could be tuned specifically to capture different fibril structures, and in particular the two extreme cases of cornea fibrils (small, uniform R and large surface twist, $\psi(R)$) and tendon fibrils (poly-disperse R and small $\psi(R)$) could be generated from this model. Furthermore, this model demonstrated that the orientation of molecules within the fibril is well represented by a mildly non-linear form of $\psi(r)$, providing evidence for the constant pitch model of fibril structure proposed in the literature [48]. However, our model neglected the axial D-banding of collagen fibrils, and so was not a complete model of fibril structure. Furthermore, we suspected that D-band structure would couple to the $\psi(r)$ field, and so our results may be significantly altered with the addition of D-band structure.

In Chapter 4, we included D-band structure in our model through generalizing the phase-field-crystal formalism [5], and found using geometrical arguments that it is coupled to the molecular orientation field $\psi(r)$ within the fibril. While our model was still able to capture the R and $\psi(R)$ values observed in the literature, the coupling between D-band and $\psi(r)$ introduced a new phase of fibril structure, which we denote the frustrated twist phase (the original phase from both this chapter and the previous being named the linear twist phase). The differences between the linear and frustrated twist phases lie in the form of $\psi(r)$, which in the former case is linear, while in the latter has an almost piece-wise $\psi(r)$, which is linear near the fibril centre $r = 0$ and the fibril surface $r = R$, but has zero slope in between the centre and surface. We characterized the phase transition between these two phases, and determined the stress-strain behaviour in both cases. We found that the Young's moduli Y were within the experimental range of suitable (non-cross-linked, hydrated) fibrils. We found that the change in the average orientation of molecules within the fibril *vs* applied D-band strain did not match up with experiment [6]. The reason for this discrepancy is an inspiration for further modifications to the collagen fibril models presented in this thesis.

5.2 Future outlook

Though the models presented in this thesis show promise in their predictions of collagen fibril structure and mechanical properties, there are two modifications that can be made (either together or separately) to this model which I think will enhance it greatly.

The first modification of this model would be to introduce cross-links into the system via an additional term in the collagen fibril free energy. Since the models of this thesis are partially based off of liquid crystal theory, it may be possible to use the rapidly developing field of liquid crystal elastomers [31] which considers ordering of cross-linked polymer liquid crystals and their corresponding mechanical properties. If this cross-linking term were to be naively added to our existing phase field collagen fibril model, it may enable us to better predict the mechanical properties of cross-linked fibrils, and perhaps could explain the discrepancy we find between our model and the experimental predictions [6] of average molecular orientation under strain (mentioned in the above section). Current theories of liquid crystal elastomers mainly consider rod-like molecules connected by Gaussian polymer cross-links; a more sophisticated iteration of the theory may be required depending on the nature of the collagen fibril cross-links.

Another possibility for introducing cross-links might be in an ad-hoc manner, inspired by the works of Grason [82] which are discussed in Chapter 2. Collagen fibrils show radial (hexagonal) crystalline packing only at certain points along the axial (z) D-band period where cross-links are most abundant [55]. If we were to assume that hexagonal packing is a proxy for cross-linking, it might be reasonable to introduce z-dependent elastic coefficients [83] and so a hexagonal packing of fibrils only at certain fibril locations. This would incorporate the differences in disorder radially at different z into our model. Currently, the maximum strain attainable in our model is as little as %.4 in our tendon fibril structure (see Figure 4.15). I expect that inclusion of cross-linking (through elastomeric or imposed local hexagonal packing treatments) will allow us to increase this maximum strain by providing a second mechanism of deformation (D-band and stretching of cross-links). My reasoning for this is through analogy with a simple system of two springs with spring constants k_1 and k_2. If the first spring (the D-band) is pulled on, it will only strain so far before it can break. However, adding a second spring (cross-linking) to the first spring in series will allow for a larger strain to be obtained before the system breaks. Though simplistic, this analogy should be applicable at least qualitatively to the PFCF model with cross-links at small strains, and will require implementation to determine whether it holds true at large enough strains to significantly alter the mechanical response of the system.

The second modification that might be useful is motivated by evidence in the literature that collagen fibrils may actually have a “hollow” cylindrical core near the fibril centre [153, 99]. We could incorporate this into model in two different ways. We could cut out the centre of the fibril in our model, and so the corresponding fibril geometry would be a cylinder with collagen molecules arranged in a double-twist fashion, but with an empty core. This approach would be analogous to the treatment of defects in the director field formalism of liquid crystal theory, where singularities in the director field at defects would be removed from the integration region, and the energy there would be replaced by an isotropic core energy [150]. The question of what the core energy might be in the fibril would add another parameter (and so another level of complexity) to our model.

A second method of introducing an “empty” core fibril, would be to have a radially dependent D-band amplitude which is zero within the core, and non-zero outside of the core, i.e.

$$\delta(r) = \begin{cases} 0 & 0 \leq r \leq R_c \\ \delta_R & R_c < r \leq R. \end{cases} \tag{5.1}$$

From a mechanical perspective, this may be similar to having an empty core, as there is no D-band present in the fibril centre. However, physically this could be interpreted as a disorder core (no D-band) of tropocollagen in the fibril centre. It would also be appealing in that it introduces no new free parameters, as both R_c and δ_R could be determined from the free energy minimization (along R, η, and $\psi(r)$). It could be entirely possible that the minimum free energy would be obtained for $R_c = 0$ (in fact, this is likely true for the linear $\psi(r)$ phase discussed in Chapter 4). However, if would provide a more general variational guess to the fibril structure and so should in certain cases reduce the free energy of fibril when compared to the $R_c = 0$ case.

Out of all the modifications suggested above, by far the simplest is the introduction of a $\delta(r)$ which has the form of eqn 5.1. It requires little modification of the existing model and numerical framework that has been developed, and does not add additional parameters to the model. I suggest that this modification is investigated in the near future to further improve our understanding of collagen fibril structure.

Bibliography

[1] F B Mallory. A hitherto undescribed fibrillar substance produced by connective-tissue cells. *The Journal of Medical Research*, 10(3):334, December 1903.

[2] P Fratzl, editor. *Collagen: Structure and Mechanics.* Springer US, Boston, MA, 2008.

[3] A I Brown, L Kreplak, and A D Rutenberg. An equilibrium double-twist model for the radial structure of collagen fibrils. *Soft Matter*, 10:8500–8511, October 2014.

[4] S Cameron, L Kreplak, and A D Rutenberg. Polymorphism of stable collagen fibrils. *Soft Matter*, 14(23):4772–4783, 2018.

[5] K R Elder and M Grant. Modeling elastic and plastic deformations in nonequilibrium processing using phase field crystals. *Physical Review E*, 70(5):357, November 2004.

[6] J S Bell, S Hayes, C Whitford, J Sanchez-Weatherby, O Shebanova, C Vergari, C P Winlove, N Terrill, T Sorensen, A Elsheikh, and K M Meek. The hierarchical response of human corneal collagen to load. *Acta Biomaterialia*, 65:216–225, January 2018.

[7] P Fratzl and R Weinkamer. Nature's hierarchical materials. *Progress in Materials Science*, 52(8):1263–1334, November 2007.

[8] Creative commons attribution-noncommerical-noderivatives international.

[9] H Lodish, A Berk, S L Zipursky, P Matsudaira, D Baltimore, and J Darnell. *Molecular Cell Biology.* W. H. Freeman, New York, 4th edition, 2000.

[10] A Rich and F H C Crick. The Structure of Collagen. *Nature*, 176(4489):915–916, November 1955.

[11] J A Petruska and A J Hodge. A subunit model for the tropocollagen molecule. *Proceedings of the National Academy of Sciences of the United States of America*, 51(5):871–876, May 1964.

[12] E J Miller and S Gay. Collagen: An overview. *Methods in Enzymology*, 82:3–32, January 1982.

[13] E J Miller and S Gay. The collagens: An overview and update. *Methods in Enzymology*, 144:3–41, January 1987.

[14] S Ricard-Blum. The collagen family. *Cold Spring Harbor Perspectives in Biology*, 3(1):a004978–a004978, January 2011.

[15] S Ricard-Blum and F Ruggiero. The collagen superfamily: from the extracellular matrix to the cell membrane. *Pathologie Biologie*, 53(7):430–442, September 2005.

[16] K E Kadler, C Baldock, J Bella, and R P Boot-Handford. Collagens at a glance. *Journal of Cell Science*, 120(12):1955–1958, May 2007.

[17] P G De Gennes and J Prost. *The Physics of Liquid Crystals*. International series of monographs on physics. Oxford Univ. Press, London, 2nd edition, 1997.

[18] L Bozec, G van der Heijden, and M Horton. Collagen fibrils: Nanoscale ropes. *Biophysical Journal*, 92(1):70–75, January 2007.

[19] M Asgari, N Latifi, H K Heris, H Vali, and L Mongeau. In vitro fibrillogenesis of tropocollagen type III in collagen type I affects its relative fibrillar topology and mechanics. *Scientific Reports*, 7(1):1392, May 2017.

[20] Y Komai and T Ushiki. The three-dimensional organization of collagen fibrils in the human cornea and sclera. *Investigative Ophthalmology & Visual Science*, 32(8):2244–2258, July 1991.

[21] A K Nair, A Gautieri, S Chang, and M J Buehler. Molecular mechanics of mineralized collagen fibrils in bone. *Nature Communications*, 4(1):1724, April 2013.

[22] E P Katz and S Li. Structure and function of bone collagen fibrils. *Journal of Molecular Biology*, 80(1):1–15, October 1973.

[23] S Weiner and H D Wagner. The material bone: structure-mechanical function relations. *Annual Review of Material Science*, 28(1):271–298, 1998.

[24] T A L Wren, S A Yerby, G S Beaupré, and D R Carter. Mechanical properties of the human achilles tendon. *Clinical Biomechanics*, 16(3):245–251, March 2001.

[25] H Kuivaniemi, G Tromp, and D J Prockop. Mutations in fibrillar collagens (types I, II, III, and XI), fibril-associated collagen (type IX), and network-forming collagen (type X) cause a spectrum of diseases of bone, cartilage, and blood vessels. *Human Mutation*, 9(4):300–315, January 1997.

[26] V R Sherman, W Yang, and M A Meyers. The materials science of collagen. *Journal of the Mechanical Behavior of Biomedical Materials*, 52:22–50, December 2015.

[27] J A J van der Rijt, K O van der Werf, M L Bennink, P J Dijkstra, and J Feijen. Micromechanical testing of individual collagen fibrils. *Macromolecular Bioscience*, 6(9):697–702, September 2006.

[28] S Viguet-Carrin, P Garnero, and P D Delmas. The role of collagen in bone strength. *Osteoporosis International*, 17(3):319–336, December 2005.

[29] A J Bailey, R G Paul, and L Knott. Mechanisms of maturation and ageing of collagen. *Mechanisms of Ageing and Development*, 106(1-2):1–56, December 1998.

[30] R B Svensson, H Mulder, V Kovanen, and S P Magnusson. Fracture mechanics of collagen fibrils: influence of natural cross-links. *Biophysical Journal*, 104(11):2476–2484, June 2013.

[31] M Warner and E M Terentyev. *Liquid Crystal Elastomers*. International series of monographs on physics. Oxford Univ. Press, Oxford, 2003.

[32] B Depalle, Z Qin, S J Shefelbine, and M J Buehler. Influence of cross-link structure, density and mechanical properties in the mesoscale deformation mechanisms of collagen fibrils. *Journal of the Mechanical Behavior of Biomedical Materials*, 52:1–13, December 2015.

[33] J S Graham, A N Vomund, C L Phillips, and M Grandbois. Structural changes in human type I collagen fibrils investigated by force spectroscopy. *Experimental Cell Research*, 299(2):335–342, October 2004.

[34] D F Holmes, Y Lu, T Starborg, and K E Kadler. Collagen fibril assembly and function. *Current Topics in Developmental Biology*, 130:107–142, January 2018.

[35] P Vanamee and K R Porter. Observations with the electron microscope on the solvation and reconstitution of collagen. *Journal of Experimental Medicine*, 94(3):255–268, September 1951.

[36] G C Wood and M K Keech. The formation of fibrils from collagen solutions 1. The effect of experimental conditions: kinetic and electron-microscope studies. *Biochemical Journal*, 75(3):588–598, June 1960.

[37] G C Wood. The formation of fibrils from collagen solutions. 2. A mechanism for collagen-fibril formation. *Biochemical Journal*, 75(3):598–605, June 1960.

[38] G C Wood. The formation of fibrils from collagen solutions. 3. Effect of chondroitin sulphate and some other naturally occurring polyanions on the rate of formation. *Biochemical Journal*, 75(3):605–612, June 1960.

[39] M L Stearns. Studies on the development of connective tissue in transparent chambers in the rabbit's ear. II. *American Journal of Anatomy*, 67(1):55–97, July 1940.

[40] K R Porter and G D Pappas. Collagen formation by fibroblasts of the chick embryo dermis. *The Journal of Cell Biology*, 5(1):153–165, January 1959.

[41] B Goldberg and H Green. An analysis of collagen secretion by established mouse fibroblast lines. *The Journal of Cell Biology*, 22(1):227–258, July 1964.

[42] R L Trelstad and K Hayashi. Tendon collagen fibrillogenesis: Intracellular subassemblies and cell surface changes associated with fibril growth. *Developmental Biology*, 71(2):228–242, September 1979.

[43] E G Canty, Y Lu, R S Meadows, M K Shaw, D F Holmes, and K E Kadler. Coalignment of plasma membrane channels and protrusions (fibripositors) specifies the parallelism of tendon. *The Journal of Cell Biology*, 165(4):553–563, May 2004.

[44] D E Birk and R L Trelstad. Extracellular compartments in matrix morphogenesis: collagen fibril, bundle, and lamellar formation by corneal fibroblasts. *The Journal of Cell Biology*, 99(6):2024–2033, December 1984.

[45] H K Graham, D F Holmes, R B Watson, and K E Kadler. Identification of collagen fibril fusion during vertebrate tendon morphogenesis. The process relies on unipolar fibrils and is regulated by collagen-proteoglycan interaction. *Journal of Molecular Biology*, 295(4):891–902, January 2000.

[46] B B Doyle, D J S Hulmes, A Miller, D A D Parry, K A PIEZ, and J Woodhead-Galloway. A D-periodic narrow filament in collagen. *Proceedings of the Royal Society of London B: Biological Sciences*, 186(1082):67–74, May 1974.

[47] F O Schmitt, C E Hall, and M A Jakus. Electron microscope investigations of the structure of collagen. *Journal of Cellular and Comparative Physiology*, 20(1):11–33, August 1942.

[48] M Raspanti, M Reguzzoni, M Protasoni, and P Basso. Not only tendons: The other architecture of collagen fibrils. *International Journal of Biological Macromolecules*, 107:1668–1674, 2018.

[49] K M Meek, G F Elliott, Z Sayers, S B Whitburn, and M H J Koch. Interpretation of the meridional X-ray diffraction pattern from collagen fibrils in corneal stroma. *Journal of Molecular Biology*, 149(3):477–488, July 1981.

[50] L J Gathercole, J S Shah, and C Nave. Skin-tendon differences in collagen D-period are not geometric or stretch-related artefacts. *International Journal of Biological Macromolecules*, 9(3):181–183, June 1987.

[51] Creative commons attribution international.

[52] T Starborg, Y Lu, R S Meadows, K E Kadler, and D F Holmes. Electron microscopy in cell-matrix research. *Methods*, 45(1):53–64, May 2008.

[53] J W Smith. Molecular pattern in native collagen. *Nature*, 219(5150):157–158, July 1968.

[54] D J S Hulmes, J C Jesior, A Miller, C Berthet-Colominas, and C Wolff. Electron microscopy shows periodic structure in collagen fibril cross sections. *Proceedings of the National Academy of Sciences of the United States of America*, 78(6):3567–3571, June 1981.

[55] D F Holmes, C J Gilpin, C Baldock, U Ziese, A J Koster, and K E Kadler. Corneal collagen fibril structure in three dimensions: Structural insights into fibril assembly, mechanical properties, and tissue organization. *Proceedings of the National Academy of Sciences of the United States of America*, 98(13):7307–7312, June 2001.

[56] G C Na, L J Butz, D G Bailey, and R J Carroll. In vitro collagen fibril assembly in glycerol solution: evidence for a helical cooperative mechanism involving microfibrils. *Biochemistry*, 25(5):958–966, 1986.

[57] G C Na, L J Butz, and R J Carroll. Mechanism of in vitro collagen fibril assembly. Kinetic and morphological studies. *Journal of Biological Chemistry*, 261(26):12290–12299, September 1986.

[58] J P R O Orgel, T C Irving, A Miller, and T J Wess. Microfibrillar structure of type I collagen in situ. *Proceedings of the National Academy of Sciences of the United States of America*, 103(24):9001–9005, June 2006.

[59] D M Maurice. The structure and transparency of the cornea. *The Journal of Physiology*, 136(2):263–286, April 1957.

[60] M A Jakus. Studies on the cornea: I. The fine structure of the rat cornea. *American Journal of Ophthalmology*, 38(1):40–53, July 1954.

[61] S Akhtar. Effect of processing methods for transmission electron microscopy on corneal collagen fibrils diameter and spacing. *Microscopy Research and Technique*, 75(10):1420–1424, October 2012.

[62] A Daxer, K Misof, and B Grabner. Collagen fibrils in the human corneal stroma: structure and aging. *Investigative Ophthalmology & Visual Science*, 39:644–648, December 1997.

[63] G Zhang, S Chen, S Goldoni, B W Calder, H C Simpson, R T Owens, D J McQuillan, M F Young, R V Iozzo, and D E Birk. Genetic evidence for the coordinated regulation of collagen fibrillogenesis in the cornea by decorin and biglycan. *Journal of Biological Chemistry*, 284(13):8888–8897, March 2009.

[64] S Akhtar, T Almubrad, I Paladini, and R Mencucci. Keratoconus corneal architecture after riboflavin/ultraviolet A cross-linking: Ultrastructural studies. *Molecular Vision*, 19:1526, 2013.

[65] D A D Parry, G R G Barnes, and A S Craig. A comparison of the size distribution of collagen fibrils in connective tissues as a function of age and a possible relation between fibril size distribution and mechanical properties. *Proceedings of the Royal Society of London B: Biological Sciences*, 203(1152):305–321, December 1978.

[66] S P Davankar, N J Deane, A S Davies, E C Firth, H Hodge, and D A D Parry. Collagen fibril diameter distributions in ligaments and tendons of the carpal region of the horse. *Connective Tissue Research*, 34(1):11–21, 1996.

[67] J C Patterson-Kane, A M Wilson, E C Firth, D A Parry, and A E Goodship. Comparison of collagen fibril populations in the superficial digital flexor tendons of exercised and nonexercised thoroughbreds. *Equine veterinary journal*, 29(2):121–125, March 1997.

[68] K L Goh, D F Holmes, Y Lu, P P Purslow, K E Kadler, D Bechet, and T J Wess. Bimodal collagen fibril diameter distributions direct age-related variations in tendon resilience and resistance to rupture. *Journal of Applied Physiology*, 113(6):878–888, September 2012.

[69] D A D Parry, A S Craig, and G R G Barnes. Tendon and ligament from the horse: an ultrastructural study of collagen fibrils and elastic fibres as a function of age. *Proceedings of the Royal Society of London B: Biological Sciences*, 203(1152):293–303, December 1978.

[70] J R Harris, A Soliakov, and R J Lewis. In vitro fibrillogenesis of collagen type I in varying ionic and pH conditions. *Micron*, 49:60–68, June 2013.

[71] D E Birk, J M Fitch, J P Babiarz, K J Doane, and T F Linsenmayer. Collagen fibrillogenesis in vitro: interaction of types I and V collagen regulates fibril diameter. *Journal of Cell Science*, 95:649–657, April 1990.

[72] A Ruggeri, F Benazzo, and E Reale. Collagen fibrils with straight and helicoidal microfibrils: a‘ freeze-fracture and thin-section study. *Journal of Ultrastructure Research*, 68(1):101–108, July 1979.

[73] V Ottani, M Raspanti, and A Ruggeri. Collagen structure and functional implications. *Micron*, 32(3):251–260, April 2001.

[74] E Reale, F Benazzo, and A Ruggeri. Differences in the microfibrillar arrangement of collagen fibrils. Distribution and possible significance. *Journal of Submicroscopic Cytology*, 13(2):135–143, April 1981.

[75] K J Stewart. A quantitative ultrastructural study of collagen fibrils in human skin normal scars, and hypertrophic scars. *Clinical Anatomy*, 8(5):334–338, January 1995.

[76] D F Holmes and K E Kadler. The 10+4 microfibril structure of thin cartilage fibrils. *Proceedings of the National Academy of Sciences of the United States of America*, 103(46):17249–17254, November 2006.

[77] J Galloway. Structure of Collagen Fibrils. In *Biology of Invertebrate and Lower Vertebrate Collagens*, pages 73–82. Springer, Boston, MA, Boston, MA, 1985.

[78] W Folkhard, D Christmann, W Geercken, E Knörzer, M H J Koch, E Mosler, H Nemetschek-Gansler, and T Nemetschek. Twisted fibrils are a structural principle in the assembly of interstitial collagens, chordae tendineae Included. *Zeitschrift für Naturforschung C*, 42(11-12):1303–1306, December 1987.

[79] D J S Hulmes, T J Wess, D J Prockop, and P Fratzl. Radial packing, order, and disorder in collagen fibrils. *Biophysical Journal*, 68(5):1661–1670, May 1995.

[80] M Raspanti, V Ottani, and A Ruggeri. Different architectures of the collagen fibril: morphological aspects and functional implications. *International Journal of Biological Macromolecules*, 11(6):367–371, December 1989.

[81] D C Wright and N D Mermin. Crystalline liquids: the blue phases. *Reviews Of Modern Physics*, 61(2):385–432, 1989.

[82] G M Grason and R F Bruinsma. Chirality and equilibrium biopolymer bundles. *Physical Review Letters*, 99(9):796, August 2007.

[83] L D Landau and E M Lifshitz. *Theory of Elasticity.* Course of theoretical physics. Pergamon, Oxford, 2nd edition, 1970.

[84] M J Buehler. Atomistic and continuum modeling of mechanical properties of collagen: Elasticity, fracture, and self-assembly. *Journal of Materials Research*, 21(08):1947–1961, 2006.

[85] M J Buehler. Nature designs tough collagen: Explaining the nanostructure of collagen fibrils. *Proceedings of the National Academy of Sciences of the United States of America*, 103(33):12285–12290, August 2006.

[86] A Gautieri, S Vesentini, A Redaelli, and M J Buehler. Hierarchical structure and nanomechanics of collagen microfibrils from the atomistic scale up. *Nano Letters*, 11(2):757–766, January 2011.

[87] M J Buehler. Nanomechanics of collagen fibrils under varying cross-link densities: Atomistic and continuum studies. *Journal of the Mechanical Behavior of Biomedical Materials*, 1(1):59–67, January 2008.

[88] F Rauch and F H Glorieux. Osteogenesis imperfecta. *The Lancet*, 363(9418):1377–1385, April 2004.

[89] G M Grason. Braided bundles and compact coils: The structure and thermodynamics of hexagonally packed chiral filament assemblies. *Physical Review E*, 79(4):041919, April 2009.

[90] J V Selinger and R F Bruinsma. Hexagonal and nematic phases of chains. I. Correlation functions. *Physical Review A*, 43(6):2910–2921, March 1991.

[91] R D Kamien and D R Nelson. Iterated Moiré Maps and braiding of chiral polymer crystals. *Physical Review Letters*, 74(13):2499–2502, March 1995.

[92] J K Mouw, G Ou, and V M Weaver. Extracellular matrix assembly: a multiscale deconstruction. *Nature Reviews Molecular Cell Biology*, 15(12):771–785, November 2014.

[93] D E Birk, J M Fitch, and T F Linsenmayer. Organization of collagen types I and V in the embryonic chicken cornea. *Investigative Ophthalmology & Visual Science*, 27(10):1470–1477, October 1986.

[94] D J Stephens. Cell biology: Collagen secretion explained. *Nature*, 482(7386):474–475, February 2012.

[95] G Mosser, A Anglo, C Helary, Y Bouligand, and M Giraud-Guille. Dense tissue-like collagen matrices formed in cell-free conditions. *Matrix Biology*, 25(1):3–13, January 2006.

[96] M Fang and M M Banaszak Holl. Variation in type I collagen fibril nanomorphology: the significance and origin. *BoneKEy reports*, 2:394, August 2013.

[97] O Antipova and J P R O Orgel. In situ D-periodic molecular structure of type II collagen. *Journal of Biological Chemistry*, 285(10):7087–7096, February 2010.

[98] V Ottani, D Martini, M Franchi, A Ruggeri, and M Raspanti. Hierarchical structures in fibrillar collagens. *Micron*, 33(7-8):587–596, January 2002.

[99] T Gutsmann, G E Fantner, M Venturoni, A Ekani-Nkodo, J B Thompson, J H Kindt, D E Morse, D K Fygenson, and P K Hansma. Evidence that collagen fibrils in tendons are inhomogeneously structured in a tubelike manner. *Biophysical Journal*, 84(4):2593–2598, April 2003.

[100] J H Lillie, D K MacCallum, L J Scaletta, and J C Occhino. Collagen structure: evidence for a helical organization of the collagen fibril. *Journal of Ultrastructure Research*, 58(2):134–143, February 1977.

[101] D R Baselt, J P Revel, and J D Baldeschwieler. Subfibrillar structure of type I collagen observed by atomic force microscopy. *Biophysical Journal*, 65(6):2644–2655, December 1993.

[102] S Yamamoto, H Hashizume, J Hitomi, M Shigeno, S Sawaguchi, H Abe, and T Ushiki. The Subfibrillar Arrangement of Corneal and Scleral Collagen Fibrils as Revealed by Scanning Electron and Atomic Force Microscopy. *Archives of Histology and Cytology*, 63(2):127–135, 2000.

[103] M Hirsch, G Prenant, and G Renard. Three-dimensional supramolecular organization of the extracellular matrix in human and rabbit corneal stroma, as revealed by ultrarapid-freezing and deep-etching methods. *Experimental Eye Research*, 72(2):123–135, February 2001.

[104] A Ferrarini. The theory of elastic constants. *Liquid Crystals*, 37(6-7):811–823, July 2010.

[105] I C Khoo. Nonlinear optics, active plasmonics and metamaterials with liquid crystals. *Progress in Quantum Electronics*, 38(2):77–117, March 2014.

[106] J Nehring and A Saupe. On the elastic theory of uniaxial liquid crystals. *The Journal of Chemical Physics*, 54(1):337–343, January 1971.

[107] P Sheng. Introduction to the Elastic Continuum Theory of Liquid Crystals. In E B Priestly, P J Wojtowicz, and P Sheng, editors, *Introduction to Liquid Crystals*, pages 103–127. Springer US, Boston, MA, 1976.

[108] Creative commons attribution-sharealike.

[109] S Lee and R B Meyer. Light scattering measurements of anisotropic viscoelastic coefficients of a main-chain polymer nematic liquid crystal. *Liquid Crystals*, 7(1):15–29, January 1990.

[110] L Onsager. The effects of shape on the interaction of colloidal particles. *Annals of the New York Academy of Sciences*, 51(1):627–659, May 1949.

[111] Joseph P Straley. Theory of piezoelectricity in nematic liquid crystals, and of the cholesteric ordering. *Physical Review A*, 14(5):1835–1841, November 1976.

[112] M M C Tortora and J P K Doye. Perturbative density functional methods for cholesteric liquid crystals. *The Journal of Chemical Physics*, 146(18):184504, May 2017.

[113] S Zhou, K Neupane, Y A Nastishin, Alan R Baldwin, S V Shiyanovskii, O D Lavrentovich, and S Sprunt. Elasticity, viscosity, and orientational fluctuations of a lyotropic chromonic nematic liquid crystal disodium cromoglycate. *Soft Matter*, 10(34):6571–6581, September 2014.

[114] W H Press, S A Teukolsky, W T Vetterling, and B P Flannery. Chapter 17. Two Point Boundary Value Problems. In *Numerical recipes in C*, pages 753–787. Cambridge Univ. Press, Cambridge, 1992.

[115] T J Wess. Collagen fibril form and function. *Advances in protein chemistry*, 70:341–374, 2005.

[116] D R Keene, L Y Sakai, H P Bächinger, and R E Burgeson. Type III collagen can be present on banded collagen fibrils regardless of fibril diameter. *The Journal of Cell Biology*, 105(5):2393–2402, November 1987.

[117] R Fleischmajer, J S Perlish, R E Burgeson, F Shaikh-Bahai, and R Timpl. Type I and type III collagen interactions during fibrillogenesis. *Annals of the New York Academy of Sciences*, 580(1):161–175, February 1990.

[118] A Geerts, D Schuppan, S Lazeroms, R De Zanger, and E Wisse. Collagen type I and III occur together in hybrid fibrils in the space of Disse of normal rat liver. *Hepatology*, 12(2):233–241, 1990.

[119] E Adachi and T Hayashi. In vitro formation of hybrid fibrils of type V collagen and type I collagen. Limited growth of type I collagen into thick fibrils by type V collagen. *Connective Tissue Research*, 14(4):257–266, 1986.

[120] R D Young, P A Lawrence, V C Duance, T Aigner, and P Monaghan. Immunolocalization of collagen types II and III in single fibrils of human articular cartilage. *Journal of Histochemistry & Cytochemistry*, 48(3):423–432, March 2000.

[121] T Itoh, H Tsuchiya, Y Yoshimura, M Hashimoto, and T Konishi. Structural analysis of collagen fibrils in rat skin based on small-angle x-ray-diffraction pattern. *Japanese Journal of Applied Physics*, 35(Part 1, No. 12A):6172–6179, December 1996.

[122] P De Sa Peixoto, A Deniset-Besseau, M Schanne-Klein, and G Mosser. Quantitative assessment of collagen I liquid crystal organizations: role of ionic force and acidic solvent, and evidence of new phases. *Soft Matter*, 7(23):11203–11210, 2011.

[123] M Kahlweit and W Ostner. An estimation of the interfacial tension between the nematic and isotropic states of a liquid crystal. *Chemical Physics Letters*, 18(4):589–591, February 1973.

[124] D Langevin and M A Bouchiat. Molecular order and surface tension for the nematic-isotropic interface of MBBA, deduced from light reflectivity and light scattering measurements. *Molecular Crystals and Liquid Crystals*, 22(3-4):317–331, January 1973.

[125] W Chen, T Sato, and A Teramoto. Measurement of the interfacial tension between coexisting isotropic and nematic phases of a lyotropic polymer liquid crystal. *Macromolecules*, 29(12):4283–4286, January 1996.

[126] P Oswald and G Poy. Droplet relaxation in Hele-Shaw geometry: Application to the measurement of the nematic-isotropic surface tension. *Physical Review E*, 92(6):062512, December 2015.

[127] S Faetti and V Palleschi. Measurements of the interfacial tension between nematic and isotropic phase of some cyanobiphenyls. *The Journal of Chemical Physics*, 81(12):6254–6258, December 1984.

[128] R Williams. Research note: measurement of the interfacial tension between nematic and isotropic phases of MBBA. *Molecular Crystals and Liquid Crystals*, 35(3-4):349–351, January 1976.

[129] M Doi and N Kuzuu. Structure of the interface between the nematic phase and the isotropic phase in the rodlike molecules. *Journal of Applied Polymer Science*, pages 65–68, 1985.

[130] D B DuPré and R W Duke. Temperature, concentration, and molecular weight dependence of the twist elastic constant of cholesteric poly-γ-benzyl-L-glutamate. *The Journal of Chemical Physics*, 63(1):143–148, 1975.

[131] V G Taratuta, F Lonberg, and R B Meyer. Anisotropic mechanical properties of a polymer nematic liquid crystal. *Physical Review A*, 37(5):1831–1834, March 1988.

[132] H Toriumi, K Matsuzawa, and I Uematsu. Magnetotropic cholesteric-to-nematic transition in lyotropic poly(γ-benzyl L-glutamate) liquid crystals in m-cresol. *The Journal of Chemical Physics*, 81(12):6085–6089, 1984.

[133] D W Allender, G P Crawford, and J W Doane. Determination of the liquid-crystal surface elastic constant K24. *Physical Review Letters*, 67(11):1442–1445, September 1991.

[134] R D Polak, G P Crawford, B C Kostival, J W Doane, and S Zumer. Optical determination of the saddle-splay elastic constant K24 in nematic liquid crystals. *Physical Review E*, 49(2):R978–R981, February 1994.

[135] M Kröger and P Ilg. Derivation of Frank-Ericksen elastic coefficients for polydomain nematics from mean-field molecular theory for anisotropic particles. *The Journal of Chemical Physics*, 127(3):034903, July 2007.

[136] W Schwarz. Elektronenmikroskopische Untersuchungen über die Differenzierung der cornea- und Sklerafibrillen des Menschen. *Zeitschrift für Zellforschung und Mikroskopische Anatomie*, 38(1):78–86, 1953.

[137] K M Meek and C Knupp. Corneal structure and transparency. *Progress in Retinal and Eye Research*, 49:1–16, November 2015.

[138] B Brodsky, E F Eikenberry, and K Cassidy. An unusual collagen periodicity in skin. *Biochimica et Biophysica Acta*, 621(1):162–166, January 1980.

[139] Y Bouligand, J P Denefle, J P Lechaire, and M Maillard. Twisted architectures in cell-free assembled collagen gels: study of collagen substrates used for cultures. *Biology of the Cell*, 54(2):143–162, 1985.

[140] N S Kalson, Y Lu, S H Taylor, T Starborg, D F Holmes, and K E Kadler. A structure-based extracellular matrix expansion mechanism of fibrous tissue growth. *elife*, 4:8427, May 2015.

[141] C Alves, A D Araújo, C L N Oliveira, J Imsirovic, E Bartolák-Suki, J S Andrade, and B Suki. Homeostatic maintenance via degradation and repair of elastic fibers under tension. *Scientific Reports*, 6(1):677, June 2016.

[142] J R Harris and A Reiber. Influence of saline and pH on collagen type I fibrillogenesis in vitro: Fibril polymorphism and colloidal gold labelling. *Micron*, 38(5):513–521, July 2007.

[143] F Berenguer, R J Bean, L Bozec, J Vila-Comamala, F Zhang, C M Kewish, O Bunk, J M Rodenburg, and I K Robinson. Coherent x-ray imaging of collagen fibril distributions within intact tendons. *Biophysical Journal*, 106(2):459–466, January 2014.

[144] M Fang, E L Goldstein, E K Matich, B G Orr, and M M Banaszak Holl. Type I collagen self-assembly: the roles of substrate and concentration. *Langmuir*, 29(7):2330–2338, February 2013.

[145] K M Meek. Corneal collagen—its role in maintaining corneal shape and transparency. *Biophysical Reviews*, 1(2):83–93, June 2009.

[146] R A Farrell and R W Hart. Light scattering in the cornea. *Journal of the Optical Society of America*, 59(6):766–774, June 1969.

[147] G B Benedek. Theory of transparency of the eye. *Applied Optics*, 10(3):459–473, 1971.

[148] M. et al Galassi. Gnu scientific library reference manual. In *GNU Scientific Library Reference Manual*. 3rd edition, 2018.

[149] S. Cameron. pfcfibrils. `https://github.com/samueljmcameron/pfc_fibrils`, 2019.

[150] P M Chaikin and T C Lubenskii. *Principles of Condensed Matter Physics*. Cambridge Univ. Press, Cambridge, 1995.

[151] D Silver, J Miller, R Harrison, and D J Prockop. Helical model of nucleation and propagation to account for the growth of type I collagen fibrils from symmetrical pointed tips: a special example of self-assembly of rod-like monomers. *Proceedings of the National Academy of Sciences of the United States of America*, 89(20):9860–9864, October 1992.

[152] L M Jawerth, M Ijavi, M Ruer, S Saha, M Jahnel, A A Hyman, F Jülicher, and E Fischer-Friedrich. Salt-dependent rheology and surface tension of protein condensates using optical traps. *Physical Review Letters*, 121(25):258101, December 2018.

[153] M Raspanti, M Reguzzoni, M Protasoni, and D Martini. Evidence of a discrete axial structure in unimodal collagen fibrils. *Biomacromolecules*, 12(12):4344–4347, November 2011.

[154] L Yang, K O van der Werf, C F C Fitié, M L Bennink, P J Dijkstra, and J Feijen. Mechanical properties of native and cross-linked type I collagen fibrils. *Biophysical Journal*, 94(6):2204–2211, March 2008.

[155] P E Gill, W Murray, and M H. Wright. *Practical Optimization.* Academic Press Limited, London, 1st edition, 1981.

[156] R. Fletcher. *Practical Methods of Optimization.* John Wiley and Sons, Chichester, 2nd edition, 1987.

Appendix A

Derivation of phase field collagen fibril model

In it's most general form, the free energy of a material which is composed of chiral, rod-like molecules, and has a crystalline density structure is given by the free energy

$$\mathcal{F} = \int_{\text{all space}} d^3r \left(f(Q_{ij}, \partial_k Q_{lm}, \phi) + g(\nabla_{\parallel}\phi, \nabla_{\perp}\phi) + h(\phi) \right), \tag{A.1}$$

where Q_{ij} is the (tensor) order parameter, which is spatially non-uniform and non-zero in any non-isotropic phase, and ϕ is the density order parameter which quantifies the change in density from some reference phase, ρ_{ref} (which is homogeneous in density). There are two pieces to this free energy:

i The liquid crystal free energy, $f(Q_{ij}, \partial_k Q_{lm}, \phi)$, which penalizes any non-equilibrium distortions (e.g. splay, bend, twist) in the average molecular orientation.

ii The phase-field crystal free energy, $g(\nabla_{\parallel}\phi, \nabla_{\perp}\phi)$, which in general allows for periodic modulations in the density below some transition temperature, T_m. ∇ is broken up into it's parallel and perpendicular components with respect to the local orientation of the rod-like molecules. A potential well term, $h(\phi)$, which encourages non-zero ϕ values below some transition temperature T_p (which could in general be different than T_m) is also included.

A.1 Liquid crystal free energy

In what follows, I will make several large assumptions in the form of the first term in eqn A.1.

1.a The first assumption concerns the liquid crystal free energy term, and in particular the order parameter Q_{ij}. Although we are interested in modelling chiral (collagen) molecules, I will assume that the biaxiality of the molecules is small so we are in the "low chirality limit" [81]. In this limit, Q_{ij} is uniaxial, and can be written as

$$Q_{ij} = \lambda(3n_in_j - \delta_{ij}), \tag{A.2}$$

where λ is a position-dependent quantity related to the usual uniaxial scalar order parameter $S(r) = 1/2\,\langle 3cos^2\theta - 1\rangle$, and n_i are the components of the director field $\boldsymbol{n}(\boldsymbol{r})$ which gives the average, local molecular orientation. Since a uniaxial Q_{ij} tends to minimize the bulk (gradient-less) components of the liquid crystal free energy, we ignore these terms in our system, and consider only the deformation free energy, which for uniaxial liquid crystals can most recognizably be written in the form

$$\begin{aligned} f_{\text{frank}}(\nabla\boldsymbol{n}) =& \frac{1}{2}\tilde{K}_{11}(\nabla\cdot\boldsymbol{n})^2 + \frac{1}{2}\tilde{K}_{22}(\boldsymbol{n}\cdot\nabla\times\boldsymbol{n}+q)^2 + \frac{1}{2}\tilde{K}_{33}(\boldsymbol{n}\times(\nabla\times\boldsymbol{n}))^2 \\ &+ \tilde{k}_{13}\nabla\cdot(\nabla\cdot\boldsymbol{n})\boldsymbol{n} - \frac{1}{2}(\tilde{K}_{22}+\tilde{k}_{24})\nabla\cdot(\boldsymbol{n}\times(\nabla\times\boldsymbol{n})+\boldsymbol{n}(\nabla\cdot\boldsymbol{n})) \end{aligned} \tag{A.3}$$

where I have assumed λ is constant and absorbed it into the definitions of the elastic constants $\tilde{K}_{ii}$ and $\tilde{k}_{ij}$.

1.b Since we are interested in modelling collagen fibril structure, I am going to assume that the liquid crystal material is constrained to be within a cylinder of radius R and infinite length, and so the elastic constants are essentially zero outside of the cylinder (as it is surrounded by some isotropic fluid). This changes the domain of integration from all of space to just that of the cylinder. This also has the consequence that the divergence terms (preceded by the elastic constants $\tilde{k}_{13}$ and $\tilde{K}_{22}+\tilde{k}_{24}$) cannot be neglected, and introduces an interfacial free energy per unit length $\mathcal{F}_s/L = 2\pi\gamma R$. This assumption is justified as long as the interactions between individual rod-like molecules and the surrounding isotropic fluid cause the molecules to aggregate, which we will take to be true. Furthermore, with this assumption we are neglecting interactions between cylinder (fibrils).

1.c The director field is constrained to be that of a double-twist structure, with $\boldsymbol{n} = \sin\psi(r)\hat{\theta} + \cos\psi(r)\hat{z}$.

1.d The density modulations ϕ are small enough perturbations as to not affect the values of the elastic constants within the cylinder. Thus, to first order the liquid crystal free energy does not depend on ϕ.

Thus, integrating over the first integral in equation A.1 with the above assumptions,

$$\begin{aligned}\int_{\text{all space}} d^3r f(Q_{ij}, \partial_k Q_{lm}, \phi) &= 2\pi L \int_0^{\infty} r dr f_{\text{frank}}(\nabla \boldsymbol{n}_{\text{double-twist}}) \\ &= 2\pi L \int_0^R r dr \left(\frac{1}{2} K_{22} \Big(q - \psi' - \frac{\sin 2\psi}{2r} \Big)^2 \right. \\ &\qquad \left. + \frac{1}{2} K_{33} \frac{\sin^4 \psi}{r^2} \right) \\ &\quad - \pi L (K_{22} + k_{24}) \sin^2 \psi(R) + 2\pi\gamma R L. \end{aligned} \tag{A.4}$$

A.2 Phase field crystal free energy

The most general terms obeying the symmetries of the rod-like molecules for the second term in eqn A.1 is

$$g(\nabla_{\parallel}\phi, \nabla_{\perp}\phi) = -a_{\parallel}(\nabla_{\parallel}\phi)^2 + b_{\parallel}(\nabla_{\parallel}^2\phi)^2 - a_{\perp}(\nabla_{\perp}\phi)^2 + b_{\perp}(\nabla_{\perp}^2\phi)^2 + c(\nabla_{\parallel}\phi)^2(\nabla_{\perp}\phi)^2. \tag{A.5}$$

For thermodynamic stability, both $b_{\parallel} > 0$ and $b_{\perp} > 0$. $a_{\parallel}$ and $a_{\perp}$ can be of either sign in principle, and could change sign according to some transition temperature T_m [1].

To ensure that non-zero density amplitude modulations are thermodynamically favourable, we also include the third term in eqn A.1, which can most simply be written as

$$h(\phi) = d\phi^2(e\phi^2 - 1). \tag{A.6}$$

The simplifying assumptions I make to reduce the number of terms in the above two equations are:

---

[1] It could be possible that there are two transition temperatures, one for ordering parallel to the molecules, and a second for ordering perpendicular to the molecules.

2.a The local density modulations perpendicular to the rod-like molecules are unimportant and can be coarse-grained out of the free energy, so that I can set any terms containing $\nabla_\perp$ to 0. For collagen fibrils in particular, this ignores the hexagonal-like packing of molecules within the fibril, which occur over length scales of $\approx 1\,\text{nm}$.

2.b The density modulations only occur along the axis of the cylinder, so $\phi(\boldsymbol{r}) \to \phi(z)$. This axis is in general not along the average local direction of the molecules ($\hat{z} \neq \boldsymbol{n}$).

2.c The density modulations have a well defined period, which I will call η. This allows us to average over a period of the structure.

The definitions of $\nabla_\parallel$ and $\nabla_\perp$ are given in terms of the local coordinates of the average orientation of molecules $\boldsymbol{x}'$, which I define with respect to the lab reference frame $\boldsymbol{x}$. The geometry of this local coordinate system is illustrated in Figure A.1. The transformation $\boldsymbol{x}' = A\boldsymbol{x}$ written explicitly is

$$\begin{bmatrix} x' + r \\ y' \\ z'. \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta\cos\psi(r) & \cos\theta\cos\psi(r) & -\sin\psi(r) \\ -\sin\theta\sin\psi(r) & \cos\theta\sin\psi(r) & \cos\psi(r) \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \tag{A.7}$$

and the inverse transformation is

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta\cos\psi(r) & -\sin\theta\sin\psi(r) \\ \sin\theta & \cos\theta\cos\psi(r) & \cos\theta\sin\psi(r) \\ 0 & -\sin\psi(r) & \cos\psi(r) \end{bmatrix} \begin{bmatrix} x' + r \\ y' \\ z'. \end{bmatrix}. \tag{A.8}$$

Importantly, θ and r (and so $\psi(r)$) are held constant in this definition, and should not be thought of as functions of x and y when defining local derivatives $\nabla_\parallel$ and $\nabla_\perp$.

With these transformations, I can define more concretely the gradient

$$\begin{aligned} \nabla_\parallel &= \hat{z}'\frac{\partial}{\partial z'} \\ &= \hat{z}'\left(-\sin\theta\sin\psi(r)\frac{\partial}{\partial x} + \cos\theta\sin\psi(r)\frac{\partial}{\partial y} + \cos\psi(r)\frac{\partial}{\partial z}\right), \end{aligned} \tag{A.9}$$

With assumption 2.a, I don't need to evaluate $\nabla_\perp$, but I will for completeness as it draws attention to the fact that $\partial/\partial z$ terms exist in $\nabla_\perp$, so both assumptions 2.a and

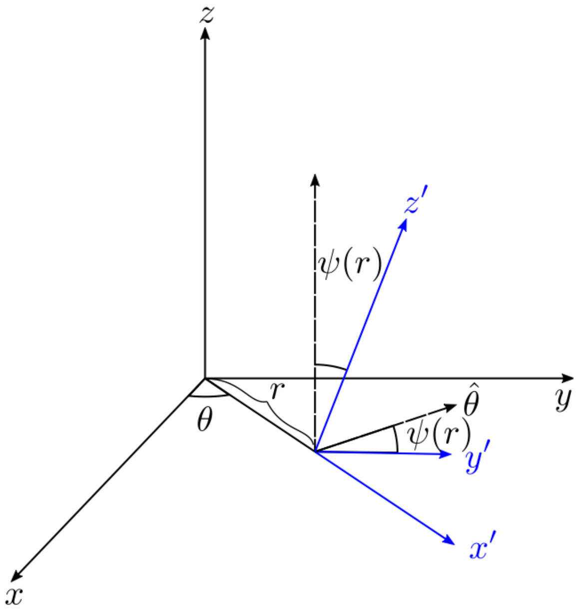

Figure A.1: Defining the local (primed) coordinates of the average molecular orientation with respect to the lab (fibril) reference frame.

2.b are required to safely ignore any $\nabla_\perp$ terms, as

$$\begin{aligned}\nabla_\perp =&\hat{x}'\frac{\partial}{\partial x'}+\hat{y}'\frac{\partial}{\partial y'}\\ =&\hat{x}'\left(\cos\theta\frac{\partial}{\partial x}+\sin\theta\frac{\partial}{\partial y}\right)\\ &+\hat{y}'\left(-\sin\theta\cos\psi(r)\frac{\partial}{\partial x}+\cos\theta\cos\psi(r)\frac{\partial}{\partial y}-\sin\theta\frac{\partial}{\partial z}\right).\end{aligned}\tag{A.10}$$

Using assumptions 2.a and 2.b and the identities $(\nabla\phi)^2=\nabla\cdot(\phi\nabla\phi)-\phi\nabla^2\phi$ and $(\nabla^2\phi)^2=\nabla\cdot\nabla(\phi\nabla^2\phi)-2\nabla\cdot(\phi\nabla\nabla^2\phi)+\phi\nabla^4\phi$, I can re-write the phase field crystal parts of the free energy as

$$\begin{aligned}\int_{\text{all space}}d^3rg(\nabla_\parallel\phi,\nabla_\perp\phi)=&2\pi\frac{\Lambda}{2}\int_0^R rdr\int_0^{\frac{2\pi}{\eta}}dz\phi(z)\left(\left(\frac{2\pi}{d_\parallel}\right)^2+cos^2\psi(r)\frac{\partial^2}{\partial z^2}\right)^2\phi(z)\\ &+\mathcal{F}_{surf},\end{aligned}\tag{A.11}$$

$$\int_{\text{all space}}d^3rh(\phi)=\omega\pi R^2\int_0^{\frac{2\pi}{\eta}}dz\left(\frac{\phi(z)}{\delta_0}\right)^2\left(\left(\frac{\phi(z)}{\delta_0}\right)^2-1\right).\tag{A.12}$$

The terms $\mathcal{F}_{surf}$ are obtained through integration by parts, and in general cannot be neglected for the same reasons that the divergence terms in the Frank free energy cannot be neglected. However, since all of these terms are integrated through a dot product of $\hat{d\boldsymbol{S}} = \hat{r}$ and some vector with direction $\hat{\boldsymbol{V}} = \hat{\nabla\phi} = \hat{z}$ by assumption 2.b, they are zero in our case. I have also taken the free parameter $\delta_0^2 > 0$, which is not necessary in general. $\delta_0^2 > 0$ is equivalent physically to only looking at our system below the critical temperature where density modulations occur (which is where we are interested in investigating). The parameter $d_{\parallel}$ is just the preferred period of the density modulations, and will in general not be equivalent to the true period η because of the $\cos^2\psi(r)$ term in eqn A.11.

A.3 Final form of the free energy per unit volume

In this section, I will use hats to denote dimensional variables, and no hats as dimensionless variables. For the dimensional case, I can write the free energy per fibril volume ($\pi R^2 2\pi/\eta$) bulk phase (i.e. array) of cylindrical, liquid crystalline collagen fibrils as

$$\begin{aligned}\tilde{E}(\tilde{R},\tilde{\eta};\tilde{\psi}(\tilde{r}),\tilde{\phi}(\tilde{z})) = {} & \frac{2}{\tilde{R}^2}\int_0^{\tilde{R}}\tilde{r}d\tilde{r}\left(\frac{1}{2}\tilde{K}_{22}\left(\tilde{q}-\tilde{\psi}'-\frac{\sin 2\tilde{\psi}}{2\tilde{r}}\right)^2+\frac{1}{2}\tilde{K}_{33}\frac{\sin^4\tilde{\psi}}{\tilde{r}^2}\right)\\ & +\frac{\tilde{\Lambda}\tilde{\delta}_0^2}{\tilde{R}^2 2\pi/\tilde{\eta}}\int_0^{\tilde{R}}\tilde{r}d\tilde{r}\int_0^{\frac{2\pi}{\tilde{\eta}}}d\tilde{z}\,\frac{\tilde{\phi}(\tilde{z})}{\tilde{\delta}_0}\left(\left(\frac{2\pi}{\tilde{d}_{\parallel}}\right)^2+cos^2\tilde{\psi}(\tilde{r})\frac{\partial^2}{\partial\tilde{z}^2}\right)^2\frac{\tilde{\phi}(\tilde{z})}{\tilde{\delta}_0}\\ & +\frac{\tilde{\omega}}{2\pi/\tilde{\eta}}\int_0^{\frac{2\pi}{\tilde{\eta}}}d\tilde{z}\left(\frac{\tilde{\phi}(\tilde{z})}{\tilde{\delta}_0}\right)^2\left(\left(\frac{\tilde{\phi}(\tilde{z})}{\tilde{\delta}_0}\right)^2-1\right)\\ & -(\tilde{K}_{22}+\tilde{k}_{24})\frac{\sin^2\tilde{\psi}(\tilde{R})}{\tilde{R}^2}+\frac{2\tilde{\gamma}}{\tilde{R}^2}. \end{aligned} \tag{A.13}$$

Note that the reason we are interested in the free energy per fibril volume instead of the total free energy, is that we assume the fibril phase has already been minimized with respect to volume $V_f = N_f\pi R^2 L$, and since the total free energy is

$$\begin{aligned}\mathcal{F} &= N_f\pi\tilde{R}^2\tilde{L}\tilde{E}\\ &= V_f\tilde{E},\end{aligned} \tag{A.14}$$

we need to find the minimum of $\mathcal{F}$ with V_f held constant.

If I take a single mode approximation for the density, $\tilde{\phi}(\tilde{z}) = \tilde{\delta}\cos(\tilde{\eta}\tilde{z})$ and replace all variables with their dimensionless forms

$$E = \frac{\tilde{E}}{\tilde{K}_{22}\tilde{q}^2}, \tag{A.15}$$

$$R = \tilde{R}\tilde{q}, \tag{A.16}$$

$$r = \tilde{r}\tilde{q}, \tag{A.17}$$

$$\psi(r) = \tilde{\psi}(\tilde{r}), \tag{A.18}$$

$$K_{33} = \frac{\tilde{K}_{33}}{\tilde{K}_{22}}, \tag{A.19}$$

$$L = \frac{\tilde{L}}{\tilde{d}_{\parallel}}, \tag{A.20}$$

$$\Lambda = \frac{2\tilde{\Lambda}\tilde{\delta}_0^2}{3\tilde{K}_{22}\tilde{q}^2\tilde{d}_{\parallel}^4}, \tag{A.21}$$

$$\rho_\delta = \frac{\tilde{\rho}_\delta}{\tilde{\delta}_0}, \tag{A.22}$$

$$\delta = \sqrt{\frac{3}{2}}\frac{\tilde{\delta}}{\tilde{\delta}_0}, \tag{A.23}$$

$$\eta = \tilde{\eta}\tilde{d}_{\parallel}, \tag{A.24}$$

$$\omega = \frac{2\tilde{\omega}\tilde{\delta}_0^4}{3\tilde{K}_{22}\tilde{q}^2}, \tag{A.25}$$

$$\gamma = \frac{\tilde{\gamma}}{\tilde{K}_{22}\tilde{q}}. \tag{A.26}$$

then this becomes

$$\begin{aligned} E(R,\eta,\delta;\psi(r)) = {} & \frac{2}{R^2}\int_0^R rdr\left(\frac{1}{2}\left(1-\psi'-\frac{\sin 2\psi}{2r}\right)^2+\frac{1}{2}K_{33}\frac{\sin^4\psi}{r^2}\right) \\ & + \frac{\Lambda\delta^2}{2R^2}\int_0^R rdr\left(4\pi^2-\eta^2\cos^2\psi(r)\right)^2 \\ & + \frac{\omega\delta^2}{2}\left(\frac{\delta^2}{2}-1\right)-(1+k_{24})\frac{\sin^2\psi(R)}{R^2}+\frac{2\gamma}{R}. \end{aligned} \tag{A.27}$$

Appendix B

Numerical methods

B.1 Free energy per fibril volume minimization with respect to $\psi(r)$

To minimize eqn A.27, we start by taking

$$\frac{\delta E}{\delta \psi} = 0 \tag{B.1}$$

which gives the ODE

$$\begin{aligned} \frac{d\psi}{dr} + r\frac{d^2\psi}{dr^2} =& 1 - \cos(2\psi)\left(1 - \frac{\sin(2\psi)}{2r}\right) + K_{33}\frac{\sin(2\psi)\sin^2\psi}{r} \\ & + \Lambda\delta^2\eta^2 r(4\pi^2 - \eta^2\cos^2\psi)\cos\psi\sin\psi, && \text{(B.2a)} \\ \psi(0) =& 0, && \text{(B.2b)} \\ \left.\frac{d\psi}{dr}\right|_{r=R} =& 1 + k_{24}\frac{\sin(2\psi(R))}{2R}. && \text{(B.2c)} \end{aligned}$$

There is no known analytical solution to this ODE system, and so it must be solved numerically. Solving eqns B.2a - B.2c is a nonlinear boundary value problem with nonlinear boundary conditions, and so simple ODE solvers (e.g. Runge-Kutta) are not useful here. In previous work [3], a shooting method of solution was used to solve the $\delta = 0$ version of this boundary value problem. This method of solution is very robust and can usually find a solution to a given boundary value problem. However, it does not take advantage of any prior knowledge one might have about the system. In our case, we expect that for a small perturbation in parameter space, the functional form of $\psi(r)$ should not change significantly. Therefore, we would like a method of solution that uses this non-chaotic behaviour to our advantage.

With this in mind, we have selected a numerical relaxation method of solution, which takes an initial guess of $\psi(r)$, and attempts to relax this guess to an approximate solution of eqns B.2a - B.2c. This is advantageous, as after we determine $\psi(r)$ for a specific set of parameters, we can use its form as an initial guess for a new set of parameters. In the remainder of this section, we outline the procedure of numerical relaxation.

B.1.1 Discretization of $\psi(r)$ and its ODE

To solve eqn B.2a with the boundary condition eqns B.2b and B.2c, We use an ODE numerical relaxation technique [114]. The discretization has M points, on the evenly spaced grid $r_1, ..., r_M$ with spacing h. To simplify notation, we define a vector $\boldsymbol{y}_k$ as

$$\boldsymbol{y}_k = \begin{bmatrix} y_k^{(1)} \\ y_k^{(2)} \end{bmatrix} = \begin{bmatrix} \psi(r_k) \\ \frac{d}{dr}\psi(r_k) \end{bmatrix} \tag{B.3}$$

where r_k is grid point k ($k = 1, 2, ..., M$). We also define a vector function $\boldsymbol{g}(x, \boldsymbol{y})$ with components

$$g^{(1)}(r, \boldsymbol{y}) = y^{(2)}, \tag{B.4a}$$

$$\begin{aligned} g^{(2)}(r, \boldsymbol{y}) =& \frac{1}{r}\Bigg(1 - y^{(2)} - \cos(2y^{(1)})\left(1 - \frac{\sin(2y^{(1)})}{2r}\right) + K_{33}\frac{\sin(2y^{(1)})\sin^2 y^{(1)}}{r} \\ &+ \Lambda\delta^2\eta^2 r(4\pi^2 - \eta^2\cos^2 y^{(1)})\cos y^{(1)}\sin y^{(1)}\Bigg). \end{aligned} \tag{B.4b}$$

With this notation, the finite difference version of eqn B.2a is

$$\boldsymbol{E}_k = \boldsymbol{y}_k - \boldsymbol{y}_{k-1} - (r_k - r_{k-1})\boldsymbol{g}[0.5(r_k + r_{k-1}), 0.5(\boldsymbol{y}_k + \boldsymbol{y}_{k-1})] = \boldsymbol{0}, \;\; k = 2, 3, ..., M, \tag{B.5}$$

where the finite difference is symmetric about the midpoint of the grid points r_k and r_{k-1}, and $\boldsymbol{E}_k = \boldsymbol{0}$ at every k is an approximate solution to eqn B.2a (it is exact in the limit $M \to \infty$).

We define

$$\boldsymbol{E}_1 = \begin{bmatrix} 0 \\ y_1^{(1)} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{B.6}$$

and

$$\boldsymbol{E}_{M+1} = \begin{bmatrix} y_M^{(2)} - 1 - k_{24}\frac{\sin(2y_M^{(1)})}{r_M} \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \tag{B.7}$$

which correspond to the boundary conditions in eqns B.2b and B.2c, respectively. This set of $\boldsymbol{E}_k$ for $k = 1, 2, ..., M+1$ is a non-linear system of $2M$ unknowns $\boldsymbol{y}_k$[1]. To solve this system, a multi-dimensional Newton's method relaxation scheme is applied. This requires an initial guess for the $\boldsymbol{y}_k$, which we take to be linear (i.e. $\psi(r) = \psi_0' r$ for some chosen ψ_0') or constant ($\psi(r) = \psi_0$). Then corrections $\Delta\boldsymbol{y}_k$ are applied by first linearizing the $\boldsymbol{E}_k$ near the initial guess,

$$E_k^{(i)}(\boldsymbol{y}_k + \Delta\boldsymbol{y}_k, \boldsymbol{y}_{k-1} + \Delta\boldsymbol{y}_{k-1}) \simeq E_k^{(i)}(\boldsymbol{y}_k, \boldsymbol{y}_{k-1}) + \frac{\partial E_k^{(i)}}{\partial y_{k-1}^{(j)}} \delta y_{k-1}^{(j)} + \frac{\partial E_k^{(i)}}{\partial y_k^{(j)}} \delta y_k^{(j)} \tag{B.8}$$

(where $i = 1, 2$ and repeated *superscript* indices are summed), and then setting $\boldsymbol{E}(\boldsymbol{y} + \Delta\boldsymbol{y}) = 0$. The resulting form is a linear system of equations, $A\boldsymbol{x} = \boldsymbol{b}$, where the unknown $\boldsymbol{x}$ is the first order corrections $\Delta\boldsymbol{y}_k$ to the current form of $\boldsymbol{y}_k$.

B.1.2 Block diagonal matrix form

It can be shown that the linear system above has a block diagonal form due to the coupling of only two grid points k and $k-1$ in eqn B.8. There are $M+1$ blocks, total, each of which can be conveniently defined in terms of a matrix $S_k^{(i,j)}$ with $i = 1, 2$ and $j = 1, 2, 3, 4$ (dimension 2×4). From eqn B.8 above, the definition of $S_k^{(i,j)}$ is

$$S_k^{(i,j)} = \frac{\partial E_k^{(i)}}{\partial y_{k-1}^{(j)}}, \quad i = 1, 2, \quad j = 1, 2 \tag{B.9}$$

$$S_k^{(i,j+2)} = \frac{\partial E_k^{(i)}}{\partial y_k^{(j)}}, \quad i = 1, 2, \quad j = 1, 2 \tag{B.10}$$

with $k = 2, 3, ..., M$.

The boundary conditions cause the first and last row of the (full) matrix to have components which are defined differently, as eqn B.6 only depends on $\boldsymbol{y}_1$ and contains only one equation[2], while eqn B.7 only depends on $\boldsymbol{y}_M$ and similarly only contains one equation. For this reason, Taylor expanding these two equations gives rise to the

---

[1] $\boldsymbol{E}_1$ and $\boldsymbol{E}_{M+1}$ each only have one equation in them (*vs* two for general $\boldsymbol{E}_k$) since $0 = 0$ is automatically satisfied.

[2] $0 = 0$ is always true.

matrix components

$$S_1^{(2,j)} = \frac{\partial E_1^{(2)}}{\partial y_1^{(j)}}, \;\; j = 1, 2, \tag{B.11}$$

$$S_{M+1}^{(1,j)} = \frac{\partial E_{M+1}^{(1)}}{\partial y_M^{(j)}}, \;\; j = 1, 2. \tag{B.12}$$

To illustrate this point further, I have below written the equation for the full matrix with its block diagonal components in the case of $M = 4$ grid points.

$$\begin{bmatrix} S_1^{(2,1)} & S_1^{(2,2)} & 0 & 0 & 0 & 0 & 0 & 0 \\ S_2^{(1,1)} & S_2^{(1,2)} & S_2^{(1,3)} & S_2^{(1,4)} & 0 & 0 & 0 & 0 \\ S_2^{(2,1)} & S_2^{(2,2)} & S_2^{(2,3)} & S_2^{(2,4)} & 0 & 0 & 0 & 0 \\ 0 & 0 & S_3^{(1,1)} & S_3^{(1,2)} & S_3^{(1,3)} & S_3^{(1,4)} & 0 & 0 \\ 0 & 0 & S_3^{(2,1)} & S_3^{(2,2)} & S_3^{(2,3)} & S_3^{(2,4)} & 0 & 0 \\ 0 & 0 & 0 & 0 & S_4^{(1,1)} & S_4^{(1,2)} & S_4^{(1,3)} & S_4^{(1,4)} \\ 0 & 0 & 0 & 0 & S_4^{(2,1)} & S_4^{(2,2)} & S_4^{(2,3)} & S_4^{(2,4)} \\ 0 & 0 & 0 & 0 & 0 & 0 & S_5^{(1,1)} & S_5^{(1,2)} \end{bmatrix} \begin{bmatrix} \Delta y_1^{(1)} \\ \Delta y_1^{(2)} \\ \Delta y_2^{(1)} \\ \Delta y_2^{(2)} \\ \Delta y_3^{(1)} \\ \Delta y_3^{(2)} \\ \Delta y_4^{(1)} \\ \Delta y_4^{(2)} \end{bmatrix} = \begin{bmatrix} -E_1^{(2)} \\ -E_2^{(1)} \\ -E_2^{(2)} \\ -E_3^{(1)} \\ -E_3^{(2)} \\ -E_4^{(1)} \\ -E_4^{(2)} \\ -E_5^{(1)} \end{bmatrix} \tag{B.13}$$

This block diagonal form can be solved very efficiently using standard linear algebra techniques, see ref. [114] for details. Each time the matrix equation is solved, the form of $\psi(r)$ is updated, i.e. $\boldsymbol{y}_k \leftarrow$ $bmy_k + \Delta \boldsymbol{y}_k$. This process is iterated until convergence is obtained. More specifically, we define the error of the current iteration as

$$\text{err} = \frac{1}{2M} \sum_{k=1}^{M} \sum_{j=1}^{2} \frac{|\Delta y_k^{(j)}|}{\text{scale}_j}, \tag{B.14}$$

where $\text{scale}_1 = 0.1$ and $\text{scale}_2 = 4.0$ are scale factors that should be (and are) around the same magnitude as $\psi(r)$ and $d\psi/dr$, respectively. Once $\text{err} < 1 \times 10^{-10}$, we assume the current values of $\boldsymbol{y}_k$ are an acceptable approximation to the solution of eqns B.2a-B.2c.

B.1.3 Computation of block diagonal ($S_k^{(i,j)}$) elements

Using the definition $\boldsymbol{g}$ from eqns B.4a and B.4b, and taking the grid space $h = R/(M-1)$ to be constant, the form of the $E_k^{(i)}$ in B.5 are

$$
\begin{aligned}
E_{1,k} &= y_{1,k} - y_{1,k-1} - h\frac{(y_{2,k} + y_{2,k-1})}{2}, && \text{(B.15a)}\\
E_{1,M+1} &= y_{2,M} - 1 - k_{24}\frac{\sin(2y_{1,M})}{2r_M}, && \text{(B.15b)}\\
E_{2,k} &= y_{2,k} - y_{2,k-1}\\
&\quad - \frac{2h}{r_k + r_{k-1}}\Bigg\{1 - \frac{(y_{2,k} + y_{2,k-1})}{2} - \cos(y_{1,k} + y_{1,k-1})\Bigg(1 - \frac{\sin(y_{1,k} + y_{1,k-1})}{r_k + r_{k-1}}\Bigg)\\
&\qquad + 2K_{33}\frac{\sin^2\big(\frac{y_{1,k}+y_{1,k-1}}{2}\big)\sin(y_{1,k} + y_{1,k-1})}{r_k + r_{k-1}}\\
&\qquad + \frac{\Lambda\delta^2\eta^2(r_k + r_{k-1})}{2}\big(4\pi^2 - \eta^2\cos^2\big(\frac{y_{1,k} + y_{1,k-1}}{2}\big)\big)\\
&\qquad\qquad \cdot \cos\big(\frac{y_{1,k} + y_{1,k-1}}{2}\big)\\
&\qquad\qquad \sin\big(\frac{y_{1,k} + y_{1,k-1}}{2}\big)\Bigg\}, && \text{(B.15c)}\\
E_{2,1} &= y_{1,1} && \text{(B.15d)}
\end{aligned}
$$

Taking the relevant derivatives of these equations for $k = 2, ..., M$

$$S_{1,1} = -1, \tag{B.16a}$$

$$S_{1,2} = \frac{-h}{2}, \tag{B.16b}$$

$$S_{1,3} = 1, \tag{B.16c}$$

$$S_{1,4} = \frac{-h}{2}, \tag{B.16d}$$

$$\begin{aligned} S_{2,1} = \frac{-2h}{r_k + r_{k-1}} \Bigg\{ & \frac{2K_{33}}{r_k + r_{k-1}} \Bigg(\frac{1}{2} \sin^2(y_{1,k} + y_{1,k-1}) \\ & + \sin^2\big(\frac{y_{1,k} + y_{1,k-1}}{2}\big) \cos(y_{1,k} + y_{1,k-1}) \Bigg) \\ & + \sin(y_{1,k} + y_{1,k-1}) \Bigg(1 - \frac{\sin(y_{1,k} + y_{1,k-1})}{r_k + r_{k-1}} \Bigg) + \frac{\cos^2(y_{1,k} + y_{1,k-1})}{r_k + r_{k-1}} \\ & + \frac{\Lambda \delta^2 \eta^2 (r_k + r_{k-1})}{2} \Bigg(\eta^2 \cos^2\big(\frac{y_{1,k} + y_{1,k-1}}{2}\big) \sin^2\big(\frac{y_{1,k} + y_{1,k-1}}{2}\big) \\ & \qquad + \frac{1}{2} \big(4\pi^2 - \eta^2 \cos^2\big(\frac{y_{1,k} + y_{1,k-1}}{2}\big) \big) \\ & \qquad \cdot \big(\cos^2\big(\frac{y_{1,k} + y_{1,k-1}}{2}\big) \\ & \qquad - \sin^2\big(\frac{y_{1,k} + y_{1,k-1}}{2}\big) \big) \Bigg) \Bigg\}, \end{aligned} \tag{B.16e}$$

$$S_{2,2} = -1 + \frac{h}{r_k + r_{k-1}}, \tag{B.16f}$$

$$S_{2,3} = S_{2,1}, \tag{B.16g}$$

$$S_{2,4} = 1 + \frac{h}{r_k + r_{k-1}}. \tag{B.16h}$$

At the first boundary $k = 1$ (corresponding to $r = 0$),

$$S_{2,1} = 0, \tag{B.17a}$$

$$S_{2,2} = 0, \tag{B.17b}$$

$$S_{2,3} = 1, \tag{B.17c}$$

$$S_{2,4} = 0. \tag{B.17d}$$

and at the second boundary $k = M + 1$ (corresponding to $r = R$),

$$S_{1,1} = 0, \tag{B.17e}$$

$$S_{1,2} = 1, \tag{B.17f}$$

$$S_{1,3} = -k_{24}\frac{\cos(2y_{1,M})}{r_M}, \tag{B.17g}$$

$$S_{1,4} = 1. \tag{B.17h}$$

B.2 Energy minimization with respect to R, η, and δ

The energy per unit volume, E, defined in eqn A.27, is a function of R, η, δ, and a functional of $\psi(r)$. Consistent with the main text, we simplify the minimization process by defining

$$E^*(R, \eta, \delta) \equiv E(R, \eta, \delta, \psi^*(r)) \tag{B.18}$$

where $\psi^*(r)$ is the $\psi(r)$ function[3] which minimizes E at given values of R, η and δ. With the unknown $\psi(r)$ still implicitly being determined at each R, η, and δ, it is not possible to derive analytical forms for the derivatives of E with respect to R, η, and δ. Therefore, the final minimization of E reduces to a simple three variable optimization problem, with

$$\nabla E^* = \begin{bmatrix} \frac{\partial E^*}{\partial R} \\ \frac{\partial E^*}{\partial \eta} \\ \frac{\partial E^*}{\partial \delta} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \tag{B.19}$$

where the partial derivatives are evaluated numerically[4]. There are several robust numerical optimization algorithms designed for this type of problem, including the Conjugate Gradient Descent (CG) algorithm and the BFGS optimization algorithm [155]. After initial tests of computational efficiency which indicated that the BFGS algorithm was marginally faster at finding E^* compared to the CG algorithm, we elected to use the former. We use the BFGS2 algorithm[5] provided by GNU Scientific

---

[3] We solve for $\psi^*(r)$ using the numerical relaxation outlined in section B.1.

[4] The errors associated with these derivatives are typically $\sim 1 \times 10^{-8}$. The error estimate is found by taking the maximum difference of $\mathcal{O}(dx^3)$ and $\mathcal{O}(dx^5)$ numerical derivatives, with code modified from the GNU Scientific Library [148].

[5] The "2" in the algorithm name is a minor difference in one step of the algorithm and was advertised as more efficient.

Library [148]. I provide an outline of the underlying BFGS algorithm next, which closely follows ref. [155].

B.2.1 BFGS algorithm for minimization

To simplify discussion here, I will define $\boldsymbol{x}_k = (R, \eta, \delta)^T$ (with ^{T} indicating the transpose of a vector or matrix)[6], where k is current iteration of the algorithm, and so the energy function is now $E^*(\boldsymbol{x}_k)$. Regardless of the algorithm, all optimization schemes attempt to find $\boldsymbol{x}_k$ which minimizes $E^*(\boldsymbol{x}_k)$ by taking a series of steps, $\boldsymbol{s}_k = \boldsymbol{x}_{k+1} - \boldsymbol{x}_k$. The difference between algorithms such as CG and BFGS is the choice of direction and magnitude for $\boldsymbol{s}_k$.

The calculation of the optimal step at each iteration is best understood by Taylor expanding the energy to second order in $\boldsymbol{s}_k$,

$$E^*(\boldsymbol{x}_{k+1}) \simeq E^*(\boldsymbol{x}_k) + s_k^{(i)} \nabla^{(i)} E^*(\boldsymbol{x}_k) + \frac{1}{2} s_k^{(i)} H^{(i,j)}(\boldsymbol{x}_k) s_k^{(j)}, \tag{B.20}$$

where $H^{(i,j)}$ is the Hessian matrix of second order derivatives and summation over repeated superscript indices is implied.. For $\boldsymbol{x}_k$ sufficiently close to the optimal value, where $\nabla E^*(\boldsymbol{x}_{eq}) = 0$, one can minimize eqn B.20 with respect to the step size to obtain the optimal step size value,

$$s_k^{(i)} = -(H^{-1})^{(i,j)} \nabla^{(j)} E^*(\boldsymbol{x}_k) \tag{B.21}$$

In higher order methods (particularly Newton's method) [155], the Hessian is computed either analytically or numerically. However, we do not have an analytical form for the Hessian $\boldsymbol{H}$ in our model, and higher order numerical derivatives become increasingly expensive and inaccurate to compute numerically. Therefore, we require a different method of choosing the direction and magnitude of $\boldsymbol{s}_k$. This is where the BFGS algorithm is useful.

In qualitative terms, the BFGS algorithm uses information at the previous iteration $k-1$ to build an increasingly accurate approximation to the Hessian at the current step, k. At any given step k, define the current approximation to the Hessian as $Q_k^{(i,j)}$. From eqn B.21, an approximation to the optimal step, $\boldsymbol{t}_k \simeq \boldsymbol{s}_k$, would be

$$t_k^{(i)} = -(Q^{-1})^{(i,j)} \nabla^{(j)} E^*(\boldsymbol{x}_k). \tag{B.22}$$

---

[6] In our calculations, we scale or "precondition" R, η, and *delta* to ensure that numerical error is not too large. See Chapter 8 of [155].

This step can be further improved by taking adjusting the magnitude of this step with a scalar, i.e. $\alpha_k \boldsymbol{t}_k$, where α_k is chosen to minimize $E^*(\boldsymbol{x}_k - \alpha \boldsymbol{t}_k)^7$. Determining this value of α_k, the variables of the $k+1$ iteration can be computed as $\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \alpha_k \boldsymbol{t}_k$.

To improve the approximation to the Hessian at step k, $Q_k^{(i,j)}$, we can take the derivative of eqn B.20 with respect to $\boldsymbol{x}_{k+1}$, substituting in the unknown value of $Q_{k+1}^{(i,j)}$ in for the Hessian matrix. This gives an approximate relationship of the form

$$\nabla_{k+1}^{(i)} E^*(\boldsymbol{x}_{k+1}) - \nabla_k^{(i)} E^*(\boldsymbol{x}_k) \simeq Q_{k+1}^{(i,j)}(x_{k+1}^{(j)} - x_k^{(j)}). \tag{B.23}$$

However, this relationship does not uniquely determine $Q_{k+1}^{(i,j)}$, which is $n \times n$ in size, but eqn B.23 only constrains n variables. The matrix is also required to be symmetric (as the true Hessian is symmetric), but this reduction in the number of unique components (down to $n(n+1)/2$ is still not enough for eqn B.23 to yield a unique solution. The BFGS algorithm introduces an additional constraint that the inverse of $Q_{k+1}^{(i,j)}$ is "close" to the previous $Q_k^{(i,j)}$, i.e.

$$\min_Q = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} (Q_{k+1}^{(i,j)} - Q_k^{(i,j)})}. \tag{B.24}$$

With this final condition, $Q_{k+1}^{(i,j)}$ can be calculated.

Once $Q_{k+1}^{(i,j)}$ has been determined, this whole process can be repeated for $k+1, k+2, ...$ until some convergence criteria are met. For our purposes we select the norm $|(\nabla E^*)| < 1 \times 10^{-8}(1 + E^*)$, a suitable value when considering the error accumulated in taking numerical derivatives.

---

[7] This calculation of α_k is known as a *line search*, and varies mildly between different algorithms. The details are a bit too messy to get into here, but see ref. [156]