

Yale University

EliScholar – A Digital Platform for Scholarly Publishing at Yale

Yale Graduate School of Arts and Sciences Dissertations

Spring 2022

Computational Studies of Packing and Jamming in Biological Systems

John Durkin Treado

Yale University Graduate School of Arts and Sciences, treado.jack@gmail.com

Follow this and additional works at: https://elischolar.library.yale.edu/gsas_dissertations

Recommended Citation

Treado, John Durkin, "Computational Studies of Packing and Jamming in Biological Systems" (2022). *Yale Graduate School of Arts and Sciences Dissertations*. 669.
https://elischolar.library.yale.edu/gsas_dissertations/669

This Dissertation is brought to you for free and open access by EliScholar – A Digital Platform for Scholarly Publishing at Yale. It has been accepted for inclusion in Yale Graduate School of Arts and Sciences Dissertations by an authorized administrator of EliScholar – A Digital Platform for Scholarly Publishing at Yale. For more information, please contact elischolar@yale.edu.

Abstract

Computational Studies of Packing and Jamming in Biological Systems

John D. Treado
2022

The application of coarse-grained computational models to the study of physical systems has exploded in recent years, in part due to the relative simplicity of such models compared to the drastic complexity that can be found in the natural world. These models have been of particular use to the study of biological systems, as living things are typically highly complex and live far from thermodynamic equilibrium. In this thesis, I will present several coarse-grained computational models of different biological systems with the aim of identifying the role played by the purely physical constraints of packing and jamming. In the first part of this thesis, I will describe a model of globular protein cores based on jammed granular materials. I will show that this model displays the same void structure as real globular protein cores, and that the inherent protocol-dependence of packing generation yields insights into systematic differences between various experimental techniques used to resolve protein structures. In the second part of this thesis, I will describe a computational model of particles that can deform their shape in response to applied stress. I will first analyze how the rigidity of single particles in this model affects the collective behavior of many co-interacting deformable particles, as well as indicate how this model may be adapted for the study of tissue fluidization. I end the thesis with an analysis of packing constraints across the development and phylogeny of the spongy mesophyll tissue of leaves and flowers.

Computational Studies of Packing and Jamming in Biological Systems

A Dissertation
Presented to the Faculty of the Graduate School
of
Yale University
in Candidacy for the Degree of
Doctor of Philosophy

by
John D. Treado

Dissertation Director: Corey S. O'Hern

May 2022

Copyright © 2022 by John D. Treado
All rights reserved.

Acknowledgments

My favorite joke to make over the past few years, whenever someone would ask, “are you done yet?”, would be to reply that PhD work is like a mother’s work: it is never done. Certainly, “raising” this dissertation has taken a village. Here are the people who made my graduate career bearable, enjoyable, and without whom I would not have succeeded.

I extend my utmost and foremost gratitude to my advisor, Corey. I have learned an immeasurable amount from being his student. I will never forget playing soccer in Amherst, research meetings with Subway and Cinnamon, Java Station in Santa Barbara, or our end-of-day discussions, when I was last in the office and Corey poked his head in, saying, “the f*** are you doing?” I will always appreciate his candor, his support, his insights, his advice, and most importantly, his example. Corey has taught me not only how to perform research, but how to be a researcher. It is a gift that I can never repay, and for which I am eternally grateful.

I also thank Mark Shattuck for being one of my most important advisors and mentors. I have learned that no matter how smart you are, you probably don’t know as much as Mark. I thank Mark for sharing his incredible wisdom and insights with me regarding my research and beyond, for always being up for a good debate, and for his uncanny ability to ask the most insightful questions. I thank my committee members, Benjamin Machta and Michael Murrell, for supporting me during my graduate work in a myriad ways, but most importantly for fostering an inclusive and exciting biological physics environment at Yale. I also thank the many wonderful individuals I have had the privilege of collaborating with during my studies, including Lynne Regan, Zach Levine, Eric Weeks, Scott Holley, Claudia Fischbach and Garrett Beeghly. In particular, I thank Adam Roddy, Craig Brodersen, Jeroen Schreel and Aleca Borsuk for their collaboration and patience, and for giving me a newfound love of plant science.

I thank my early mentors in the Jamming group, Abe Clark and Jennifer Gaines, as well as more recent collaborators Alex Grigas, Zhe Mei, Dong Wang, Yuxuan Cheng, Francesco Arceri, Andrew Ton, and the rest of the deformable particle subgroup. I look forward to what you all make next. I also thank Doro Noble for running Yale’s PEB program and Daniel Seara for always being there to talk science.

I thank Sam Bryant for living with me for so long, despite how hard that must have been on him, and for being one of the funniest, kindest, and smartest people I have ever known. I thank Gwen Gage for rounding out our quaran-team, for leaving pool noodles everywhere, and for being there to talk to me when I needed it most. I also thank the rest of the KBT crew for many things, but most of all, for providing unwavering joy in my life.

I am blessed to have two wonderful parents, Pat and Carey, who have been my #1 fans since the beginning. I thank them for their love, support, advice, and for providing safe harbor during my studies regardless of the storm. I also thank my sister Grace and her husband Seth for their love and support, as well as for always welcoming me with open arms in their Brooklyn apartment.

Finally, my dissertation could not have been written without the love of my life, Margie Fuchs. Margie, thank you for everything. Thank you for supporting me during my most difficult moments, for rejoicing in my accomplishments, for commiserating in my failures, and for sacrificing all you have sacrificed for us. I look forward to our exciting new life together in Dresden and beyond.

For Margie & my family.

Contents

1	Introduction	1
2	Packing and voids in protein cores	3
2.1	Introduction	3
2.2	Methods	5
2.3	Results	8
2.3.1	Local Void Analysis	8
2.3.2	Connected void analysis of protein cores	12
2.3.3	Connected void analysis of packings	16
2.4	Conclusions	18
2.5	Appendix	20
2.5.1	Packing-generation Protocol	20
2.5.2	Protein core size distribution	21
2.5.3	Measurement of Fisher Exponent	22
3	Packing in NMR and X-ray crystal structures	24
3.1	Introduction	24
3.2	Methods	26
3.2.1	Protein structure datasets	26
3.2.2	NMR structural quality	26
3.2.3	Relative solvent accessible surface area (rSASA)	26
3.2.4	Packing fraction	27
3.2.5	Side chain repacking	27
3.2.6	Jammed packings of amino-acid-shaped particles	28
3.3	Results	28
3.4	Conclusions	34
3.5	Appendix	34
3.5.1	Protein structure datasets	35
3.5.2	Placement of hydrogen atoms in protein structures determined using NMR	39
3.5.3	Fraction of identical core residues and the distribution of root-mean-square deviations in atomic positions	41

3.5.4	Single residue sidechain repacking using the hard-sphere plus stereochemical constraint model	42
3.5.5	Resolution dependence of RMSD between x-ray crystal structure and NMR structure pairs	44
3.5.6	Protocol dependence of packings of amino-acid-shaped particles	45
4	Jamming and response of packings of deformable polygons	48
4.1	Introduction	48
4.2	Methods	50
4.3	Results	52
4.3.1	Rigidity	52
4.3.2	Vibrational response	55
4.3.3	Shear response	57
4.3.4	Preliminary studies on non-linear behavior	58
4.4	Conclusions	61
4.5	Appendices	62
4.5.1	Force definitions	62
4.5.2	Stress definitions	65
4.5.3	Dynamical matrix elements	69
4.5.4	Counting degrees of freedom from vibrational modes	76
4.5.5	Particle buckling with bending energy	80
4.5.6	System size dependence of the vibrational density of states and shear modulus	80
4.5.7	Mode decomposition	81
5	Beyond packing and jamming: diversity and development of spongy mesophyll	84
5.1	Introduction	84
5.2	Spongy mesophyll in flowers	87
5.2.1	Microstructural analysis	88
5.2.2	Computational modelling of anisotropic mesophyll	91
5.3	Spongy mesophyll in leaves	94
5.4	Conclusion	99
5.5	Appendices	100
5.5.1	Mesophyll bending force	100
5.5.2	Derivation of enthalpy minimization	102
6	Conclusions	105
6.1	Thesis summary	105
6.2	Future Directions	107
	Bibliography	109

List of Figures

2.1	Visualization of two kinds of voids used to probe protein core packing structure . . .	4
2.2	Comparison between core packing in x-ray crystal structures of single protein molecules and jammed packings of amino -cid-shaped particles	6
2.3	Fitting parameter k for Voronoi cell volume distributions in packings and protein cores	9
2.4	Local porosity distribution in packings and protein cores	10
2.5	Comparison of void percolation in packings and protein cores	11
2.6	Critical porosity at void percolation as a function of lattice resolution	12
2.7	Scaling behavior of percolating connected voids	13
2.8	Visualization of percolating connected voids in model systems	16
2.9	Core residue size distribution in Dunbrack 1.0 database	21
2.10	Fisher exponent as a function of lattice resolution	22
2.11	Cluster size distribution as a function of system size	23
3.1	Root-mean-square deviations (RMSD) of NMR structures as a function of the number of distance restraints	25
3.2	RMSD of all structures and between pairs of proteins resolved by different experimental techniques	29
3.3	Side-chain repacking and clashes for core residues in NMR and crystal structures . .	30
3.4	Core packing fraction in NMR and crystal structures	31
3.5	Packing fraction at jamming onset of thermalized packings of amino acid-shaped particles	33
3.6	Packing fraction with different hydrogen placements	39
3.7	Core identity comparison	41
3.8	RMSD comparison between all datasets	42
3.9	NMR-crystal RMSD as a function of crystal structure resolution in our datasets . .	44
3.10	NMR-crystal RMSD as a function of crystal structure resolution in the CoDNaS database	45
3.11	Protocol dependence of packing fraction at jamming of amino acid-shaped particles .	47
4.1	Vibrational spectrum of single deformable polygons	49
4.2	Contact counting in packings of deformable polygons with and without bending energy	51
4.3	Density of vibrational states in jammed packings of deformable polygons	54

4.4	Collective shape degrees of freedom in different vibrational eigenvectors	55
4.5	The rigid-particle limit	56
4.6	Static shear modulus for jammed packings of deformable polygons at different pressures	57
4.7	Effect of crawling motility and cell mechanics on glassy dynamics	59
4.8	Geometry of DPb particles	62
4.9	Eigenvalues of the dynamical matrix used for constraint counting	77
4.10	Counting constraints in packings of DPb particles	78
4.11	Buckling in DPb particles	79
4.12	Effect of system size on VDOS and the shear modulus	80
4.13	Decomposition of normal models into translational, rotational and shape degrees of freedom	82
5.1	Overview of spongy mesophyll microstructure in leaves and flowers	85
5.2	Examples of correlation analysis on several microCT samples	87
5.3	Anisotropy, packing and cell size across the angiosperms	89
5.4	Computational models for complex three-dimensional microstructures	91
5.5	Developmental trajectory of spongy mesophyll cells in a developing <i>Arabidopsis thaliana</i> leaf.	94
5.6	Cell shapes in <i>Arabidopsis</i> mesophyll	97
5.7	Progression of cell shapes during simulated development of a small system	99
5.8	Geometry of bending angle in the mesophyll cell model	102
6.1	Schematic evolutionary simulation scheme using deformable polygons	108

List of Tables

2.1	Table of void percolation critical exponents	15
3.1	Dataset of NMR models with crystal structure pairs from the Protein Data Bank (PDB)	36
3.2	Dataset of crystal structure duplicates from the PDB	37
3.3	Dataset of sole NMR structures from the PDB	38
3.4	Dataset of mutant crystal structures with a resolved wildtype pair	40
3.5	Comparison of RMSD between datasets	41
3.6	Atomic radii	43

Chapter 1

Introduction

The study of biological systems using theoretical physics has grown substantially over the last half century, in part due to the relative simplicity of model physical systems compared to the drastic complexity inherent in living things. As advanced computational resources and methods have also increased during this same time frame, computational physics has become the tool of choice for studying complex physical systems of all varieties, from complex materials to climate change. Computational biological physics has therefore emerged as an important area of research, uniting theoretical concepts in the areas of soft condensed matter and statistical physics to experimental observations of biological processes across scales, from individual biomolecules to the scale of ecosystems. The ability of computational methods to capture relevant phenomena across such a broad range of scales relies on the simple fact that computational models simplify biological systems down to their barest.

For example, the Vicsek Model of collective migration [1] seeks to model the ability of individual biological agents to move together collectively, yet in its simplest form its only relevant parameter is the amount of error each agent makes in aligning their local velocities. Despite this incredibly simple form, the model displays a novel phase transition that helps explain complex migration behavior across a wide variety of biological systems, from bacterial colonies [2, 3] to the developing zebrafish tailbud [4] to flocks of starling [5]. While the Vicsek Model itself has been heavily studied since its inception, the fact of *why* it is successful has also garnered significant attention [6]. While this remains an important direction of research in and of itself, the consensus as to why simple models can be applied to complex biological phenomena is emerging that there must be some dominant physical contribution that can explain at least a significant portion of data in any given biological system. It therefore remains vital to construct simplified (even overly-simplified) physical models of biological systems to seek out these dominant contributions. With such models in hand, we can begin to build a fundamentally physical understanding of biological systems where the knowledge of certain constraints allows one to predict and describe vastly different biological phenomena.

In this thesis, I will describe two separate research directions in this spirit that attempt to boil down different biological down to their essentials in order to better understand problems of biological diversity and universality. The first direction will focus on studying the arrangement

of amino acids in the cores of folded proteins, and the second direction will focus on constructing models of densely-packed cells in tissues. While these models are focused on phenomena separated by orders of magnitude of length and time, they are united by the importance of *packing* in biological phenomena. Packing, or the spatial arrangement of discrete objects into a confined space, plays a vital role in many material and mathematical problems, but is an essential constraint across biology for the simple reason that life often seeks to efficiently store lots of material in a small space. From the packing of DNA in chromosomes to the packing of buildings in organized human settlement, space remains an ever-important constraint to which Life must always adapt. When things become packed too tightly, they become *jammed*, which can either aid or hinder biological processes. In this thesis, I will describe the development and application of computational models meant for the study of packing and jamming to protein cores and cells in tissues.

In Chapter 2, I will demonstrate how jamming can be used as a framework for understanding the structure of protein cores in globular (i.e. folded) proteins. I will describe a computational model of jammed packings of amino acid-shaped particles, and show that the distribution of void space and local packing environments in our model system will be nearly identical to that observed in high-resolution X-ray crystal structures of globular proteins. In Chapter 3, I will compare X-ray crystal structures to protein structures obtained by nuclear magnetic resonance (NMR) spectroscopy, and use protocol-dependence of packings of amino acid-shaped particles to help understand systematic differences observed between the two classes of experimentally-observed protein structures. In Chapter 4, I will turn my attention to the development and study of a model of deformable polygons that can be used to study a host of soft matter systems composed of objects that can change shape in response to stress. In particular, I will focus on how mechanical properties at the single-particle level impact mechanical properties of the bulk. Finally, in Chapter 5, I will describe how to look beyond jamming to how packing constraints across developmental stage and species affects the spongy mesophyll, a porous tissue that provides stability and airflow to the interiors of leaves and flowers. Chapter 2 will cover material published in Ref. [7], Chapter 3 will cover material published in Ref. [8], and Chapter 4 will cover material published in Ref. [9].

Chapter 2

Packing and voids in protein cores

2.1 Introduction

A significant driving force in protein folding is the sequestration of hydrophobic amino acids from solvent. Moreover, these buried amino acids are densely packed in the protein core [10]. In fact, the packing of core residues has been linked directly to protein stability [11]. For example, large-to-small amino acid mutations, which can increase interior protein cavities, or voids, are known to destabilize proteins when they are subjected to hydrostatic pressure [12–14] and chemical denaturants [15, 16]. Understanding the connection between dense core packing and voids is therefore crucial to understanding the physical origins of protein stability and reliably designing new protein structures that are stable [17]. However, no such quantitative understanding yet exists, and it is currently difficult to distinguish computational protein designs that are not stable in experiments from experimentally observed structures [18].

In previous studies [19–21], we found, using collective side chain repacking, that the side chain conformations of residues in protein cores (from a collection of high-resolution protein crystal structures) are uniquely specified by hard-sphere, steric interactions. Moreover, we have shown that, when considering hard-sphere optimized atomic radii, the core regions in proteins possess the same packing fraction $\phi \approx 0.56$ as that found in simulations of dense, random packings of purely repulsive, amino acid-shaped particles. This result suggests that the packing fraction of protein cores is determined by the bumpy and non-symmetric shape of amino acids, and not by the backbone or local secondary structure.

However, materials that share the same packing fraction do not necessarily possess the same internal structure. In this article, we characterize the void space in experimentally obtained and computationally generated protein cores to further test the geometric similarities between these two systems. We show below that dense random packings of amino acid-shaped particles have the same local packing fraction, void distribution, and percolation of connected void space as protein cores, which indicates structural equivalence.

Our results suggest that the computationally generated packings can be used as mechanical analogs of protein cores to predict their collective mechanical response. Further, our results em-

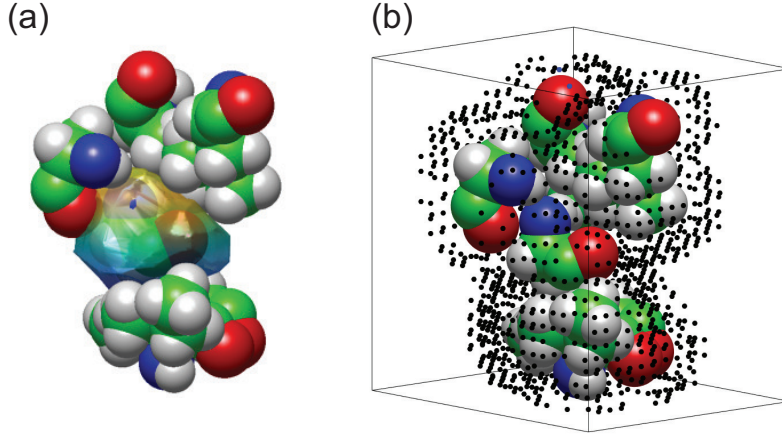


Figure 2.1: Visualization of (a) local and (b) connected voids from the same computationally generated packing of $N = 64$ amino acid-shaped particles. Only the central Alanine (Ala) with the neighboring Alanine, 2 Isoleucines (Ile), Leucine (Leu), and Valine (Val) are shown for clarity. The neighboring amino acids share at least one common surface Voronoi cell face with the central Ala. In (a), the central Ala is enclosed by its surface Voronoi cell. In (b), the connected void space is visualized using points on a grid. For clarity only 75% of the points are shown, and the grid spacing ($g = 0.7\text{\AA}$) is large compared to values used in the text. In both (a) and (b), the atoms are colored as follows: C (green), O (red), N (blue), and H (white). See Fig. 2.8 for visualizations of the connected void space throughout the entire simulation domain.

phasize the connection between structurally arrested, yet thermally fluctuating, protein cores and the jamming transition of highly nonspherical particles [22]. Although the similarity between structural glasses and proteins at low temperatures has been known for several decades [23–27], prior computational studies have mainly focused on the transition from harmonic to anharmonic conformational fluctuations on length scales spanning the full protein. In contrast, our studies identify key structural similarities between jammed packings of amino acid shaped particles and the cores of protein crystal structures.

This article is organized into four sections and three appendices. In Sec. 2.2, we describe the database of high-resolution protein crystal structures that we use for our structural analyses and the computational methods we use to generate jammed packings of amino acid-shaped particles. We also outline two methods to measure the void distribution in the two systems: a local measure of void space using surface Voronoi tessellation, and a non-local or “connected” measure of void space similar to that used by Kertész [28] and Cuff and Martin [29]. In Sec. 2.3, we compare the results of both the local and connected void measurements for jammed packings of amino acid-shaped particles and protein cores and find that both void measurements are the same for both systems. In Sec. 2.3.1, we show that the Voronoi cell volume distributions in both systems are described by a k -gamma distribution with similar shape factors k . In addition, we find that the distribution of the local porosity ($\eta = 1 - \phi$) is the same for protein cores and jammed packings

of amino acid-shaped particles. In Sec. 2.3.2, we identify the void percolation transition as a function of the probe particle accessibility for the connected voids, and find that protein cores and jammed packings of amino acid-shaped particles share the same critical probe size that separates the percolating and non-percolating regimes. In Sec. 2.3.3, we investigate the critical properties of this percolation transition, and show that they are similar to void percolation of systems of randomly placed spheres, but distinct from void percolation in jammed sphere packings. In Sec. 2.4, we summarize our results, discuss their importance, and identify future research directions. We include three appendices with additional details of our computational methods. In Appendix 2.5.1, we provide details for the computational method we use to generate jammed packings of amino acid-shaped particles. In Appendix 2.5.2, we discuss the differences between protein cores in the Dunbrack 1.0 database, and the core replicas we generate from jammed packings of amino acid-shaped particles. In Appendix 2.5.3, we discuss the differences between the connected void cluster size distributions in the systems considered in Sec. 2.3.3.

2.2 Methods

To benchmark our studies of local and connected void regions, we use a subset of the Dunbrack PISCES Protein Database (PDB) culling server [30, 31] of high-resolution protein crystal structures. This dataset, which we will refer to as “Dunbrack 1.0”, contains 221 proteins with $< 50\%$ sequence identity, resolution ≤ 1.0 Å, side chain B factors per residue ≤ 30 Å² and R factor ≤ 0.2 . We add hydrogen atoms to each protein crystal structure using the REDUCE software [32]. To determine core amino acids, we calculate the solvent accessible surface area (SASA) for each residue using the NACCESS software [33] with a 1.4 Å water molecule-sized probe [34]. To compare the SASA for residues with different sizes, we calculate the relative SASA (rSASA), which is the ratio of the SASA of the residue in the protein context to that of the residue *outside* the protein context, along with the C_α , C, and O atoms of the previous amino acid in the sequence and the N, H, and C_α atoms of the next amino acid in the sequence. We define core residues as those with $rSASA \leq 10^{-3}$, and we define a protein core as a set of core residues that each share at least one Voronoi cell face (defined below) with each other. We find similar results if the threshold for defining a core residue is smaller, although there will be fewer “core” residues. We showed in previous work that the local packing fraction decreases significantly for residues with $rSASA > 0.05$ [34]. See Fig. 2.2 (a) for an example core region in a protein from the Dunbrack 1.0 database, and Fig. 2.9 for the size distribution of protein cores in the Dunbrack 1.0 database.

We will compare the structural properties of the cores of protein crystal structures and jammed packings [35] of amino acid-shaped particles. In previous studies, we found that the packing fraction of core regions in proteins is $\phi \approx 0.56$, which is the same as that of jammed packings of purely repulsive amino acid-shaped particles *without* backbone constraints [20, 36]. Here, we will focus exclusively on packing the hydrophobic residues: Ala, Leu, Ile, Met, Phe, and Val. The amino acid-shaped particles will include the backbone atoms N, C_α , C, and O, as well as all of the side chain atoms, with the atomic radii given in Ref. [20], which recapitulate the side chain dihedral angles of residues in protein cores. The packings of amino acid-shaped particles contain mixtures of Ala, Leu, Ile, Met, Phe, and Val residues, with each residue treated as a purely repulsive, rigid

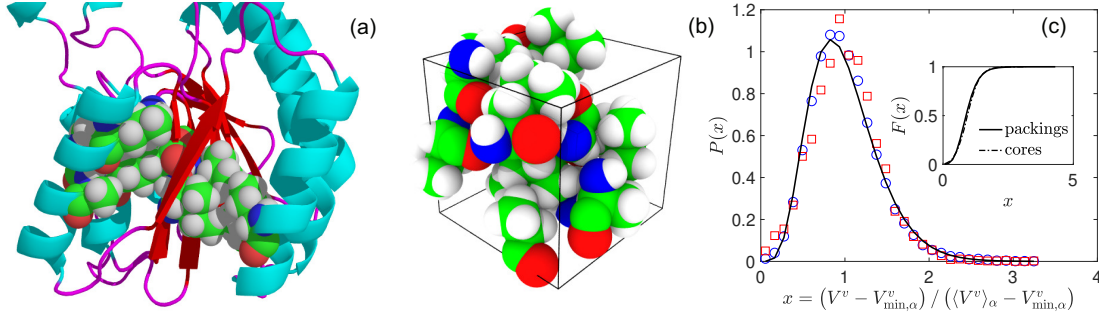


Figure 2.2: (a) Core residues in an example globular protein (PDB code: 3F1L). Non-core regions are drawn using the ribbon representation, and the 11 core amino acids are drawn in all-atom representation. (b) Jammed packing of the same 11 core residues in (a). (c) The surface Voronoi cell volume V^v distribution plotted as a function of $x = (V^v - V_{\min,\alpha}^v) / (\langle V^v \rangle_\alpha - V_{\min,\alpha}^v)$ and fit to a k -gamma distribution (black line) with $k = 6.06 \pm 0.08$ and 5.29 ± 0.27 for packings of amino acid-shaped particles (circles) and protein cores (squares), respectively. $\langle V^v \rangle_\alpha$ is the average and $V_{\min,\alpha}^v$ is the minimum SV cell volume of residue type α . The inset of (c) is the cumulative distribution function $F(x)$ for the data in the main panel.

body composed of a union of spherical atoms with fixed bond lengths, bond angles, and side-chain and backbone dihedral angles taken from instances in the Dunbrack 1.0 database.

We choose which residues are included in each packing using two methods. For method 1 (M1), we generate $\mathcal{C} = 20$ jammed packings of the exact residues found in each distinct protein core in the Dunbrack 1.0 database. For example, if protein X has a core with R residues, we produce \mathcal{C} jammed packings of those exact R residues. If r of these R residues are not one of the hydrophobic residues we consider, these residues are removed and a jammed packing is generated with the remaining $R - r$ residues. This method seeks to mimic the core size and amino acid frequency distribution found in the Dunbrack 1.0 database. Note that we do not remove r residues from the corresponding protein core in the Dunbrack 1.0 database; non-hydrophobic residues are only excluded from the initial conditions used to generate packings of amino acid-shaped particles. In method 2 (M2), we randomly select hydrophobic residues from the Dunbrack 1.0 database with frequencies set by the fraction of each amino acid type found in the Dunbrack 1.0 database. The frequencies are 0.29 (Ala), 0.19 (Leu), 0.17 (Ile), 0.05 (Met), 0.07 (Phe) and 0.23 (Val). In method 2, the identities of the residues in the jammed packings only match those in protein cores on average.

We now briefly describe the computational method for generating jammed packings of amino acid-shaped particles. We use a pairwise, purely repulsive linear spring potential to model inter-residue interactions. Because the residues are rigid particles with each composed of a union of spheres, we test for overlaps between residues μ and ν by checking for overlaps between all atoms i on residue μ and all atoms j on residue ν , respectively. Note that this potential is isotropic and depends only on the distances between atoms on different residues. (See Eq. (2.14) in Appendix 2.5.1.)

We place N residues with random initial positions and orientations at packing fraction $\phi_0 = 0.40$

in a cubic simulation box with periodic boundary conditions and then increase the packing fraction in small steps $\Delta\phi$ to isotropically compress the system. After each compression step, we relax the total potential energy using FIRE energy minimization [37]. This method is similar to a “fast” thermal quench that finds the nearest local potential energy minimum. We use quaternions to track the particle orientations for each residue, as described in Ref. [38]. If the total potential energy per residue is zero after energy minimization, i.e. $U/N\epsilon < 10^{-8}$, where ϵ is the energy scale of the atomic interactions, we continue to increase the packing fraction. If the total potential energy per residue is nonzero, i.e. $U/N\epsilon \geq 10^{-8}$ and residues have small overlaps, we decrease the packing fraction. The packing fraction increment $\Delta\phi$ is halved each time the algorithm switches from compression to decompression and vice versa. We terminate the packing-generation protocol when the residue packings satisfy $10^{-8} < U/N\epsilon < 2 \times 10^{-8}$ and possess a vanishing kinetic energy per residue (i.e. $K/N\epsilon < 10^{-20}$) [22]. (An example jammed packing of amino acid-shaped particles is shown in Fig. 2.2 (b) and further computational details are included in Appendix 2.5.1.)

To measure the distribution of local voids in packings of amino acid-shaped particles and protein cores, we use a Voronoi tessellation, which ascribes to each particle the region of space that is closer to that particle than all other particles in the system. For residues, which are highly non-spherical particles, we use a generalization of the standard Voronoi tessellation known as the *surface-* or *set-*Voronoi (SV) tessellation [39]. This tessellation partitions the empty space in the system using a bounding surface for each residue. An efficient algorithm to generate this tessellation is outlined in Ref. [39] and implemented using POMELO [40]. To construct the SV tessellation, consider a set of N particles with bounding surfaces $\{\partial K_\mu\}$ for $\mu = 1, \dots, N$. The software approximates ∂K_μ by triangulating points on the particle surfaces, and uses the standard Voronoi tessellation of the surface points to construct the SV cell for each residue μ . We find that using 400 surface points per atom, or ≈ 6400 surface points per residue, gives an accurate representation of the SV cell, which does not change significantly as more surface points are added. An example SV cell from a packing of amino acid-shaped particles is shown in Fig. 2.1 (a). For an SV cell with volume V_μ^v surrounding residue μ with volume v_μ , the local porosity is given by:

$$\eta_\mu = \frac{V_\mu^v - v_\mu}{V_\mu^v} = 1 - \phi_\mu, \quad (2.1)$$

where $\phi_\mu = v_\mu/V_\mu^v$ is the local packing fraction. This quantity measures the local *void* space associated with each residue.

We also quantify the “connected” void space shared between residues in packings of amino acid-shaped particles and protein cores. To do this, we implement a grid-based method similar to that described by Kertész [28] and Cuff and Martin [29], where the “void space” is defined as the region of a system accessible to a spherical probe particle with radius a . The geometry and distribution of void space in a system is thus a function of a , the residue positions \vec{r}_μ , and bounding surfaces ∂K_μ . We define a cubic lattice with G points in each direction within the simulation domain, which gives a lattice spacing $g = L/G$. For all lattice points \mathbf{p} , we define the set of void points \mathcal{V} to be all points that can accommodate a spherical probe particle with radius a without causing overlaps with any atoms. We label all void points with a 1, and all other points with a 0. After all grid points are

labeled, we use the Newman-Ziff algorithm [41] to cluster adjacent, similarly-labeled grid points. We consider all adjacent points on the nearest face, edge, and vertex of a cube of points surrounding each lattice point (i.e. next-to-next-to-nearest neighbor counting with 26 possible adjacencies for each point) when merging void clusters and implement periodic boundary conditions. A sketch of connected void lattice points in a subset of a packing of amino acid-shaped particles is shown in Fig. 2.1 (b).

When measuring void space in protein structures, we implement a similar procedure, but we only consider voids in core residues. We construct a box of dimension $L_x \times L_y \times L_z$ that circumscribes each protein core, with the box just outside the radii of core residues near the box edges. We pick a spherical probe particle of radius a , and label the void space as all points that are (a) not contained inside an atom, and (b) contained only within the union of the SV cells of core residues. With these constraints, we only consider connected void space specific to the core of the protein. We then use the Newman-Ziff algorithm to merge void clusters, and repeat the procedure for 100 different random protein orientations.

2.3 Results

2.3.1 Local Void Analysis

We begin with an analysis of local voids associated with each amino acid in jammed packings of amino acid-shaped particles and protein cores. We measure the distribution of the SV cell volumes and show that the distributions in both systems can be fit to a k -gamma distribution, which also describes Voronoi cell distributions in jammed packings of spheres [42, 43], ellipsoids [44], attractive emulsion droplets [45], wet granular materials [46], and model cell monolayers [47]. The k -gamma distribution for the SV cell volume V_μ^v for each residue has the form:

$$P(x) = \frac{k^k}{\Gamma(k)} x^{k-1} \exp(-kx), \quad (2.2)$$

where $x = (V_\mu^v - V_{\min,\alpha}^v) / (\langle V_\mu^v \rangle_\alpha - V_{\min,\alpha}^v)$, which sets the scale factor of the distribution to 1. Here,

$$\langle V_\mu^v \rangle_\alpha = \frac{1}{N_\alpha} \sum_{\mu=1}^{N_\alpha} V_\mu^v \quad (2.3)$$

is the average SV cell volume of residue type α . The sum involving μ is over all N_α residues of type α in all packings, and $V_{\min,\alpha}^v$ is the minimum SV cell volume of residue type α . We consider minima and averages for each residue type separately to account for the large differences in residue volumes; that is, each residue type α , when considered individually, has a SV cell volume distribution described by Eq. (2.2).

We measure the shape factor k_α for each residue type α either by fitting the SV cell volume

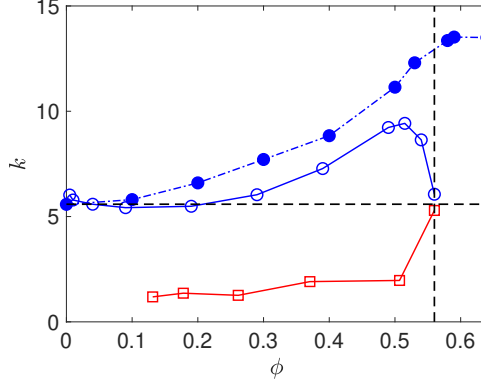


Figure 2.3: The shape parameter k for fits of the k -gamma distribution [Eq. (2.2)] to the SV cell volume distributions $P(x)$ for packings of amino acid-shaped particles (open circles), monodisperse spheres (filled circles), and both core and surface residues in the Dunbrack 1.0 database (open squares) as a function of packing fraction ϕ . The dashed horizontal line at $k = 5.59$ is the analytical value of the shape factor for the Voronoi cell volume distribution of a random Poisson point process [48], and the dashed vertical line at $\phi_J = 0.56$ is the packing fraction for protein cores and jammed packings of amino acid-shaped particles.

distribution to Eq. (2.2) using Maximum Likelihood Estimation (MLE), or by calculating

$$k_\alpha = \frac{(\langle V_\mu^v \rangle_\alpha - V_{\min, \alpha}^v)^2}{\langle (V_\mu^v)^2 \rangle_\alpha - \langle V_\mu^v \rangle_\alpha^2}. \quad (2.4)$$

We obtain similar k -values using both methods. Although the values of k_α depend on the type of amino acid α , when we average the values of k_α we recover the value of k obtained from fitting the combined distribution. We focus on the distributions of SV cell volumes averaged over all hydrophobic residues.

In Fig. 2.2 (c), we show the SV cell volume distributions $P(x)$ for packings of core amino acid-shaped particles modeled after specific protein cores (method M1) and for all core residues in the Dunbrack 1.0 database. We find that the distributions for these two systems are similar; both obey a k -gamma distribution [Eq. (2.2)] with similar shape parameters, $k = 6.06 \pm 0.08$ and $k = 5.29 \pm 0.27$, for core residues in the Dunbrack 1.0 database and packings of amino acid-shaped particles, respectively. As expected, the cumulative distributions $F(x)$ of the SV cell volumes for residues in protein cores and packings of amino acid-shaped particles are also nearly indistinguishable.

The strong similarity between the SV cell volume distributions indicates that jammed packings of amino acid-shaped particles (at $\phi_J \approx 0.56$) and protein cores possess the same underlying structure. To better understand this result, in Fig. 2.3 we plot the shape parameter k that describes the form of the Voronoi cell volume distributions for packings of $N = 10^3$ monodisperse spheres (with $\phi_J \approx 0.64$) and of $N = 64$ amino acid-shaped particles versus ϕ . When $\phi \ll \phi_J$, and the systems

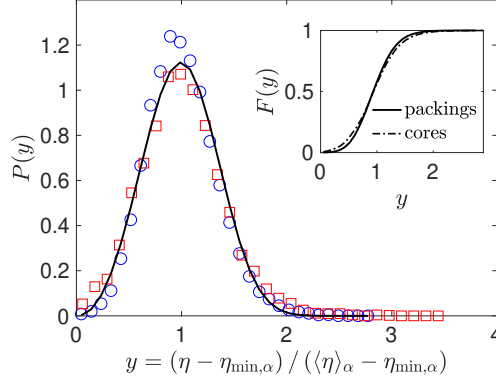


Figure 2.4: Distribution of the scaled local porosity $y = (\eta - \eta_{\min,\alpha}) / (\langle \eta \rangle_\alpha - \eta_{\min,\alpha})$, where $\langle \eta \rangle_\alpha$ is the average and $\eta_{\min,\alpha}$ is the minimum porosity of residue type α , for packings of amino acid-shaped particles (circles) and residues in protein cores in the Dunbrack 1.0 database (squares). The solid line is a Weibull distribution with shape parameter $b \approx 3.2$ [Eq. (2.7)]. The inset is the cumulative distribution function $F(y)$ of the data in the main panel.

are sufficiently dilute, the Voronoi cell volume distributions of the packings of monodisperse spheres and amino acid-shaped particles resemble that for a random Poisson point process [48] with $k \approx 5.6$. In this regime, free volume is assigned randomly to each particle since the particle positions are uncorrelated. However, as ϕ increases, the k -values for packings of monodisperse spheres and amino acid-shaped particles begin to grow, but at different rates, since the particle geometry becomes important in determining the local free volume. Near $\phi \simeq \phi_J$, the shape parameter plateaus at $k \approx 13$ for packings of monodisperse spheres, but the shape parameter decreases strongly to $k \approx 6$ for packings of amino acid-shaped particles. This decrease in k indicates a transition from having the shape of the Voronoi cell volume distribution determined by spherical particles (for $\phi \lesssim \phi_J$) to that determined by bumpy, asymmetric amino acid-shaped particles (for $\phi \simeq \phi_J$). Note, however, that the SV cell volume distribution of *jammed* packings of amino acid-shaped particles is similar (in terms of k value) to that of *randomly* placed Poisson points. This suggests that the void distribution of jammed packings of amino acid-shaped particles and protein cores share structural properties with randomly placed points. We will expand on this similarity in Sec. 2.3.3.

In addition, we calculate k for the SV cell volume distributions for residues in the Dunbrack 1.0 database as a function of packing fraction. In previous studies, we have found a one-to-one correlation between solvent accessibility and packing fraction [34]; residues with lower values of ϕ in Fig. 2.3 are therefore more solvent-exposed, i.e. closer to the protein surface. For most of the range in ϕ , $k \approx 2$, whereas $k \gtrsim 5.6$ for packings of monodisperse spheres and amino acid-shaped particles. In particular, k does not equal the value for a random Poisson point process ($k = 5.6$) in the limit $\phi \ll \phi_J$ for residues in protein cores. In protein cores, the backbone constraint gives rise to correlations in the residue positions. However, as $\phi \rightarrow \phi_J$, k increases, reaching $k \approx 6$ when $\phi = \phi_J$. This result shows that there is a fundamental change in the SV cell distribution near

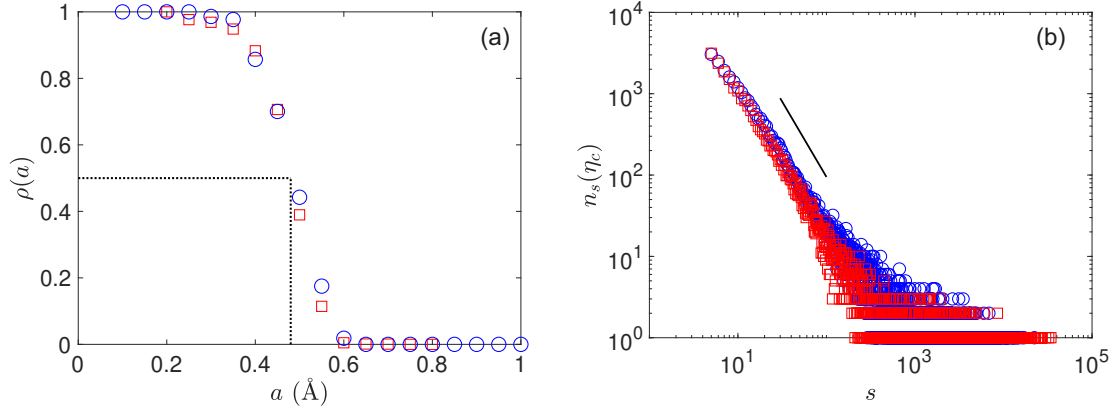


Figure 2.5: (a) Percolation probability $\rho(a)$ plotted versus the probe radius a for protein cores from the Dunbrack 1.0 database (crosses) and clusters of core residues extracted from static packings of $N = 64$ amino acid-shaped particles (circles). The horizontal and vertical dashed lines indicate the critical probe radius $a_c = 0.48$ Å that satisfies $\rho(a_c) = 0.5$. (b) Cluster size distribution n_s with size s at the critical porosity η_c , which scales as $n_s(\eta_c) \sim s^{-\tau}$. The Fisher power-law exponent $\tau = 1.95 \pm 0.06$ and 1.85 ± 0.05 for protein cores from the Dunbrack 1.0 database (crosses) and representative clusters of core residues in packings of amino acid-shaped particles (circles), respectively. The solid line has slope equal to -1.85.

the onset of jamming in protein cores. For $\phi \lesssim \phi_J$, the backbone determines the shape of the SV cell volume distribution, whereas for $\phi \rightarrow \phi_J$, the shapes of the amino acids determine the SV cell volume distribution.

We also compare the local porosity distributions for protein cores and packings of amino acid-shaped particles in Fig. 2.4. We scale the porosity (as in Eq. (2.2)) by defining

$$y = \frac{\eta_\mu - \eta_{\min, \alpha}}{\langle \eta_\mu \rangle_\alpha - \eta_{\min, \alpha}}, \quad (2.5)$$

where

$$\langle \eta_\mu \rangle_\alpha = \frac{1}{N_\alpha} \sum_{\mu=1}^{N_\alpha} \eta_\mu, \quad (2.6)$$

and $\eta_{\min, \alpha}$ is the minimum porosity over all N_α core residues of type α . Again, the porosity distributions $P(y)$ (and cumulative distributions $F(y)$) for residues in protein cores and packings of amino acid-shaped particles are similar, but here $P(y)$ has the shape of a Weibull distribution with scale factor $\lambda = 1$,

$$P(y) = by^{b-1} \exp(-y^b). \quad (2.7)$$

where b is the shape parameter of the Weibull distribution.

The small differences in $P(x)$ and $P(y)$ between core residues in protein crystal structures and packings of amino acid-shaped particles can be explained by the small differences between the

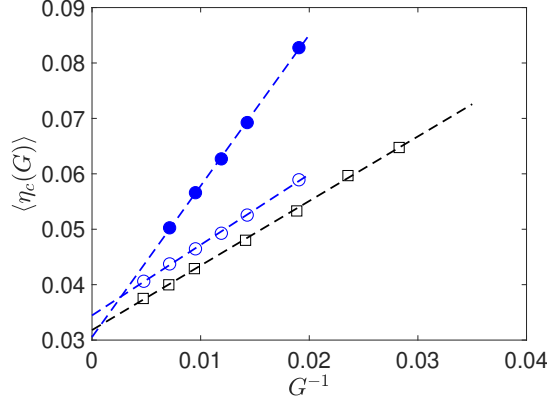


Figure 2.6: Critical porosity $\langle \eta_c(G) \rangle$ using a lattice with G points along each dimension plotted versus G^{-1} for jammed packings of $N = 64$ amino acid-shaped particles with $N_a = 1024$ atoms (open circles), $N = 10^3$ randomly placed spheres (open squares), and $N = 10^3$ monodisperse spheres (filled circles). The dashed lines have vertical intercepts that indicate $\eta_c(G \rightarrow \infty) \approx 0.0345$, 0.0318 , and 0.0305 for packings of amino acid-shaped particles, randomly placed spheres, and monodisperse sphere packings, respectively.

volumes of core residues in crystal structures and in packings. The atoms on neighboring amino acids interact differently for free amino acids in packings versus backbone atoms in protein cores, which form covalent and hydrogen bonds. Thus, we find that the volumes of residues in protein cores have larger variances and smaller means than those in packings of amino acid-shaped particles. Also, the overlaps between covalently bonded backbone atoms that link adjacent residues slightly decreases the mean SV cell volume, which gives rise to a larger population of small SV cells and a small deviation between $P(x)$ for residues in protein cores and in packings for small x in Fig. 2.2 (c).

2.3.2 Connected void analysis of protein cores

We next quantify the distribution of “connected” void space that is shared between residues. Using a grid-based method, we calculate the volume of regions of connected void space as a function of the radius a of a spherical probe particle. As we increase a , the connected void space transitions from highly connected throughout the system to compact and localized with distinct void regions. We measure the probability $\rho(a)$ of finding a percolating void region, where we define percolation as the appearance of a cluster that spans one of the system dimensions when the boundary is closed, and a cluster that both spans, wraps around the boundary, and self-intersects when the boundaries are periodic. We identify the critical probe radius a_c by setting $\rho(a_c) = 0.5$. Because the definition of connected void regions depends on the boundary condition, the value of a_c , especially in systems as small as protein cores, is affected by the boundary conditions. Thus, to calculate $\rho(a)$, we create packings of amino acid-shaped particles with similar boundary conditions as those in protein cores.

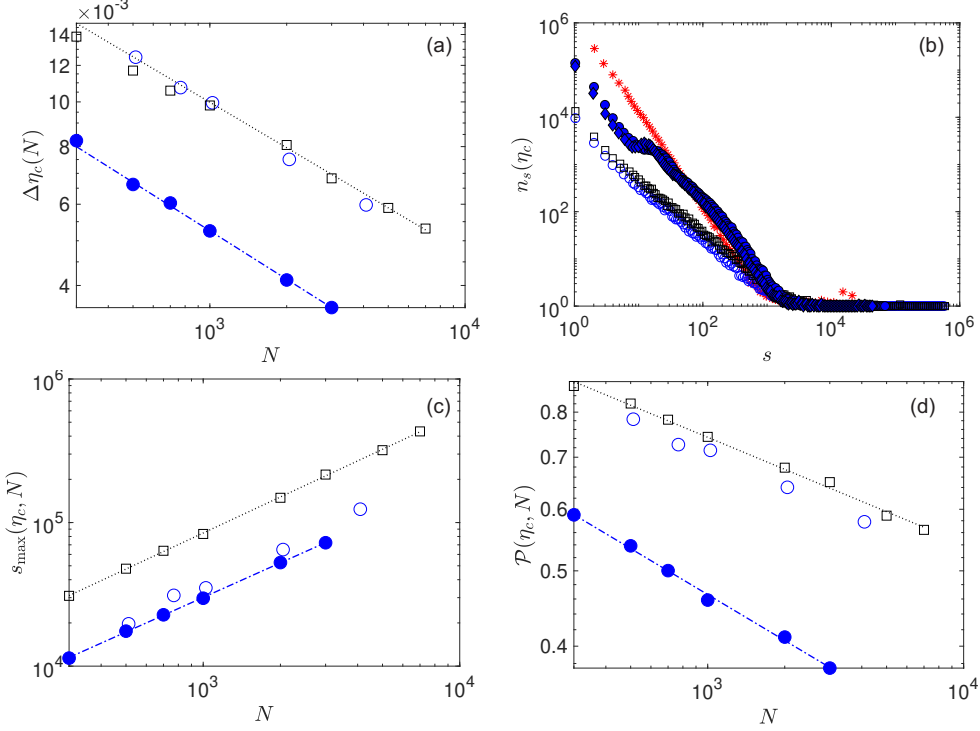


Figure 2.7: (a)-(d) Scaling behavior for jammed packings of amino acid-shaped particles (open circles), bidisperse spheres (filled circles), and randomly placed spheres (open squares). In (a), we show that the standard deviation in the critical porosity scales as $\Delta\eta_c(N) \sim N^{-1/d\nu}$, where ν is the correlation length exponent. The lines have slopes -0.33 (dotted line) and -0.31 (dot-dashed line). In (b), we show that for randomly placed spheres and jammed packings of amino acids, the distribution of connected void clusters of size s at the percolation threshold scales as $n_s(\eta_c) \sim s^{-\tau}$, with a Fisher exponent $\tau \approx 1.25$. In addition, we plot n_s at criticality for standard site percolation on a cubic lattice (red stars), which has an apparent $\tau \approx 2.07$. We also plot the same distribution for the connected voids in jammed packings of bidisperse and monodisperse spheres (filled diamonds), which do *not* display power-law scaling. In (c), we show that the maximum cluster size near the percolation onset scales as $s_{\max}(\eta_c, N) \sim N^{D/d}$, where D is the fractal dimension. The lines have slopes 0.83 (dotted line) and 0.82 (dot-dashed line). In (d), we show that the probability for a given site to be in the percolating void cluster at η_c scales as $\mathcal{P}(\eta_c, N) \sim N^{-\beta/d\nu}$, where β is the percolation strength scaling exponent. The lines have slopes -0.14 (dotted line) and -0.19 (dot-dashed line).

From a packing of amino acid-shaped particles with periodic boundary conditions ($N = 64$, method M2), we extract a representative protein core of $R - r$ residues that all share at least one SV cell face. We sample $R - r$ from the distribution of core sizes $P(R)$ found in the Dunbrack 1.0 database. (See Fig. 2.9 in Appendix 2.5.1.) The resulting packings have boundary conditions similar to protein cores in the Dunbrack 1.0 database. We then determine the connected void regions as a function of a and identify the critical probe size a_c as shown in Fig. 2.5 (a). We find the same critical probe size $a_c = 0.48 \pm 0.01$ Å for both protein cores and packings of amino acid-shaped particles with similar boundary conditions. Note that this value of the critical probe radius is smaller than that of a water molecule, which is ≈ 1.4 Å, and thus the voids we consider here are not accessible by aqueous solvents. However, as we discuss below, this value of the probe radius corresponds to a critical point; we will exploit the behavior of the voids near this critical point to understand the geometric properties of the connected voids, and to differentiate between the voids in various systems.

Thus, determining the connected void regions in protein cores is a type of percolation problem. In lattice site percolation, sites on a lattice in d spatial dimensions are either occupied randomly with probability p or not occupied with probability $1 - p$. At the percolation threshold p_c , adjacent occupied sites form a percolating cluster that spans the system and becomes infinite in the large-system limit. Continuum percolation occurs in systems that are not confined to a lattice. Both particle contact and void percolation have been studied in randomly placed overlapping spheres [28, 49, 50] and percolation of particle contacts [51, 52] has been studied in packings of repulsive [53] and adhesive particles [54].

In this article, we consider percolation of the void space accessible to a spherical probe particle with radius a in packings of spheres and amino acid-shaped particles, as well as systems composed of randomly placed spheres [49, 50]. As the probe particle radius is increased, the amount of space available to the probe is restricted and the number of void lattice sites decreases. We define an effective porosity η as the ratio of the number of void lattice sites to the total number of lattice sites G^d . We determine the percolation threshold using a bisection method, where we begin with two initial guesses for the percolation transition, a_H and a_L with $a_H > a_L$, and iteratively check for percolation of void sites at the probe radius $a = (a_H + a_L)/2$. We set $a_H = a$ if we find a percolated cluster of void sites, and $a_L = a$ if we do not find a percolated cluster. We terminate the algorithm when the difference between successive values for a_c are within a small tolerance $\delta a = 10^{-8}$ Å. Note that our use of a lattice of points to measure the connected void region does not imply that our model is a lattice model. The lattice is simply a tool to calculate the connected void space volume [28]. Furthermore, in the continuum limit (i.e. $G \rightarrow \infty$), we recover the critical porosity $\eta_c \approx 0.03$ measured using Kerstein’s method [49, 50] on systems of randomly placed spheres [55]. (See Fig. 2.6.) Since there is a one-to-one mapping between a and η , we will use η as the order parameter for continuum void percolation.

System	ν	τ	D	β	$d - \frac{\beta}{\nu}$
residue packings (full)	0.90 ± 0.24	1.29 ± 0.06	2.58 ± 0.18	0.37 ± 0.19	2.58 ± 0.33
residue packings (rep.)	—	1.85 ± 0.05	—	—	—
Protein cores, Dunbrack 1.0	—	1.95 ± 0.06	—	—	—
Mono. Spheres (jammed)	1.05 ± 0.12	—	2.46 ± 0.09	0.60 ± 0.07	2.43 ± 0.15
Bidis. Spheres (jammed)	0.93 ± 0.10	—	2.40 ± 0.08	0.56 ± 0.07	2.40 ± 0.14
Cubic Lattice [56]	0.91 ± 0.04 (0.88)	2.07 ± 0.01 (2.18)	2.49 ± 0.03 (2.53)	0.48 ± 0.02 (0.42)	2.47 ± 0.04
Poisson Spheres (Kertész's method [28])	1.10 ± 0.06	1.22 ± 0.05	2.51 ± 0.03	0.44 ± 0.03	2.60 ± 0.05
Poisson Spheres (Kerstein's method [49, 50])	0.99 ± 0.05 (0.902 ± 0.005 [50])	—	2.44 ± 0.03	0.48 ± 0.02 (0.45 ± 0.2 [49])	2.52 ± 0.05

Table 2.1: Table of critical exponents ν , τ , D , and β for several models of void percolation. In the last column, we provide the value for the hyperscaling relation, $d - \frac{\beta}{\nu}$, which matches the fractal dimension D if hyperscaling is satisfied. In the first four rows, we report the critical exponents for packings of amino acid-shaped particles with periodic boundary conditions (full) and boundary conditions representative of protein cores (rep.). We also report the critical exponents for void percolation in jammed packings of monodisperse (Mono.) and bidisperse (Bidis.) spheres. In the last four rows, we compare these results to those for void percolation in several systems that were studied previously. We report our measurements of the critical exponents for site percolation on a cubic lattice, where only nearest-neighbors are counted as adjacent sites. We also report the critical exponents for void percolation and Voronoi vertex percolation in systems composed of randomly placed spheres. Previously reported values of the exponents are given in parentheses. Throughout the table, error bars are obtained from the bootstrap method, where we randomly, independently sample 20% of the data over 200 trials, and fit the exponents over each individual trial. Values for the exponents are the average over trials, and the error bars are standard deviations.

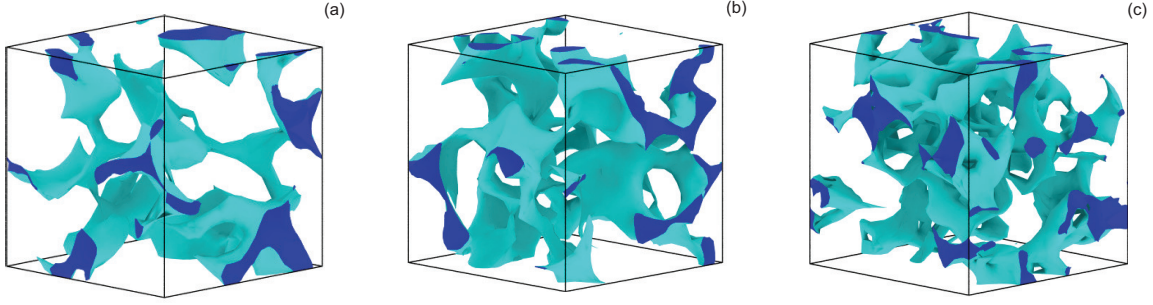


Figure 2.8: Visualization of the surface of a connected void region (light domain) at the percolation threshold $\eta_c \approx 0.03$. The dark domains are the “inside” of the void region, which connects across the periodic boundaries. These three systems are (a) a jammed packing of $N = 16$ amino acid-shaped particles, with 298 atoms in total, (b) $N = 300$ randomly placed spheres, and (c) a jammed packing of 300 bidisperse spheres.

2.3.3 Connected void analysis of packings

We now focus on the statistical properties of the connected void regions in packings of spheres and amino acid-shaped particles prepared in systems with cubic, periodic boundaries. In Fig. 2.7 and Table 2.1, we summarize the results of this analysis, and we show visualizations of the percolating connected voids in jammed packings of spheres and amino acid-shaped particles in Fig. 2.8. We first measure the correlation length exponent ν , where the correlation length ξ is defined as the average distance between two points in the largest connected void cluster. Near η_c , ξ diverges as $|\eta - \eta_c|^{-\nu}$. Using finite-size scaling [56], we can write

$$\eta_c(N) - \eta_c(\infty) \sim N^{-1/d\nu}, \quad (2.8)$$

where $\eta_c(\infty)$ is the percolation threshold in the large-system limit and $N \sim L^d$. $\eta_c(N)$ is a random variable with standard deviation $\Delta\eta_c(N)$, which will approach $\eta_c(\infty)$ as $N \rightarrow \infty$. Thus, we make the ansatz that

$$\Delta\eta_c(N) \sim N^{-1/d\nu}, \quad (2.9)$$

which can be used to measure ν . (See Fig. 2.7 (a).) We also measure the Fisher exponent τ , defined by

$$n_s(\eta_c) \sim s^{-\tau}, \quad (2.10)$$

where n_s is the number of void clusters containing s sites. While we measure this exponent for protein cores and random packings with representative boundary conditions in Fig. 2.5 (b), in Fig 2.7 (b) we measure this exponent in systems with cubic, periodic boundary conditions.

We also measure the fractal dimension and percolation strength of the percolating void clusters. The fractal dimension is defined by

$$s_{\max}(\eta_c, N) \sim N^{D/d}, \quad (2.11)$$

where $s_{\max}(\eta_c, N)$ is the number of sites contained in the largest void cluster in the system at percolation onset. If $D = d$, the largest void cluster is a compact, non-fractal object. However, if $D < d$, the void cluster is fractal [57]. (See Fig. 2.7 (c).) The percolation strength is the probability $\mathcal{P}(\eta)$ that a given lattice site is part of the percolating void cluster at a given porosity. Near η_c , the probability scales as $\mathcal{P}(\eta) \sim |\eta - \eta_c|^\beta$. The probability obeys finite size scaling,

$$\mathcal{P}(\eta_c, N) \sim N^{-\beta/d\nu}. \quad (2.12)$$

Once we determine ν using Eq. (2.9), we can determine β from Eq. (2.12). (See Fig. 2.7 (d).) We also expect β , ν and D to satisfy the hyperscaling relation,

$$D = d - \frac{\beta}{\nu}. \quad (2.13)$$

In Table 2.1, we report our measurements for the critical exponents ν , τ , D , and β for void percolation (using a spherical probe particle), as well as for $d = 3$ lattice site percolation on a cubic lattice and void percolation in systems of randomly placed spheres using two methods: the connected void method described previously and the Voronoi vertex method introduced by Kerstein [49] and implemented by Rintoul [50]. Note that protein cores and representative subsets of jammed packings of amino acid-shaped particles (denoted “rep.”) are small systems with $N < 30$, and thus we cannot use finite-size scaling to measure the critical exponents. We can, however, measure the critical exponents for full packings of amino acid-shaped particles (denoted “full”), which mimic the geometric properties of void clusters in protein cores.

We observe that across all models and methods studied, the correlation length exponent $\nu \approx 0.9$ -1.0 for void percolation. In particular, $\nu \approx 0.90$ for packings of amino acid-shaped particles is similar to that (0.90) for randomly placed spheres [50], as well as for standard site percolation [56]. In addition, the fractal dimension $D \approx 2.4$ -2.6 is similar for all models and methods for calculating void percolation. We find that the percolation strength exponent $\beta < 0.5$ for randomly placed spheres and packings of amino acid-shaped particles when using the connected void method, but $\beta > 0.5$ for packings of monodisperse and bidisperse spheres. (The bidisperse systems include $N/2$ large and $N/2$ small spheres with diameter ratio $d = 1.4$.)

However, because of the limited range of system sizes studied here, it is difficult to determine the critical exponents with high precision. Thus, given the results for the ν , D , and β exponents alone, it is difficult to distinguish the statistical properties of the void content of packings of jammed spheres, randomly placed spheres, and jammed amino acid-shaped particles from each other, or from void percolation on a cubic lattice, for that matter.

We do see a strong distinction in the Fisher exponent τ [Eq. (2.10)] between void percolation on a cubic lattice and in packings of amino acid-shaped particles. For these two systems, $\tau = 2.07 \pm 0.01$ and $\tau = 1.29 \pm 0.06$, respectively. For packings of randomly placed spheres, we find that the Fisher exponent is $\tau = 1.22 \pm 0.05$, which is similar to our result for packings of amino acid shaped particles. (See Table 1.) These values for τ were obtained from a cubic lattice with $G = 100$ sites per box length, packings of amino acid-shaped packings with $N = 128$ particles and, on average, $N_a = 2048$ atoms, and systems of $N = 2000$ randomly placed spheres. These results suggest

that the properties of connected voids are similar in packings of amino acid-shaped particles and randomly placed spheres, and in general that connected voids in these systems are distinct from those for void regions in cubic lattices near percolation onset.

We do not report values of τ for jammed packings of monodisperse and bidisperse spheres, since we observe non-power-law behavior in the cluster size distributions for these systems. As discussed in Appendix 2.5.3, this behavior is most likely due to a residual finite length scale at the percolation threshold. We also observe non-power-law behavior in the cluster size distribution for void percolation in randomly placed spheres using Kerstein’s method, and do not report a value for τ in Table 2.1. As shown in Appendix 2.5.3, this non-power-law behavior is most likely due to the sparsity of the Voronoi vertex network, which truncates the cluster size distribution.

Our results suggest that the critical exponent τ is able to distinguish the geometries of connected void regions in different systems. In particular, the connected void regions in packings of amino acid-shaped particles and systems of randomly placed spheres are similar, but distinct from that for jammed sphere packings. In Fig. 2.8, we show examples of the connected void surface in packings of (a) amino acid-shaped particles, (b) randomly placed spheres, and (c) bidisperse spheres. Qualitatively, the connected void surfaces in systems of randomly placed spheres and amino acid-shaped particles look similar, while the connected void surface in jammed packings of bidisperse spheres looks different, with a characteristic void size.

2.4 Conclusions

In this article, we analyzed local and connected void regions in protein cores and in jammed packings of purely repulsive amino acid-shaped particles and showed that these two systems share the same void structure. We first investigated the surface-Voronoi (SV) cell volume distributions and found that in both systems these distributions are well-described by a k -gamma distribution with $k \approx 6$. This k -value is much smaller than that ($k \approx 13$) obtained for jammed sphere packings, which indicates that packings of amino acid-shaped particles have a broader distribution of Voronoi volumes. We also studied the SV cell volume distribution as a function of the packing fraction, and found that only near the onset of jamming do the SV cell distributions in protein cores and packings of amino acid-shaped particles match. In the dilute case $\phi \ll \phi_J$, the local packing environment in protein cores is determined by the backbone, whereas the local packing environment of packings of free residues resembles a Poisson point process. At jamming onset, the local packing environment is determined by the “bumpy”, asymmetric shape of amino acids, not the backbone constraints.

Using a grid-based method, we also measured the distribution of non-local, connected voids in protein cores and jammed packings of amino acid-shaped particles. We found that when we consider similar boundary conditions in protein cores and jammed packings of amino acid-shaped particles, the two systems also have the same critical probe size a_c (at which the accessible, connected void region spans the system) and Fisher exponent τ (which characterizes the scaling of the size of the void clusters near percolation onset). We also compare the finite-size scaling results for void percolation in packings of amino acid-shaped particles, in packings of monodisperse and bidisperse spheres, and systems of randomly placed spheres. We find that the void percolation critical exponents in packings of amino acid-shaped particles are similar to those in randomly placed

spheres. Specifically, the Fisher exponent τ takes a similar value for these two systems (≈ 1.2 and ≈ 1.3 , respectively), while this exponent is significantly different for void percolation in cubic lattices (≈ 2.1). This result may explain why the distribution of SV cell volumes is similar for jammed packings of amino acid-shaped particles and randomly placed Poisson points with $\phi = 0$, as seen in the horizontal line Fig. 2.3. Interestingly, these results echo similar observations by Liang and Dill, where the authors recognize the similarity between the void distribution of randomly-placed spheres and of protein crystal structures, although they did not connect packing in protein cores with random close packing of nonspherical particles [11].

In future work, we will use jammed packings of amino acid-shaped particles to understand the structural and mechanical response of protein cores to amino acid mutations. We can assess the response in two ways. First, we can prepare jammed packings of amino acid-shaped particles that represent wildtype protein cores, substitute one or more of the wildtype residues with other hydrophobic residues, relax the “mutant” packing using potential energy minimization, and measure the changes in void structure. We can also measure the vibrational density of states (VDOS) in jammed packings that represent the wildtype and mutant cores. The VDOS and the associated eigenmodes can provide detailed information on how the low-energy collective motions change in response to mutations. There are several advantages for calculating the VDOS in jammed packings of amino acid-shaped particles. For example, in jammed packings, only hard-sphere-like steric interactions are included. In contrast, molecular dynamics force fields for proteins typically include many terms in addition to those that enforce protein stereochemistry, which makes it difficult to determine the interactions that control the collective motions. Studying jammed packings of amino acid-shaped particles also decouples the motions of core versus surface residues.

Studies of the VDOS in jammed packings of amino acid-shaped particles will also shed light on the protein “glass” transition, where the root-mean-square (rms) deviations in the atomic positions switch from harmonic to anharmonic behavior [58] in globular proteins near $T_g \approx 200$ K [27]. We will investigate the vibrational response of jammed packings of amino acid-shaped particles to thermal fluctuations. In particular, we will measure the Fourier transform of the position fluctuations and determine the onset of anharmonic response. Additionally, in this work we did not include backbone connectivity between amino acids in our packings, nor did we treat the side chain dihedral angles as “soft” degrees of freedom with harmonic constraints. In future work, to more accurately model the geometrical and topological properties of dynamically fluctuating protein cores, we will incorporate harmonic bond length, bond angle and dihedral angle interactions (for both backbone and side chain atoms), with stiffnesses taken from bond length, bond angle and dihedral angle distributions observed in high-resolution protein crystal structures.

In addition, our analysis of void distributions in protein cores will provide new methods for identifying protein decoys, which are computationally generated protein structures that are not observed experimentally. However, it is currently difficult to distinguish between real structures and decoys. For example, in the most recent Critical Assessment of Protein Structure Prediction (CASP12), researchers were given a set of target sequences, and were tasked with predicting the structures of those sequences using a variety of methods [59]. Each group was allowed to submit 5 structures per target sequence; when tasked with assessing which of their submissions were the most accurate, only 3 groups out of 31 had $> 50\%$ success at identifying the most accurate structure [60].

The average success rate was 30%, just slightly better than guessing at random. Thus, assessing the viability of computationally-designed structures is an incredibly difficult task.

Since the structure of void regions in the cores of protein crystal structures is the same as that found in packings of amino acid-shaped particles, the properties of void regions can serve as a benchmark for ranking computationally designed protein structures. Recent studies have suggested that protein decoys [17] possess local packing fraction inhomogeneities that are not present in protein crystal structures. In addition, the void-based analyses presented here can be used to evaluate the conformational dynamics of proteins sampled in all-atom molecular dynamics simulations. An understanding of the expected void properties from high-resolution protein crystal structures can improve our ability to identify unphysical conformational fluctuations that occur during molecular dynamics trajectories. We propose that detailed characterizations of the void space, using the methods described here, will be a sensitive metric than can be used to assess a variety of protein designs.

2.5 Appendix

2.5.1 Packing-generation Protocol

As described in Sec. 2.2, we generate jammed packings of amino acid-shaped particles using successive small steps of isotropic compression or decompression with each step followed by potential energy minimization. Each residue was modeled as a rigid union of spheres with fixed bond lengths, bond angles, and dihedral angles. Forces between atoms i and j on distinct residues μ and ν were calculated using $\vec{F}_{ij}^{\mu\nu} = -\vec{\nabla}U(r_{ij}^{\mu\nu})$, with the pairwise, purely repulsive linear spring potential energy,

$$U(r_{ij}^{\mu\nu}) = \frac{\epsilon}{2} \left(1 - \frac{r_{ij}^{\mu\nu}}{\sigma_{ij}^{\mu\nu}}\right)^2 \Theta\left(1 - \frac{r_{ij}^{\mu\nu}}{\sigma_{ij}^{\mu\nu}}\right). \quad (2.14)$$

In Eq. (2.14), ϵ is the characteristic energy scale of the repulsive interactions and $\sigma_{ij}^{\mu\nu} = (\sigma_i^\mu + \sigma_j^\nu)/2$, where σ_i^μ is the diameter of atom i on residue μ . The quantity $r_{ij}^{\mu\nu} = |\vec{r}_j^\mu - \vec{r}_i^\nu|$ is the separation distance between atoms i and j on distinct residues μ and ν , and Θ is the Heaviside step function that sets the potential energy to zero when atoms i and j are not in contact. We consider each residue as a rigid body. Forces between pairs of atoms on contacting residues generate torques. The torque on residue μ arising from a force on atom i on residue μ from atom j on residue ν is $\vec{h}_i^\mu \times \vec{F}_{ij}^{\mu\nu}$, where \vec{h}_i^μ is the position of atom i relative to the center of mass of residue μ . Note that this pair potential reduces to a hard-sphere-like interaction in the limit of small atomic overlaps [36]. The total potential energy U is given by

$$U = \sum_{\nu < \mu} \sum_{i,j} U(r_{ij}^{\mu\nu}). \quad (2.15)$$

We use the velocity-Verlet algorithm to integrate the translational equations of motion for each particle's center of mass, and a quaternion-based variant of the velocity-Verlet method described in Ref. [38] to integrate the rotational equations of motions for each residue.

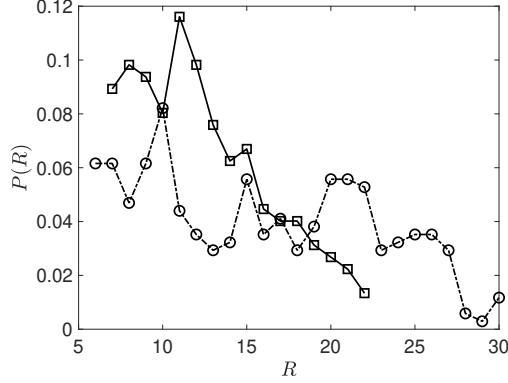


Figure 2.9: Distribution of the number of core residues $P(R)$ in the Dunbrack 1.0 database, before (circles) and after (squares) pruning non-hydrophobic residues from the core replicas as described in Sec. 2.3.1. The mean number of residues before pruning is $\langle R \rangle \approx 16$, and after pruning is $\langle R - r \rangle \approx 12$.

To simulate isotropic compression, we scale all lengths in the system (except the box edges) at each iteration m by the scale factor

$$\alpha = \left(\frac{\phi_m + \Delta\phi_m}{\phi_m} \right)^{1/3}, \quad (2.16)$$

where ϕ_m is the packing fraction and $\Delta\phi_m$ is the packing fraction increment at iteration m . This process uniformly grows or shrinks all atoms, and thus the packing fraction satisfies $\phi_{m+1} = \phi_m + \Delta\phi_m$. After each compression or decompression step, we use the FIRE algorithm [37] to minimize the potential energy in the packing. The packing fraction increment is halved each time the total potential energy switches from zero (i.e. $U/N\epsilon < 10^{-8}$) to nonzero or vice versa. We terminate the packing-generation algorithm when the total potential energy per residue satisfies $10^{-8} < U/N\epsilon < 2 \times 10^{-8}$ and the kinetic energy per residue is below a small threshold, $K/N\epsilon < 10^{-20}$. We set the initial values of the packing fraction and packing fraction increment to be $\phi_0 = 0.4$ and $\Delta\phi_0 = 10^{-3}$, but our results do not depend sensitively on these values.

2.5.2 Protein core size distribution

In this Appendix, we show the distributions of the number of core residues in protein crystal structures from the Dunbrack 1.0 database. (See Fig. 2.9.) As described in Sec. 2.2, we define protein cores as clusters of residues that all share at least one SV cell face with other residues in the cores, and every atom in each residue has an rSASA $\leq 10^{-3}$. In Method M1 for generating jammed packings of amino acid-shaped particles, we create \mathcal{C} replicas of each protein core with the specific $R - r$ residues found in that core, where R is the number of core residues and $R - r$ is the number of Ala, Ile, Leu, Met, Phe, and Val core residues. Before pruning non-hydrophobic residues, the

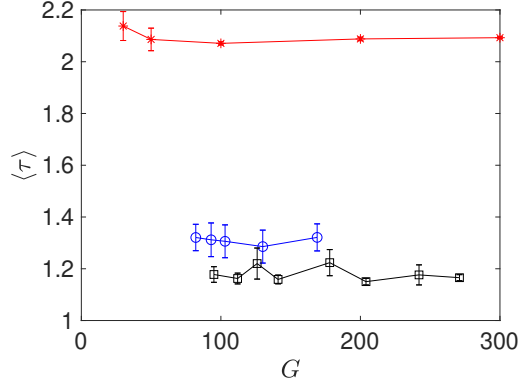


Figure 2.10: Plot of the average Fisher exponent $\langle \tau \rangle$ as a function of the number of lattice points along a given direction G for void percolation on a cubic lattice (stars), in packings of jammed amino acid-shaped particles (open circles), and packings of randomly placed spheres (open squares). Error bars are calculated by the bootstrap method, where τ is fit to 200 separate trials of independent, random subsets of the cluster size data. The average Fisher exponent $\langle \tau \rangle$ is the mean of these fits, and the error bars are the standard deviations.

average core size is $\langle R \rangle \approx 16$ residues, and $\langle R - r \rangle \approx 12$ after pruning.

2.5.3 Measurement of Fisher Exponent

In this Appendix, we explain the differences we observe in the Fisher exponent τ for different systems. In void percolation in cubic lattices, systems of randomly placed spheres and jammed packings of amino acid-shaped particles, the distribution of void cluster sizes at percolation onset $n_s(\eta_c)$ has a well defined power-law decay, as shown in Fig. 2.7 (b). We demonstrate this difference further in Fig. 2.10, where we plot the average Fisher exponent $\langle \tau \rangle$ with error bars as a function of the number of void grid points G measured along the box length. Error bars are estimated by the bootstrap method (described above). We find that the value of τ for randomly placed spheres and jammed packings of amino acid-shaped particles are similar, while markedly different from τ measured for void percolation on cubic lattices.

Non-power-law decay in the void cluster size distribution, as displayed in Fig. 2.7 (b) for jammed sphere packings, may be due to the existence of competing length scales in the system. The typical form of Eq. (2.10) at any porosity η is [56]

$$n_s = s^{-\tau} \exp(-s/s_\xi), \quad (2.17)$$

where s_ξ is the number of sites in a cluster with correlation length ξ . In systems where ξ is the only length scale, $s_\xi \rightarrow \infty$ as $\eta \rightarrow \eta_c$ and Eq. (2.17) reduces to Eq. (2.10). However, if there is another intrinsic length scale in the system that is still relevant at the void percolation transition, it is not necessarily true that $s_\xi \rightarrow \infty$. s_ξ can remain finite, and add an exponential tail to $n_s(\eta_c)$. Indeed,

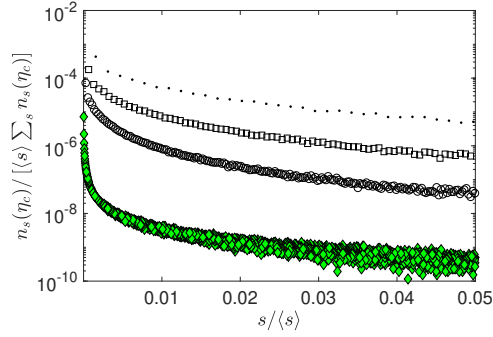


Figure 2.11: Cluster size distribution at percolation $n_s(\eta_c)$ versus $s/\langle s \rangle$, normalized such that each curve has unit area, for void percolation through randomly placed spheres using Kerstein’s method with $N = 300$ (dots), 10^3 (squares), and 10^4 (open circles). We also show the normalized $n_s(\eta_c)$ calculated using the connected void method for $N = 7 \times 10^3$ randomly placed sphere (filled diamonds).

this behavior is what we find for the connected void size distribution in jammed sphere packings. The “kink” in $n_s(\eta_c)$ in Fig. 2.7 (b) indicates that $s_\xi \approx 20$. However, the end of the n_s distribution for void percolation in jammed packings of spheres seems to gain a power-law tail which matches that of void percolation on cubic lattices. Studies of larger systems would be required to confirm this similarity, but the data in Fig. 2.7 (b) suggests that void percolation in jammed packings of spheres resembles that in systems composed solely of a cubic lattice.

This second length scale is most likely set by the nearest-neighbor distances between particles. Qualitatively, if the nearest-neighbor distance between particles is a δ -function (or a set of δ -functions, in the case of polydisperse spheres), there are a limited number of local cavities in the system. In particular, there can be small, particle-scale voids that persist even at the percolation threshold. However, in packings of amino acid-shaped particles and in systems of randomly placed spheres, there are a wide range of inter-particle distances, and a continuous range of local cavity sizes that can form. In Fig. 2.8, we show that the void regions are well-connected for jammed packings of amino acid-shaped particles and randomly placed spheres, while the void regions have a characteristic cavity size for jammed sphere packings at percolation onset. Thus, there is a well-defined Fisher exponent τ in jammed packings of amino acid-shaped particles and randomly placed spheres, but not in jammed monodisperse and bidisperse sphere packings.

Chapter 3

Packing in NMR and X-ray crystal structures

3.1 Introduction

It is generally acknowledged that protein structures determined by x-ray crystallography versus NMR exhibit small but significant differences. It is by no means resolved, however, whether these differences stem from differences in the experimental methods themselves, or if they reflect physical differences in proteins under the different conditions in which the measurements are made [61–68]. To begin to answer this question, one must directly compare high-quality structures of the same protein solved by both methods. Choosing x-ray crystal structures based on their resolution is a straightforward way to identify well-specified structures. In our database of structures solved by both x-ray crystallography and NMR, we only include structures that have been solved by x-ray crystallography at a resolution of 2.1Å or less. We also show that our results do not depend on this resolution threshold as long as it is 3Å or less.

There is, however, no universally accepted metric to assess the quality of NMR structures. We therefore defined one; we determined the number of NMR restraints per residue beyond which structures do not change significantly with the addition of more restraints, and only used structures with at least this number of restraints per residue on average. (See Fig. 3.1.) Applying these selection criteria, we created a data set of 21 proteins whose structures have been determined by both x-ray crystallography and NMR. We created an additional dataset of 51 high-quality NMR protein structures (defined in the same way), for which there is no companion x-ray crystal structure, in an attempt to exclude any influence of ‘crystallizability’ on the NMR protein structures. In addition, as a reference set of high-resolution protein structures solved by x-ray crystallography, we use a dataset of 221 high-resolution protein structures collected by Wang and Dunbrack [30]. Finally, we created a dataset of structures that have been solved multiple times by x-ray crystallography, with resolution of 2.0Å or less and the same crystal forms and space groups, to allow us to assess structural variations that arise from thermal fluctuations.

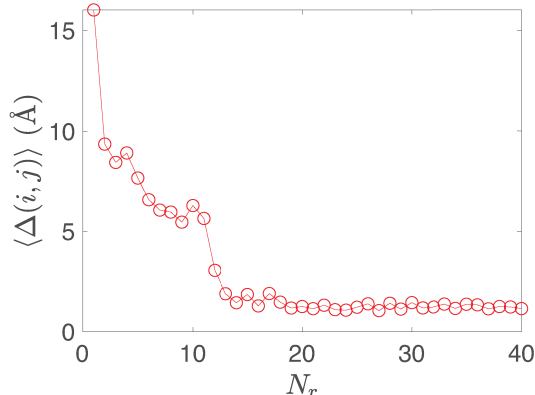


Figure 3.1: Average root-mean-square deviations (RMSD) in the C_α positions $\langle \Delta(i, j) \rangle$ (in Å) of all residues in the larger database of NMR structures without x-ray crystal structure pairs, plotted as a function of the number of restraints on each residue N_r . The average is taken over the multiple structures (~ 20) in each bundle.

We find that the root-mean-square deviations (RMSD) of the positions of core C_α atoms within an NMR ‘bundle’ is greater than the RMSD of core C_α atoms of the set of protein crystal structures that have been solved multiple times, a result found by researchers in prior work [61]. Also, the difference between an x-ray crystal structure and each structure in the NMR ‘bundle’ is greater than the spread within the NMR bundle. To gain deeper insight into these differences, we performed side chain repacking studies on core residues in both x-ray crystal and NMR structures using the hard-sphere plus stereochemical constraint model developed in our previous work [19, 34]. We find that the hard-sphere plus stereochemical constraint model can predict the side chain dihedral angle conformations of core residues equally well in both NMR and x-ray crystal structures, predicting $\Delta\chi$ values to within 30° of the experimental structures. In our previous work, we found that the predictability of side chain conformations is strongly correlated with the local packing fraction ϕ , i.e. where we obtain almost 100% prediction accuracy of side chain conformations for core residues with packing fraction $\phi \geq 0.55$. We therefore also calculate the core packing fractions in NMR and x-ray crystal structures, and find that the cores of NMR structures are more tightly packed than the cores of x-ray crystal structures [61].

To further explore the physical basis for these observations, we generated jammed packings of amino-acid-shaped particles computationally, and determined whether we can tune their packing fraction using protocols with different degrees of thermalization. We find that depending on the thermalization protocol we use, we can match the packing fraction to that which we observe in the cores of structures determined by x-ray crystallography and NMR. Specifically, the packing fraction of amino acid-shaped particles in the athermal limit corresponds to that in the cores of protein crystal structures, whereas the packing fraction we observe in cores of NMR structures is higher, but less than that achieved in the limit of strong thermalization. Thus, the core packing fraction for protein structures determined by x-ray crystallography and NMR are both physically

reasonable, and we speculate that the higher packing fraction for NMR structures reflects the different conditions under which NMR structures are determined.

3.2 Methods

3.2.1 Protein structure datasets

All experimental proteins used in this study were culled from the RCSB Protein Data Bank (PDB). We used datasets of (a) high resolution crystal structures, (b) x-ray crystal-NMR structure pairs, (c) duplicate crystal structures, (d) high-quality, non-paired NMR structures, (e) mutated crystal structures, and (f) structural prediction decoys from the 12th Critical Assessment of Protein Structure Prediction (CASP12). We show the full PDB id’s in the Supplementary Information (SI) for all datasets except the high-resolution crystal structures and the CASP12 decoys and targets. Detailed descriptions of the datasets are provided in the Supplementary Information.

3.2.2 NMR structural quality

There is no universally accepted metric to assess the quality of NMR structures [62]. To define one, we determined the number of NMR restraints per residue beyond which the structures do not change significantly with the addition of more restraints. We measured the root-mean-square deviation (RMSD) of the C_α positions of a given set of residues defined by their sequence location on two models i and j within an NMR bundle:

$$\Delta(i, j) = \sqrt{\frac{1}{N_S} \sum_{\mu=1}^{N_S} (\vec{c}_{\mu,j} - \vec{c}_{\mu,i})^2}, \quad (3.1)$$

where $\vec{c}_{\mu,i}$ is the position of the C_α atom on residue μ in model i , and N_S is the number of residues being compared. We can calculate the average RMSD $\langle \Delta(i, j) \rangle$ by averaging over all pairs of models i and j . As shown in Fig. 3.1, Δ plateaus to a value near 1.5 Å when the average number of restraints per residue reaches $N_r \gtrsim 15$. Thus, we restrict our NMR datasets (Tables S1 and S3 in the SI) to proteins for which the NMR structures possess on average ≥ 15 restraints per residue.

3.2.3 Relative solvent accessible surface area (rSASA)

We define core residues based on their solvent-accessible surface area (SASA). To calculate the SASA, we use the NACCESS software package [33] that implements an algorithm originally proposed by Lee and Richards [69]. The algorithm takes z -slices of the protein, determines the solvent accessibility of the sets of contours using a probe molecule of a given radius, and integrates the SASA over the slices. We use a water-molecule-sized probe with radius 1.4 Å and z -slices with thickness $\Delta z = 10^{-3}$ Å, which were used in previous work [34]. We calculate the SASA for a given residue μ in both the context of the surrounding protein ($\text{SASA}_\mu^{\text{context}}$) and for the residue “extracted” from the protein and modeled as a dipeptide mimetic ($\text{SASA}_\mu^{\text{dipeptide}}$), with all bond

lengths, bond angles, and dihedral angles preserved. We define the relative SASA (rSASA) for residue μ as the ratio

$$\text{rSASA}_\mu = \frac{\text{SASA}_\mu^{\text{context}}}{\text{SASA}_\mu^{\text{dipeptide}}}. \quad (3.2)$$

We define core residues as those with $\text{rSASA} < 10^{-3}$, which was found in previous work [34] to be the largest value of rSASA such that the packing fraction and side chain repacking predictability no longer depend on the value of the rSASA cutoff when it is decreased.

3.2.4 Packing fraction

The most direct way to characterize packing in protein cores is to measure the dimensionless volume fraction, or packing fraction ϕ . The packing fraction ϕ_μ of a single residue μ in a protein core is defined as

$$\phi_\mu = \frac{v_\mu}{V_\mu^v}, \quad (3.3)$$

where v_μ is the volume of residue μ , and V_μ^v is the volume of the Voronoi cell surrounding residue μ . To calculate the Voronoi tessellation for a given protein core, we employ surface Voronoi tessellation [39], which defines a Voronoi cell as the region of space in a given system that is closer to the bounding *surface* of residue μ than to the bounding surface of any other residue in the system. We calculate the surface Voronoi tessellations using the POMELO software package [40]. This software approximates the bounding surfaces of each residue by triangulating points on the residue surfaces. We find that using ~ 400 points per atom, or ~ 6400 surface points per residue, gives an accurate representation of the surface Voronoi cells and the results do not change if more surface points are included. Note that to calculate the average packing fraction of a protein core, we define

$$\langle \phi \rangle = \frac{\sum_\mu v_\mu}{\sum_\mu V_\mu^v}, \quad (3.4)$$

where the sum over μ includes only core residues. In this work, we define a protein core as the set of residues with $\text{rSASA} < 10^{-3}$ that all share at least one surface Voronoi cell face with each other.

3.2.5 Side chain repacking

To better understand the dominant forces determining the side chain conformations in protein cores, we have developed a protocol that can repack the side chains of core residues assuming that the non-bonded atomic interactions are hard-sphere-like, and that bond lengths and angles are tightly constrained around experimentally-observed values. The hard-sphere plus stereochemical constraint model has been used extensively in previous work (e.g. Refs. [19, 34] and references therein) to accurately place hydrophobic residue side chains in the cores of the crystal structures of globular proteins, transmembrane proteins, and protein-protein interfaces. In this model, we sample all possible combinations of the side chain dihedral angles of the core residues, and calculate the purely repulsive Lennard-Jones interaction energy (Eq. (3.6)) between non-bonded atoms for each combination. The backbone dihedral angles of each core residue are fixed to their experimental

values, as well as the side chain and backbone dihedral angles of the rest of the protein. We obtain a probability distribution for the side chain dihedral angle combinations of each core residue using Boltzmann weighting, and repeat this procedure over an ensemble of structures with core residues given different bond-length and bond-angle variants constrained around the experimental values. We then average the probability distributions over this ensemble and identify the side chain dihedral angle combination with the highest probability. We employ this model to study residue packing and side chain placement in the cores of both x-ray and NMR structures. Additional details of the method are given in the SI.

3.2.6 Jammed packings of amino-acid-shaped particles

In previous work [7], we found that the packing fraction and void distribution of protein cores is well-modeled by computer simulations of jammed packings of purely repulsive, rigid, and non-backbone-connected particles shaped like hydrophobic residues. The amino-acid-shaped particles include the backbone N, C $_{\alpha}$, C, and O atoms, as well as all side chain atoms and hydrogens placed using the REDUCE software [32]. Atomic radii are listed in Table S6 in the SI. To prepare the jammed packings, we first place N amino-acid-shaped particles with random positions and orientations in a cubic box with periodic boundary conditions at an initially dilute packing fraction $\phi_0 = 0.1$. The packing fraction is increased by small steps $\Delta\phi$, with each followed by energy minimization, to mimic athermal isotropic compression of the system. We also carry out thermalized compression protocols, where we thermalize the amino-acid shaped particles between compression steps. In this method, we run molecular dynamics trajectories at constant temperature T for a fixed amount of time t_{MD} , and then minimize the total potential energy of the system U using the FIRE minimization method [37] prior to the next compression step. We terminate the packing generation protocols when the minimized potential energy per particle satisfies $10^{-16} < U/N\epsilon \leq 2 \times 10^{-16}$, where ϵ is the energy scale of the non-bonded atomic interactions, and the kinetic energy per particle $K/N\epsilon < 10^{-30}$. Further details of the packing-generation protocols are given in the SI.

3.3 Results

We first compare pairs of structures determined by x-ray crystallography and NMR spectroscopy by quantifying the root-mean-square deviation (RMSD, Eq. (3.1)) of the C $_{\alpha}$ positions of a given set of residues defined by their sequence location on two structures i and j . For the NMR datasets, i and j represent each model within a bundle, and, for the x-ray crystal duplicate dataset, i and j represent each of the duplicates. As mentioned in Sec 3.2.3, we define core residues as residues with small ($< 10^{-3}$) relative solvent-accessible surface area (rSASA), as defined in Eq. (3.2) in Sec. 3.2.3. In Fig. 3.2 (a), we compare the distributions $P(\Delta_{\text{core}})$ of RMSD values of core residues in x-ray crystal structure duplicates and RMSD values of core residues in NMR bundles. We show that the fluctuations among x-ray crystal structure duplicates are consistent with B-factor fluctuations of the C $_{\alpha}$ positions of core residues, B , which are given by $\Delta = \sqrt{3B/8\pi^2}$. We also compare x-ray crystal and NMR structures for the same proteins by calculating the RMSD between C $_{\alpha}$ atoms of core residues.

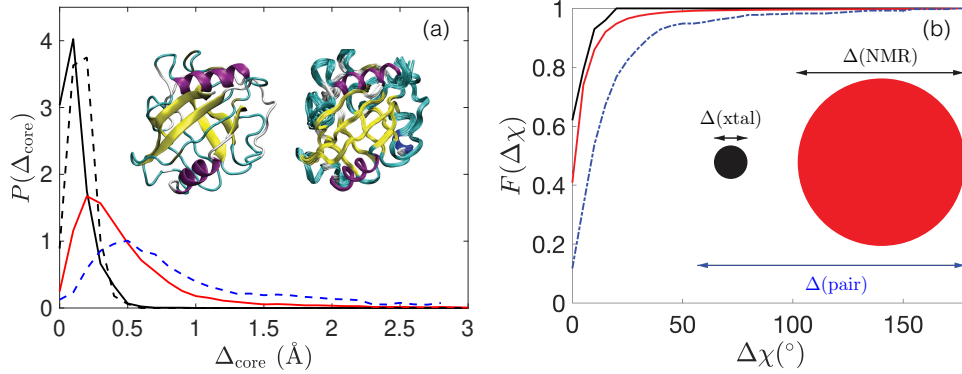


Figure 3.2: (a) Probability distributions $P(\Delta_{\text{core}})$ of the root-mean-square deviations (RMSD) in the positions of the C_α atoms (in Å) for core residues in duplicate x-ray crystal structures (solid black line), in the NMR model ensemble for each structure (solid red line), and in paired x-ray crystal and NMR structures (dot-dashed blue line). We also plot the distribution for $\Delta = \sqrt{3B/8\pi^2}$ from the B-factor for core C_α atoms in the duplicate x-ray crystal structures (dashed black line). The inset shows an example of one of the proteins in the paired x-ray crystal and NMR structure dataset, with the x-ray crystal structure on the left and the bundle of 20 NMR structures on the right (PDB codes 3K0M and 1OCA, respectively). The α -helices are colored purple, the β -sheets are yellow, and the loops are gray. (b) The fraction of core amino acids $F(\Delta_\chi)$ with root-mean-square deviations of the side chain dihedral angles less than Δ_χ (in degrees) for the pairwise comparisons in (a). The inset is a schematic in two dimensions of the high-dimensional volume in configuration space that the C_α atoms in core residues in x-ray crystal structures and NMR ensembles sample. X-ray crystal structures sample a smaller region than NMR ensembles, but the distance between these regions of configuration space is larger than the fluctuations of both the x-ray crystal and NMR structures. The relative lengths of the arrows are drawn to scale, with $\langle\Delta_{\text{core}}\rangle \approx 0.1, 0.5$, and 0.8 Å for the x-ray duplicates, NMR models, and pairs of x-ray crystal and NMR structures, respectively.

To quantify differences between each RMSD distribution, we compute the Jensen-Shannon (JS) divergence [70] for each distribution in Fig. 3.2 (a). The JS divergence between the x-ray duplicate RMSD distribution and the B-factor distribution is 0.5, while the JS divergence between the NMR intrabundle RMSD and the NMR-x-ray RMSD is 1.1, which demonstrates that the RMSD between NMR and x-ray structures is greater than the RMSD differences within a bundle of NMR structures, or between duplicate x-ray structures of the same protein. Because x-ray duplicate RMSD values are similar to B-factor RMSD values, the relatively low JS divergence indicates that fluctuations across duplicate crystal structures is dominated by the uncertainty in atomic positions arising from thermal motion. Whereas the larger JS divergence between NMR intrabundle RMSD and NMR-x-ray RMSD values, as well as the broad tail in the NMR-x-ray RMSD distribution, suggests that differences between structures solved by both NMR and crystallography are larger than those expected in both the ensemble of x-ray structures and in NMR bundles individually. That is, while the fluctuations in the ensemble of observed NMR structures is larger than those in the observed ensemble of crystal

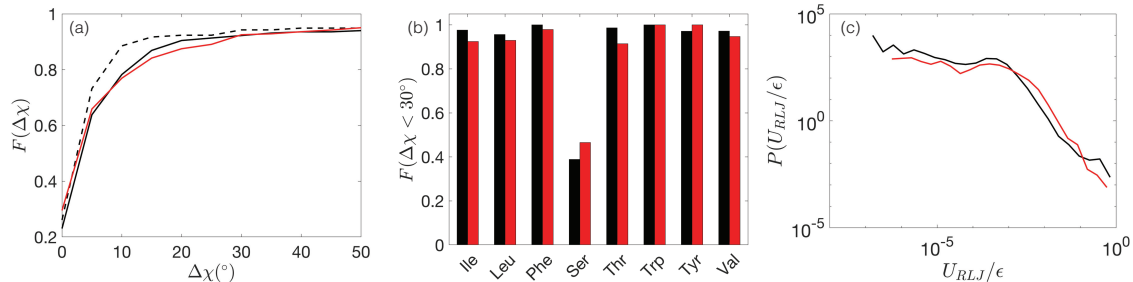


Figure 3.3: (a) Fraction of side chain conformations of core residues with predictions from the hard-sphere plus stereochemical constraint model that deviate from the experimentally observed values by less than $\Delta\chi$ (in degrees) in the dataset of x-ray crystal (solid black line) and NMR (solid red line) structure pairs, and the Dunbrack 1.0 dataset of 221 high resolution x-ray crystal structures (dashed black line) [30, 31]. (b) Fraction of core hydrophobic side chains, grouped by residue type, that can be predicted to within 30° of the corresponding experimental structure using the hard-sphere plus stereochemical constraint model for x-ray (black bars) and NMR structures (red bars). (c) Distribution of the overlap potential energy U_{RLJ}/ϵ , calculated using Eq. (3.6) for core residues in the x-ray crystal (black line) and NMR structures (red line) in the paired dataset.

structures, these two ensembles typically occupy distinct, non-overlapping regions of configuration space.

We also calculate the side chain dihedral angle fluctuations $\Delta\chi$ for the same pairs of structures; we define $\Delta\chi(\mu|i, j)$ as the distance between the side chain conformations of residue μ within structures i and j , i.e.

$$\Delta\chi(\mu|i, j) = \sqrt{(\vec{\chi}_{\mu, j} - \vec{\chi}_{\mu, i})^2}. \quad (3.5)$$

where $\vec{\chi}_{\mu, i}$ is the set of side chain dihedral angles (χ_1, \dots, χ_m) for residue μ on structure i . Note that in Fig. 3.2 (b), we measure $\Delta\chi$ between two experimental structures of the same protein, whereas in Fig. 3.3 (a) and (b) we measure $\Delta\chi$ between an experimental structure and a prediction using the hard-sphere plus stereochemical constraint model.

In Fig. 3.2, we show that the conformations of both the backbone and side chains of core residues fluctuate less in x-ray crystal structures compared to the conformations within an NMR bundle, but that the fluctuations within an NMR bundle are smaller than those *between* the x-ray crystal and NMR structure pairs [61, 67, 68]. The inset to Fig. 3.2 (b) illustrates the proportion of configuration space sampled for structures solved by both NMR and x-ray crystallography. Structures determined by x-ray crystallography sample states in a relatively small volume of configuration space compared to that sampled by structures in an NMR bundle. Moreover, these two ensembles are separated by a characteristic distance that is larger than the scale of fluctuations in either ensemble.

To put these structural differences in context, we also analyze fluctuations in a database of pairs of x-ray crystal structures of wild-type proteins and the same protein with a single core mutation and also high-scoring submissions from a recent Critical Assessment of Protein Structure Prediction (CASP) competition [59]. In the SI (see Fig. S3), we show that the fluctuations of

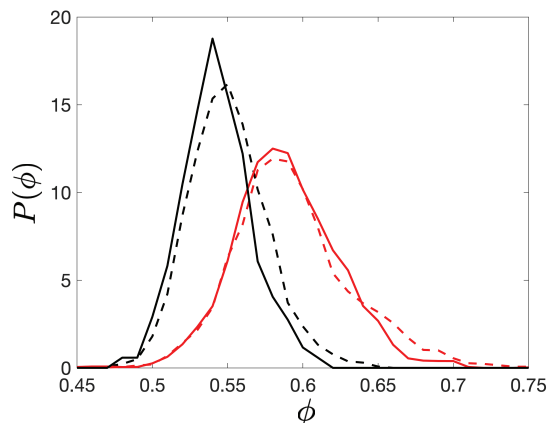


Figure 3.4: Distribution $P(\phi)$ of the packing fraction of core residues in the Dunbrack 1.0 dataset of high-resolution x-ray crystal structures (black dashed line), the dataset of high-resolution NMR structures for which there is not a corresponding x-ray crystal structure (red dashed line), and x-ray crystal structures (black solid line) and NMR structures (red solid) from the paired dataset.

single-site core mutants relative to wildtype structures is similar to that in x-ray crystal structure duplicates. In contrast, submissions to CASP12 exhibit much larger fluctuations. Because CASP12 submissions are computational predictions, not experimentally determined structures, one might expect larger fluctuations. The fluctuations among CASP12 submissions is also larger than those between structures of the same protein determined by x-ray crystallography or NMR. In the SI, we report additional measures of structural fluctuations, such as fluctuations in identities of core residues (Fig. S2). We also show in Figs. S4 and S5 that the global and core RMSD of the C_α positions do not depend on the resolution of the x-ray crystal structures, as long as the resolution is less than 3\AA .

To understand the origin of differences between x-ray crystal and NMR structures, we investigated if these differences are due to physical forces governing sidechain placement of core residues. In previous work, we showed that the hard-sphere plus stereochemical constraint model uniquely specifies the sidechain dihedral angles of core residues in protein crystal structures [34]. One potential source of differences in fluctuations in NMR and crystal structure cores could be that the cores in NMR structures are less well-resolved, and the sidechains are poorly placed due to insufficient information to uniquely define their conformations. Such methodological inaccuracies have been suggested by previous studies, where computational refinement moves NMR backbone and sidechain dihedral angles towards those of x-ray crystal structures [61–64]. However, as shown in Fig. 3.3 (a) and (b), we find that we can repack sidechains of core residues in NMR structures just as accurately as we can repack the same sidechains in high-resolution x-ray crystal structures. The side chain repacking protocol is described in Sec. 3.2.5 and in further detail in the SI. For side chain repacking, we calculate the repulsive Lennard-Jones potential energy of overlap U_{RLJ} between side chains of core residues in the pairs of structures. The potential energy of a single residue μ with

side chain confirmation $\vec{\chi}_\mu$ is defined by

$$U_{\text{RLJ}}(\vec{\chi}_\mu) = \sum_{\nu=1}^N \sum_{i,j} \frac{\epsilon}{72} \left[1 - \left(\frac{\sigma_{ij}^{\mu\nu}}{r_{ij}^{\mu\nu}} \right)^6 \right]^2 \Theta(\sigma_{ij}^{\mu\nu} - r_{ij}^{\mu\nu}), \quad (3.6)$$

where the potential energy is evaluated as a sum over all non-bonded atomic interactions. $r_{ij}^{\mu\nu}$ is the distance between atoms i and j on residues μ and ν , $\sigma_{ij}^{\mu\nu} = (\sigma_i^\mu + \sigma_j^\nu)/2$, and σ_i^μ is the diameter of atom i on residue μ . The Heaviside step function Θ enforces the potential to be purely-repulsive. We find that the distribution of repulsive Lennard-Jones energies between core side chains are almost identical when comparing x-ray crystal and NMR structures, which indicates that the NMR and crystal structure cores are statistically at the same energies. (See Fig. 3.3 (c).)

However, when we investigate the packing fraction ϕ of core residues for x-ray crystal and NMR structures, we find important differences. In Fig. 3.4, we plot the probability distribution $P(\phi)$ of the packing fraction for core residues in x-ray crystal and NMR structures. The average packing fraction of core residues in the protein structures in the datasets determined by x-ray crystallography is $\langle\phi\rangle = 0.55 \pm 0.01$, a value that is consistent with our previous results for the packing fraction of core residues in globular and transmembrane protein cores and the cores of protein-protein interfaces solved by x-ray crystallography [34, 7]. For core residues of protein structures in the NMR database, the average packing fraction is higher with $\langle\phi\rangle = 0.59 \pm 0.02$. We believe that this is the first time that such a difference in the packing fraction of core residues in high-quality protein structures determined by both x-ray crystallography and NMR has been reported.

We were concerned that the higher packing fraction of core residues in protein structures determined by NMR could be an artifact of improperly-placed side chains that overlap with neighboring residues, which would artificially increase the observed packing fraction. However, comparison of the distribution of overlap energies measured by U_{RLJ} (Eq. (3.6)) in Fig. 3.3 (c) demonstrates that the two methods result in almost identical energies, and therefore almost identical atomic overlaps. The difference in the packing fraction of core residues was at first surprising, because our previous studies showed that the cores of x-ray crystal structures pack as densely as jammed packings of purely-repulsive amino-acid-shaped particles without backbone constraints generated using a protocol of successive compressions followed by energy minimization [20, 7].

We therefore revisited the protocol with which we prepared jammed packings of amino-acid-shaped particles [7]. In our previous work, packings were prepared using an ‘‘athermal’’ protocol, where kinetic energy was drained rapidly from the system during the packing preparation. For the athermal protocol, amino acids were initialized in a cubic simulation box at a small initial packing fraction ϕ_0 and compressed by small increments $\Delta\phi$ with each followed by energy minimization (see Sec. 3.2.6 and SI for additional details.) Because the amino-acid-shaped particles were not allowed to translate and rotate significantly between each compression step, the jammed packings at $\langle\phi\rangle \approx 0.55$ were obtained at the first metastable jammed state that the protocol encounters. However, the packing fractions that can be achieved in packings of amino-acid-shaped particles are protocol-dependent; we next investigated more thermalized protocols to see how different protocols lead to different jammed packing fractions.

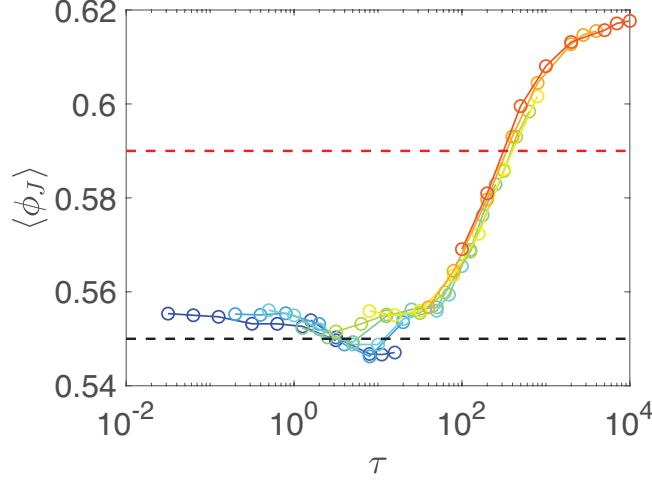


Figure 3.5: Ensemble-averaged packing fraction $\langle \phi_J \rangle$ of jammed packings of amino acid-shaped particles versus the dimensionless timescale τ for $N = 16$ particles. The colors represent simulations with different temperatures $k_B T / \epsilon$, logarithmically spaced from 10^{-7} (blue) to 1 (red). The dashed black line at $\langle \phi_J \rangle = 0.55$ is the average packing fraction of core residues in x-ray crystal structures, and the dashed red line at $\langle \phi_J \rangle = 0.59$ is the average packing fraction of core residues in NMR structures.

We chose a family of annealing packing-generation protocols. We initialize the system in a dilute configuration, and compress the system in small increments $\Delta\phi$ between periods of molecular dynamics simulations of purely repulsive amino acids-shaped particles in the canonical ensemble for a time period t_{MD} at thermal energy $k_B T$. (See SI for details.) We find that temperature only acts to renormalize t_{MD} , i.e. a longer simulation at a lower temperature will yield the same results as a shorter simulation at higher temperatures. Thus, there is another time scale associated with the annealing protocol, $t_{QA} = c(T)t^*$, where $c(T)$ is a dimensionless quantity that depends on temperature, $t^* = \sqrt{m_R \sigma_R^2 / \epsilon}$, and m_R and σ_R are the mass and diameter of the smallest residue. We find that plotting the ensemble-averaged packing fraction $\langle \phi_J \rangle$ of jammed packings of amino acid-shaped particles versus $\tau = t_{MD} / t_{QA} = n \left(\frac{k_B T}{\epsilon} \right)^\alpha$, collapses the data for different temperatures and time periods onto a single curve (Fig. 3.5). The exponent $\alpha = 0.4 \pm 0.01$ and n is the number of time steps between compression increments.

Two limits of packing fractions emerge over the range of annealing protocols we tested; an athermal limit, which corresponds to packing fractions in cores of x-ray crystal structures [34], and the thermalized limit with $\langle \phi \rangle \approx 0.62$. The packing fraction in the cores of protein structures solved by NMR fall between these two extremes with $\langle \phi \rangle = 0.59$. The states at exceedingly high packing fractions exist only in the limit of extremely long annealing times. The results of simulations using different protocols are consistent with the differences observed in cores of protein structures solved by x-ray crystallography and NMR. The fact that thermalized packing protocols yield NMR-like packing fractions, and that athermal protocols generate x-ray crystal-like packing fractions, suggests

that fluctuations are distinct for these two methods.

3.4 Conclusions

In this work, we compare the fluctuations of protein structures characterized by both NMR and x-ray crystallography, and find several key results: first, we found that RMSD values between core residues in duplicated x-ray crystal structures are smaller than RMSD values between core residues across multiple structures in NMR bundles, but these RMSD values are still smaller than the RMSD values between core residues in NMR and x-ray crystal structure pairs. These findings suggest that NMR and x-ray crystal structures occupy distinct regions in configuration space. However, we also showed that the hard-sphere plus stereochemical constraint model is extremely accurate in side chain conformation prediction for core residues in both x-ray crystal and NMR protein structures. Measurements of the core packing fraction show that NMR structures possess denser cores, even though the cores in x-ray crystal and NMR structures possess the same overlap energy. To resolve this apparent discrepancy, we prepare jammed packings of amino-acid-shaped particles both athermally and with thermal agitation, and find that packings produced in the athermal limit resemble the cores of x-ray crystal structures, while thermalized packings resemble cores in NMR structures. This result suggests that there are subtle yet real differences in the fluctuations between structures characterized by x-ray crystallography and NMR spectroscopy. The fluctuations are larger in NMR structures than in x-ray crystal structures, and these fluctuations lead to slightly denser packing in the core.

A previous study that also compared protein structures determined by x-ray crystallography and NMR suggested that the crystal environment restricts dynamical fluctuations, whereas bundles of NMR structures in solution contain the full dynamics one would expect from elastic network models for proteins [66]. The work we present here provides further evidence to support this conclusion, but whether the differences are due to crystalline contacts [66, 67, 71] or the different temperatures at which the protein structures are characterized [72] remains to be determined. Interestingly, several structures used in our dataset of duplicate crystal structures were resolved at room temperature (~ 300 K), as opposed to the cryogenic temperatures typically used in x-ray crystallography. We found that core RMSD values do not change significantly when considering duplicate x-ray crystal pairs solved at different temperatures, which suggests that the crystal environment is the dominant cause of the differences between structures solved by NMR and x-ray crystallography. To fully resolve this question, however, further characterization of protein structure fluctuations at different temperatures is required.

3.5 Appendix

This section includes six Appendices that provide important details about the studies described in the main text. In Appendix 3.5.1, we summarize the six datasets of protein structures that we analyzed. In Appendix 3.5.2, we show that the higher packing fraction of core residues in NMR structures is not an artifact of hydrogen placement. In Appendix 3.5.3, we show the fraction

of common core residues and the root-mean-square deviations (RMSD) of the core C_α positions among structures in several of the datasets. In Appendix 3.5.4, we describe the method to perform sidechain repacking using the hard-sphere plus stereochemical constraint model. In Appendix 3.5.5, we present data on the extent to which the RMSD of the C_α positions of core residues in x-ray crystal structure and NMR structure pairs depends on the resolution of the x-ray crystal structure. In Appendix 3.5.6, we provide additional details concerning the protocols that we use to generate packings of amino-acid-shaped particles in the numerical simulations.

3.5.1 Protein structure datasets

In this Appendix, we describe the six datasets of protein structures used in this work: the Dunbrack 1.0 dataset, the NMR-x-ray dataset, the x-ray duplicate dataset, the NMR dataset, the mutant-WT dataset, and the CASP12 dataset.

Dunbrack 1.0 dataset

The Dunbrack 1.0 dataset [30, 31] is a collection of 221 high-resolution x-ray crystal structures with resolution ≤ 1.0 Å, R-factor ≤ 0.2 , side-chain B-factor per residue ≤ 30 Å², and sequence identity between proteins in the dataset $\leq 50\%$.

Dataset of protein structures solved by both x-ray crystallography and NMR (NMR-x-ray dataset)

We constructed a protein structure dataset (NMR-x-ray) from the Protein Data Bank (PDB) that contains 21 structures that have been solved by both x-ray crystallography and NMR. (See Table 3.1.) The numbers of models and restraints per residue for the NMR data were obtained from the Biological Magnetic Resonance Data Bank (BMRB). The sequence identity between each NMR and x-ray structure pair is $\geq 95\%$, and the sequence identity between different proteins in the dataset is $\leq 50\%$. The sequence identity cutoff for paired structures is chosen to be 95% because NMR structures can have modified residues near the N- or C-terminus to improve the quality of the NMR spectra, and these small changes in sequence generally do not alter the overall structure. The x-ray crystal structures have resolution ≤ 2.1 Å. If an x-ray crystal structure contained two identical chains, only one of the chains was included in all analyses to avoid double-counting. Each NMR structure bundle contains ≥ 20 models and on average ≥ 15 restraints per residue.

Dataset of proteins solved multiple times by x-ray crystallography (x-ray duplicate dataset)

We constructed an x-ray duplicate dataset of 39 proteins from the PDB for which there are at least two independently generated x-ray crystal structures. (See Table 3.2.) All x-ray crystal structures in this dataset satisfy the following conditions: (1) the resolution < 2.0 Å, (2) the structures contain no ligands or modified residues, and (3) the biological assembly and asymmetric unit of the crystal structure contain only a single chain.

NMR PDB ID	# models	# restraints/residue	X-ray PDB ID	Resolution (Å)	R-factor
1jbh	25	21.0	1crb	2.1	0.18
1bu9	21	20.3	1ibh	2.00	0.20
1z2d	20	26.2	1jl3	1.60	0.22
2aas	32	16.1	1kf5	1.10	0.18
1t3v	22	17.9	1o13	1.83	0.28
1b4m	24	24.4	1opa	1.90	0.20
2kl2	20	24.4	1vkb	1.90	0.13
2ka0	20	20.9	1zx8	1.90	0.16
2msn	20	24.4	2go7	2.10	0.17
2y4w	20	16.2	2yb6	1.50	0.16
1oca	20	27.2	3k0m	1.25	0.13
1zgg	20	33.0	4kk3	1.79	0.18
1g7o	21	19.8	4kx4	1.60	0.16
2mpb	20	19.0	4oa3	1.39	0.17
1ael	20	25.2	1ifc	1.19	0.17
1wfr	20	18.2	2cx7	1.75	0.21
2l59	20	17.6	2i1u	1.30	0.21
2nln	20	20.5	1rro	1.30	0.18
2lsu	20	16.2	4cgq	2.00	0.16
2i9h	20	22.8	3f3q	1.76	0.19
1f16	20	18.1	4s0o	1.90	0.20

Table 3.1: Protein Data Bank (PDB) identifiers, number of models in each bundle and number of restraints per residue for the NMR structures, and the resolution and R-factor for x-ray crystal structures in the NMR-x-ray dataset.

PDB ID 1	Resolution (Å)	R-factor	PDB ID 2	Resolution (Å)	R-factor
1ihg	1.80	0.18	1iip	2.00	0.24
1kf3	1.05	0.10	2e3w	1.05	0.19
3d8s	2.00	0.20	3d8r	2.00	0.18
3irz	1.70	0.19	3is0	1.75	0.18
3sn2	1.48	0.19	2yh5	1.90	0.20
4grz	1.37	0.17	4gry	1.69	0.19
5y98	1.36	0.10	4lwq	1.38	0.13
1aa2	2.00	0.21	1bkr	1.10	0.19
1b7d	1.73	0.18	1npi	1.16	0.08
1bpi	1.09	0.19	5pti	1.00	0.19
1ftg	2.00	0.18	1qhe	2.00	0.18
lixh	0.98	0.14	2abh	1.70	0.18
1krn	1.67	0.15	1pk4	1.60	0.14
1kth	0.95	0.14	2knt	1.20	0.21
1l42	1.80	0.16	1l43	1.80	0.16
1lds	1.80	0.19	2d4f	1.70	0.21
1qke	1.50	0.17	5ebx	1.80	0.14
1qy6	1.90	0.20	2o8l	1.50	0.19
1s8h	1.80	0.19	1s8i	1.61	0.19
1xvo	0.84	0.11	2g5l	1.84	0.16
1ymk	1.70	0.18	1ys0	1.15	0.14
lixg	1.05	0.14	1pbp	1.90	0.15
2bwl	1.62	0.18	2bwk	1.50	0.18
2g1h	1.86	0.21	2g2o	1.76	0.22
2w4f	1.30	0.17	5vwc	1.91	0.20
2xxd	1.89	0.26	4aep	1.80	0.19
3dyi	1.72	0.18	3dyv	1.81	0.20
3vhg	1.00	0.16	3vjg	1.00	0.11
4jl7	1.16	0.19	4n6x	1.05	0.15
4wg1	1.70	0.23	4wg7	1.70	0.20
4wob	1.90	0.19	4woc	1.60	0.17
4y9m	1.60	0.18	4y9o	1.81	0.18
5cnw	1.65	0.17	5co3	1.65	0.19
5dsr	1.30	0.12	5dsp	1.4	0.12
5f68	1.23	0.18	5f69	1.37	0.37
5kjo	1.47	0.14	3x2n	1.20	0.11
5nlu	1.20	0.18	5nlw	1.50	0.20
5pg1	1.49	0.21	5pfm	1.54	0.18

Table 3.2: PDB identifiers, resolution, and R-factor for the pairs of protein structures in the x-ray duplicate dataset.

PDB ID	# models	# restraints/residue
2msk	25	21.1
1g7o	21	19.8
2k3f	20	26.4
1xn6	20	24.5
1z9v	20	16.9
2grg	20	25.0
2k3a	20	16.1
2mpb	20	19
2msn	20	24.4
1n3k	20	22.6
2luo	20	29
2hna	21	33.5
2f3i	21	20.1
2jsy	21	37.7
2kc5	20	26.2
1k0s	20	15.8
2mui	20	15.9
2kcw	21	38.0
1fzt	21	16.2
1jbh	20	21.0
1z2d	20	26.2
5z4e	20	20.0
2mzb	20	17.6
1bu9	21	20.3
1t17	20	15.3

PDB ID	# models	# restraints/residue
1q8g	20	20.7
1b4m	24	24.4
2kl2	20	24.2
2mlw	20	24.4
2ka7	20	15.4
2k4k	20	22.0
2kz8	20	30.4
2kmg	20	18.7
2mtb	20	32.3
1jaJ	25	17.9
1t3v	22	17.9
2jr0	20	24.2
2ka0	20	20.9
2i8l	20	45.4
1xo8	20	23.8
2if1	29	17.6
1wki	30	18.4
2ai6	21	29.8
1oca	20	27.1
1qkl	22	15.0
2y4w	20	16.2
2gqb	20	17.9
1ni7	20	19.6
1f16	20	18.1
2mph	20	18.3
1pux	20	16.1

Table 3.3: PDB identifiers, number of models in the bundle, and number of restraints per residue in our dataset of NMR structures without companion x-ray crystal structures.

Dataset of NMR structures for which there is no x-ray crystal structure companion (NMR dataset)

We also generated a (NMR) dataset of NMR structures containing 51 proteins that have ≥ 20 models per bundle and on average ≥ 15 restraints per residue, but do not have a corresponding structure solved by x-ray crystallography. (See Table 3.3.) The sequence identity among proteins in this dataset is $\leq 50\%$.

Dataset of wildtype proteins and mutants with single amino acid mutations solved by x-ray crystallography (mutant-WT dataset)

We generated a mutant-WT dataset of 32 pairs of wildtype and corresponding single amino acid mutant x-ray crystal structures. The mutations were limited to amino acids (in both the wildtype and mutant) with low relative solvent accessible surface area ($\text{rSASA} < 10^{-1}$) and mutations were

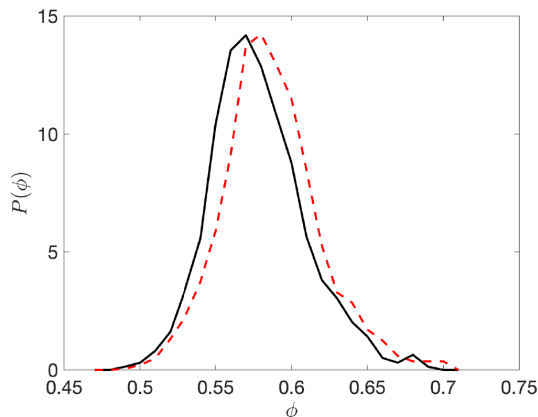


Figure 3.6: Packing fraction of core residues in protein structures solved by NMR with the original hydrogen placement in the PDB file (red dotted) and hydrogen placement obtained using REDUCE (black).

made only between the following residues: Ala, Gly, Ile, Leu, Met, Phe, Ser, Thr, Trp, Tyr, and Val. These choices ensure that we focus on fluctuations in the positions of core residues in the mutant-WT dataset. All of the x-ray crystal structures in this dataset have resolution $\leq 2.1\text{\AA}$.

Dataset of CASP models and their associated targets (CASP12 dataset)

The Critical Assessment of protein Structure Prediction (CASP) is a computational protein design competition organized every two years by the Protein Structure Prediction Center. The Center collects protein structures that have been solved experimentally, but have not yet been publicly released, and uses these structures as targets for computational protein design competitions. We will focus on high-scoring predictions made in the free modeling category in CASP12, in which groups generate predictions using only the primary sequence of the target. Note that CASP collects targets that were solved by different methods. In the CASP12 dataset, we chose targets that were solved via x-ray crystallography to a resolution $\leq 2.0\text{ \AA}$. Out of CASP12's 117 targets, 12 targets satisfied this constraint on resolution. In CASP12, each group that competes was allowed to submit up to five predicted structures for each target. CASP12 ranks structures by combining the z-scores of several structural quantities. In our analysis, we selected the top 25% of the ranked structures for each target. Our CASP12 dataset includes 1059 decoy structures for 12 different protein targets.

3.5.2 Placement of hydrogen atoms in protein structures determined using NMR

In this Appendix, we compare core packing as a function of the placement of hydrogen atoms. We compare the distribution of packing fractions $P(\phi)$ for protein structures solved by NMR, where the hydrogen placement is the same as that in the PDB file and the hydrogen placement is given by

WT PDB	Resolution (Å)	R-factor	MT PDB	Resolution (Å)	R-factor	Mutation
4p9i	1.34	0.14	4p9l	1.43	0.15	A1107M
1tyv	1.8	0.18	1qa3	2.00	0.18	A334I
1tyv	1.8	0.18	1qa2	2.00	0.18	A334V
2rn2	1.48	0.19	3aa5	2.10	0.19	A52F
2rn2	1.48	0.19	3aa2	1.90	0.20	A52I
2rn2	1.48	0.19	3aa4	1.79	0.19	A52V
2rn2	1.48	0.19	1law	1.80	0.19	V74I
2rn2	1.48	0.19	1lav	1.90	0.18	V74L
1r2d	1.94	0.20	1r2i	2.00	0.20	F146L
1kf5	1.15	0.10	lizr	1.5	0.19	F46A
1kf5	1.15	0.10	lizp	1.50	0.20	F46L
1kf5	1.15	0.10	lizq	1.80	0.19	F46V
1goa	1.90	0.19	1gob	2.00	0.19	G77A
2lzm	1.70	0.19	1l23	1.60	0.16	G77A
2lzm	1.70	0.19	256l	1.60	0.16	M6I
2lzm	1.70	0.19	1l52	1.70	0.16	T152S
1hel	1.70	0.15	1heo	1.80	0.16	I55V
1hel	1.70	0.15	1hem	1.80	0.14	S91T
1hel	1.70	0.15	1her	1.80	0.16	T40S
4axt	1.10	0.16	1ir8	1.63	0.16	I58M
4axt	1.10	0.16	1ir9	1.90	0.17	I98M
4axt	1.10	0.16	lios	1.76	0.17	M12F
4axt	1.10	0.16	liot	1.75	0.17	M12L
1agy	1.15	0.18	1cux	1.75	0.16	L114Y
3pzf	1.75	0.17	4roa	1.90	0.17	S358W
3k0m	1.25	0.13	3k0o	1.55	0.10	S99T
1ey0	1.60	0.20	1ey7	1.88	0.15	S128A
1ey0	1.60	0.20	1ey6	1.75	0.20	T41I
1dua	2.00	0.21	1due	2.00	0.21	S195A
2nwd	1.04	0.13	1ckh	2.00	0.18	T70V
2wry	1.58	0.18	3nj5	1.67	0.18	Y157F
4pti	1.50	0.16	8pti	1.80	0.16	Y35G

Table 3.4: PDB identifiers, resolution, and R-factor for the 32 pairs of wildtype and corresponding mutant x-ray crystal structures in the mutant-WT dataset. The sequence location and amino acid identities before and after the mutation are provided in the seventh column.

Dataset	$\langle \Delta_{\text{core}} \rangle$ (Å)
CASP12	4.98
x-ray duplicate	0.12
NMR	0.48
NMR-x-ray	0.83
mutant-WT	0.12
B-factor	0.16

Table 3.5: Average root-mean-square deviations in the C_α positions $\langle \Delta_{\text{core}} \rangle$ in several of the datasets.

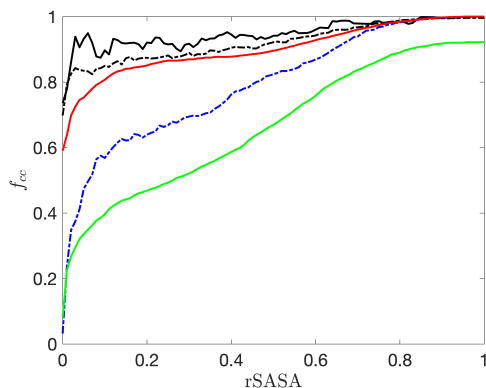


Figure 3.7: Average fraction of identical core residues f_{cc} as a function of the rSASA cutoff that determines whether an amino acid is labelled as core for pairs of structures in the x-ray duplicate dataset (black solid line), for NMR models within each bundle in the NMR-x-ray dataset (red solid line), for wildtype and single amino acid mutant pairs in the mutant-WT dataset (black dotted line), among predictions for each target in the CASP 12 dataset (green solid line), and pairs of NMR models and their corresponding x-ray crystal structures in the NMR-xray dataset (blue dotted line).

the REDUCE software package [32]. As shown in Fig. 3.6, $P(\phi)$ is the same for both sets of NMR structures. In particular, both distributions possess a peak near $\phi \sim 0.58$ - 0.59 . Thus, the high packing fraction for core residues in NMR structures is not an artifact of the hydrogen placement.

3.5.3 Fraction of identical core residues and the distribution of root-mean-square deviations in atomic positions

In this Appendix, we show results for calculations of the fraction f_{cc} of core residues that are shared among a group of protein structures as a function of the threshold rSASA that defines whether a residue is considered as a core residue and the root-mean-square deviations of the C_α atomic positions among groups of protein structures. In Fig. 3.7, we show the results for the fraction of common core residues. We find that for three groups of protein structures (x-ray duplicates, NMR models for a given protein, and mutant-WT pairs) the core regions of the proteins are well-

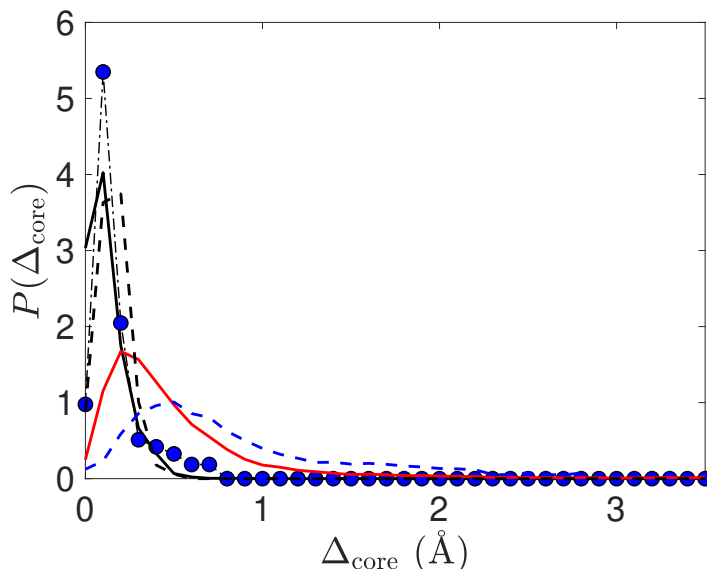


Figure 3.8: Probability distribution $P(\Delta_{\text{core}})$ of the root-mean-square deviations (RMSD) Δ_{core} in the C_{α} positions of corresponding core residues (with $\text{rSASA} \leq 10^{-3}$) between structures in the x-ray duplicate dataset (black solid line), in the NMR dataset (red solid line), in the mutant-WT dataset (blue circles), in the x-ray-NMR dataset (blue dotted line), and the RMSD obtained from the B-factors in the Dunbrack 1.0 dataset (black dotted line).

defined and $f_{cc} \gtrsim 0.8$ of the core residues are shared among structures for a small rSASA threshold $\text{rSASA} \lesssim 0.1$. In contrast, in the dataset of corresponding x-ray crystal structures and NMR models, as well as the high-score CASP12 predictions and their targets, the protein structures do not share a large fraction of core residues for a threshold $\text{rSASA} \lesssim 0.1$.

In Fig. 3.8, we show the results for the root-mean-square deviations (RMSD) Δ_{core} in the C_{α} positions of core residues between structures in several of the datasets. The x-ray duplicates and mutant-wildtype pairs have the smallest RMSD, with peaks in the distribution $P(\Delta_{\text{core}})$ that are below 0.5 Å. The RMSD distribution among NMR models is wider than that for x-ray crystal structures, but narrower than the distribution for corresponding x-ray crystal and NMR structures. A summary of the average RMSD values $\langle \Delta_{\text{core}} \rangle$ is provided in Table 3.5. Note that $\langle \Delta_{\text{core}} \rangle$ for the CASP12 dataset is approximately five times larger than the average for the other datasets.

3.5.4 Single residue sidechain repacking using the hard-sphere plus stereochemical constraint model

In this Appendix, we show the effect of repacking protocols using the hard-sphere plus stereochemical constraint model on the different protein datasets. In previous work [73, 19, 34], we showed that the conformations of the sidechains of core residues in globular proteins, as well as in transmembrane

Atom	Radius (Å)
H	1.1
N	1.3
C _{sp3} /C _{Aromatic}	1.5
C _O	1.3
O	1.4
S	1.7

Table 3.6: The atomic radii in Å used to perform the sidechain repacking studies in this work.

domains and core residues at protein-protein interfaces, can be predicted using a hard-sphere plus stereochemical constraint model. We briefly describe the computational method for single residue sidechain repacking here. To mimic hard-sphere interactions between atoms i and j on residues μ and ν , we use the purely repulsive Lennard-Jones potential,

$$U_{\text{RLJ}}(r_{ij}^{\mu\nu}) = \frac{\epsilon}{72} \left[1 - \left(\frac{\sigma_{ij}^{\mu\nu}}{r_{ij}^{\mu\nu}} \right)^6 \right]^2 \Theta(\sigma_{ij}^{\mu\nu} - r_{ij}^{\mu\nu}), \quad (3.7)$$

where ϵ is the characteristic energy scale of the interaction potential, $r_{ij}^{\mu\nu}$ is the separation between the centers of spherical atoms i and j on residues μ and ν , $\sigma_{ij}^{\mu\nu} = (\sigma_i^\mu + \sigma_j^\nu)/2$, and σ_i^μ is the diameter of atom i on residue μ . The Heaviside step function Θ enforces the constraint that the interaction is only non-zero when two atoms overlap. The probability that the sidechain for residue μ adopts conformation $\vec{\chi}_\mu = \chi_1^\mu, \dots, \chi_m^\mu$ is given by

$$P(\vec{\chi}_\mu) \propto \exp[-\beta U_{\text{RLJ}}(\vec{\chi}_\mu)], \quad (3.8)$$

where β is the inverse temperature (with Boltzmann constant $k_B = 1$) and $U_{\text{RLJ}}(\vec{\chi}_\mu)$ is the total repulsive Lennard-Jones energy for residue μ ,

$$U_{\text{RLJ}}(\vec{\chi}_\mu) = \sum_{\nu=1}^N \sum_{i,j} U_{\text{RLJ}}(r_{ij}^{\mu\nu}), \quad (3.9)$$

summing over non-bonded pairs of atoms i and j for all neighboring residues ν . We chose $\beta\epsilon = 10^2$ to penalize atomic overlaps and we enforce the following normalization condition:

$$\int d^m \vec{\chi}_\mu P(\vec{\chi}_\mu) = 1. \quad (3.10)$$

We repeat this procedure for 300 independent sets of bond-length and bond-angle combinations, which are taken from the Dunbrack 1.0 dataset. The predicted sidechain conformation for residue μ is the one that maximizes $P(\vec{\chi}_\mu)$.

The atomic radii are important model parameters in the sidechain repacking studies. As described in Ref. [20], we chose atomic radii to minimize atomic overlaps in protein cores. The

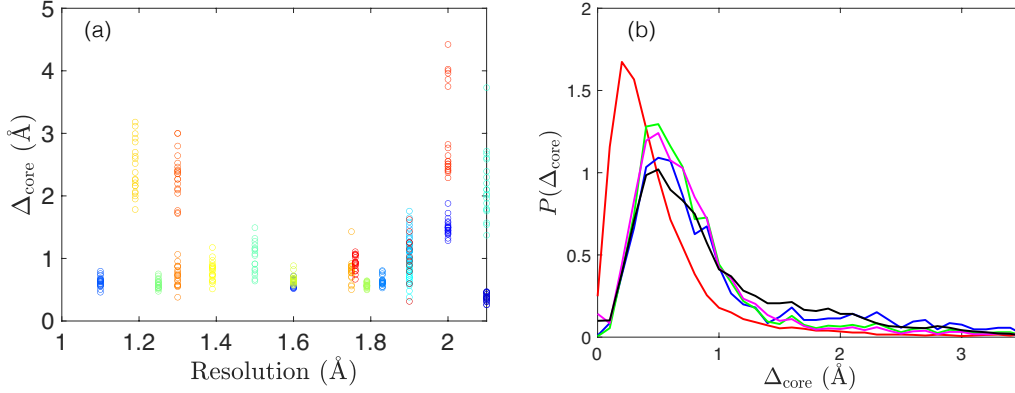


Figure 3.9: (a) Root-mean-square deviations (RMSD) of the C_α positions (Δ_{core}) between a given core residue defined in an x-ray crystal structure at a given resolution and the corresponding residue in an NMR bundle. Each data point corresponds to an average over the ~ 20 structures in each NMR bundle. (b) Probability distribution $P(\Delta_{\text{core}})$ of the RMSD between core residues in an x-ray crystal structure and the corresponding NMR structure for different resolution cutoffs. We consider x-ray crystal structures with resolutions < 1.5 Å (blue), < 1.7 Å (green), < 1.9 Å (magenta), and < 2.1 Å (black). For reference, $P(\Delta_{\text{core}})$ between core residues in a given NMR bundle for the NMR structures is also plotted (red).

positions of Hydrogen atoms are determined using the REDUCE software package [32]. The atomic radii that are used in the sidechain repacking studies are listed in Table 3.6.

3.5.5 Resolution dependence of RMSD between x-ray crystal structure and NMR structure pairs

In this Appendix, we show that the RMSD of the core C_α positions between paired x-ray crystal and NMR structures does not depend strongly on the resolution of the x-ray crystal structure. In Fig. 3.9 (a), we plot the average RMSD (Δ_{core}) of core C_α atoms between paired x-ray and NMR structures, averaged over the > 20 structures in a given NMR bundle, as a function of the x-ray crystal structure resolution. We find no strong correlation between the resolution of the x-ray crystal structure and the RMSD. In Fig. 3.9 (b), we show the same data, but plotted as a distribution of the RMSD $P(\Delta_{\text{core}})$ for each x-ray crystal structure resolution. Again, there is no correlation between the RMSD and the x-ray crystal structure resolution. In addition, it is clear that the average of the RMSD distribution between the x-ray crystal structure and NMR structure pairs is larger than the average of the RMSD distribution between structures within NMR bundles.

We also show that the *global* RMSD of the C_α positions between x-ray crystal structure and NMR structure pairs is not strongly correlated with the resolution of the x-ray crystal structures. In Fig 3.10 (a), we plot the global RMSD of the C_α positions between all residues in the paired high-resolution x-ray crystal structure and high-quality NMR structure dataset (with 21 pairs). Again,

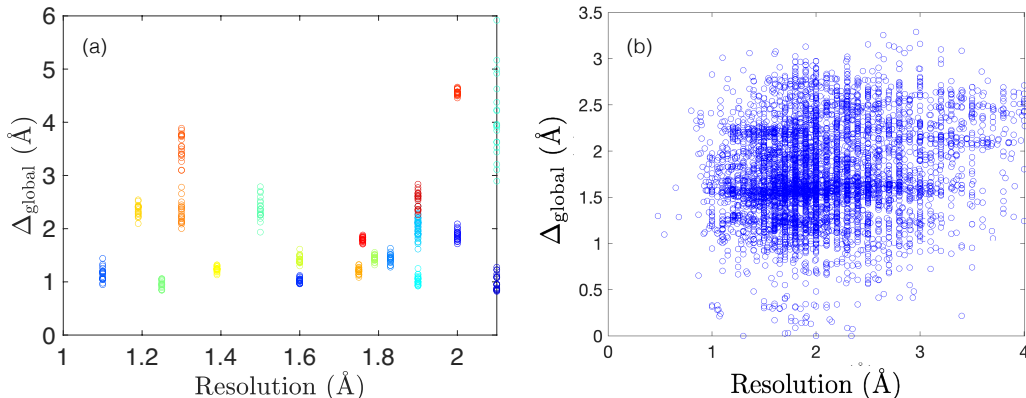


Figure 3.10: (a) Root-mean-square deviations (RMSD, Δ_{global}) of the C_{α} positions between all residues in the x-ray crystal structure at a given resolution and the corresponding residue in an NMR bundle for the same protein for the dataset of paired high-resolution x-ray crystal structure and high-quality NMR structure pairs. (b) Similar plot as that shown in panel (a), except we report the global C_{α} RMSD for all x-ray crystal structure and NMR structure pairs from the CoDNaS database [74].

we do not find a strong correlation between the RMSD values and the x-ray crystal structure resolution. This result is not restricted to the dataset of 21 high-resolution x-ray crystal structure and high-quality NMR structure pairs. In Fig. 3.10 (b), we plot the global RMSD of C_{α} positions versus x-ray crystal structure resolution for all x-ray crystal structure and NMR structure pairs found in the CoDNaS database [74] (with 2384 pairs). We find no correlation between the global C_{α} RMSD and the x-ray crystal structure resolution until ~ 3.0 Å, which indicates that the resolution cutoff used in the main text (i.e. 2.1 Å) does not affect our results.

3.5.6 Protocol dependence of packings of amino-acid-shaped particles

In this Appendix, we describe the simulation techniques used to prepare jammed packings of amino-acid-shaped particles as described in the main text. We model each amino-acid-shaped particle as a rigid union of spheres with fixed bond lengths, bond angles, and dihedral angles. Each particle has sidechain atoms as well as the backbone N, C_{α} , C, and O atoms. The atomic radii are given in Table 3.6. The packings include N residues using a representative sample of core residues from the cores of x-ray crystal structures in the Dunbrack 1.0 dataset. We restrict the amino-acid-shaped particles in our packings to the six hydrophobic amino acids most frequently encountered in protein cores, i.e. Alanine (Ala), Isoleucine (Ile), Leucine (Leu), Methionine (Met), Phenylalanine (Phe), and Valine (Val). To generate the packings, we randomly sample N amino acids of the six types from the protein cores in the Dunbrack 1.0 dataset such that the frequency of the amino acid types in the packings matches that in protein cores in the Dunbrack 1.0 dataset, which are 0.29 for Ala, 0.19 for Leu, 0.17 for Ile, 0.05 for Met, 0.07 for Phe, and 0.23 for Val. The packings for fixed N

thus differ in amino-acid content, and only match the residue identities of Dunbrack 1.0 cores on average. Also note that because the amino-acid-shaped particles are sampled from different cores for each initial condition, the amino-acid-shaped particles possess different bond length, bond angle, and dihedral angle combinations across the packings, even though the individual amino-acid-shaped particles are treated as rigid bodies.

To prepare jammed packings, we start with an initial configuration of amino-acid-shaped particles with random positions and orientations in a three-dimensional cubic domain with periodic boundary conditions at packing fraction $\phi_0 = 0.1$. The particles interact via the pairwise, purely repulsive harmonic spring potential,

$$U_s(r_{ij}^{\mu\nu}) = \frac{\epsilon}{2} \left(1 - \frac{r_{ij}^{\mu\nu}}{\sigma_{ij}^{\mu\nu}} \right)^2 \Theta \left(1 - \frac{r_{ij}^{\mu\nu}}{\sigma_{ij}^{\mu\nu}} \right). \quad (3.11)$$

Here, ϵ is the characteristic energy scale of the interaction potential, $r_{ij}^{\mu\nu}$ is the inter-atomic separation between atoms i and j on distinct residues μ and ν , and $\sigma_{ij}^{\mu\nu}$ is the distance between the atomic centers at contact. The Heaviside $\Theta(\cdot)$ function enforces that the interaction is purely repulsive, i.e. that the interaction potential $U_s(r_{ij}^{\mu\nu}) = 0$ when $r_{ij}^{\mu\nu} \geq \sigma_{ij}^{\mu\nu}$. Note that the purely repulsive Lennard-Jones interaction potential in Eq. (3.7) reduces to Eq. (3.11) in the limit of small particle overlaps. Pairwise forces between atoms i and j are then defined by $\vec{F}_{ij}^{\mu\nu} = -\vec{\nabla} U_s(r_{ij}^{\mu\nu})$, and we integrate Newton's equations of motion to update the positions and velocities of the residue's centers of mass using the velocity-Verlet algorithm [75]. Since the amino-acid shaped particles are treated as rigid bodies, the forces will also generate torques on the particles. The torque on residue μ arising from the pairwise force $\vec{F}_{ij}^{\mu\nu}$ is $\vec{h}_i^\mu \times \vec{F}_{ij}^{\mu\nu}$, where \vec{h}_i^μ is the position of atom i relative to the center of mass of residue μ . We integrate the rotational equations of motion using a quaternion-based variant of the velocity-Verlet algorithm [38].

As described in the main text, we generate jammed packings of amino-acid-shaped particles using several protocols. We first use an athermal compression protocol that has been used to study jammed packings of granular materials [76]. The process begins with an initially dilute configuration with $\phi_0 \ll \phi_J$. To simulate isotropic compression, we grow the particles in small packing fraction increments $\Delta\phi$ and use the FIRE algorithm [37] to minimize the total potential energy of the packing $U = \sum_{\mu < \nu}^N \sum_{i,j} U_s(r_{ij}^{\mu\nu})$ after each compression step. The packing fraction increment is halved each time the total potential energy per particle switches from zero (i.e. $U/N\epsilon < 10^{-16}$) to nonzero or vice versa. We terminate the packing algorithm when the total potential energy per residue satisfies $10^{-16} < U/N\epsilon < 2 \times 10^{-16}$ and the kinetic energy per residue is below a very small threshold of $K/N\epsilon < 10^{-30}$. We set the initial packing fraction to be $\phi_0 = 0.1$ and $\Delta\phi = 10^{-3}$, although our results do not depend sensitively on these values.

We also implemented a quasi-annealing (QA) protocol, where we perform athermal isotropic compression between periods of thermalization. The protocol is similar to what is described above, but we run constant-temperature molecular dynamics (i.e. in an NVT ensemble) for n simulation steps at a dimensionless temperature $\tilde{T} = k_B T / \epsilon$. We use velocity-rescaling to maintain constant temperature, where the velocities \vec{v}_i and angular velocities $\vec{\omega}_i$ are rescaled every time step by a factor $\sqrt{\tilde{T}/T_0}$, where $T_0 = K/\epsilon$ is the dimensionless kinetic energy of the system before rescaling. After n

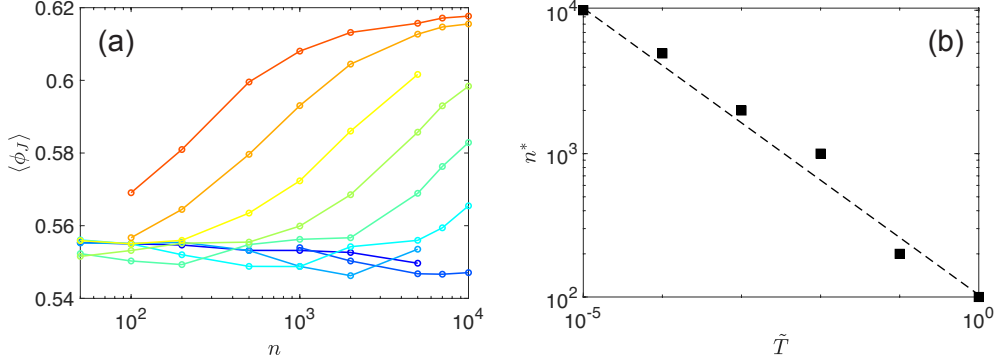


Figure 3.11: (a) Ensemble-averaged packing fraction $\langle\phi_J\rangle$ at jamming onset plotted versus the number of simulation steps n per compression increment using the quasi-annealing protocol. Each color represents a different value of the dimensionless temperature $\tilde{T} = k_B T / \epsilon$ that is logarithmically spaced from 10^{-8} (blue) to 1 (red). (b) The number of steps n^* at which the curves in (a) cross $\langle\phi_J\rangle = 0.57$ is plotted as a function of \tilde{T} . The dashed line has slope -0.4 on a logarithmic scale.

time steps in the NVT ensemble simulation, we rapidly quench the system to $T = 0$ by minimizing the total potential energy. If the total minimized potential energy per particle $U/N\epsilon > 10^{-16}$, the protocol searches for the nearest jammed state at packing fraction ϕ_J using the athermal protocol described above; otherwise, the system is reheated to temperature T and the protocol continues. More than 150 independent configurations were generated for each parameter set considered.

The ensemble-averaged $\langle\phi_J\rangle$ found from the QA protocol is plotted for several n and \tilde{T} values in Fig. 3.11. Each curve has a similar shape, and is shifted from one another along the horizontal axis. We confirm this by plotting n^* , the value of n such that $\phi_J(n^*) = 0.57$, as a function of \tilde{T} . As we expect, n^* decays as a power law $n^* \sim \tilde{T}^{-\alpha}$ with $\alpha = 0.4 \pm 0.01$. Thus, we can define the scaling function, $\langle\phi_J\rangle \propto \mathcal{F}(n\tilde{T}^\alpha)$, which is confirmed by the collapse of the data in Fig. 4 in the main text.

We rationalize this collapse by the fact that the temperature scale \tilde{T} provides a time scale t_{QA} that can be compared to the annealing time, $t_{MD} = n\sqrt{m_R^2\sigma_R^2/\epsilon}$. We can then define the dimensionless time scale $\tau = t_{MD}/t_{QA} = n\tilde{T}^\alpha$. Note that by dimensional analysis one could expect that the thermal time scale is

$$t'_{QA} = \sqrt{\frac{m_R\sigma_R^2}{k_B T}}, \quad (3.12)$$

which would allow us to define a dimensionless time scale $\tau' = t_{MD}/t'_{QA} = n\sqrt{\tilde{T}}$. Thus, the time scale τ that gives the best collapse of the packing fraction data contains an anomalous exponent $\alpha \approx 0.4$, not the square-root dependence in Eq. (3.12).

Chapter 4

Jamming and response of packings of deformable polygons

4.1 Introduction

All soft, athermal solids deform in response to applied stress, yet much of our understanding of these systems relies on computational models using particles with fixed shapes [77, 78]. While extensive work has focused on the effect of varying soft *interparticle* interactions, less attention has been placed on how *intraparticle* degrees of freedom affect collective behavior. Foams [79, 80], emulsions [81, 82], and a wide array of living tissues [83–99] are composed of deformable objects. The complexity and variety of the shape degrees of freedom across these systems emphasizes the importance of investigating how single-particle deformability affects collective properties of soft solids, such as rigidity and linear response.

Athermal systems composed of soft particles form rigid solids at the jamming transition when all non-trivial deformations cost energy [35]. If the particles are spherical, frictionless, and purely repulsive, it is well known that jamming occurs at an isostatic point; mechanically stable configurations at jamming onset in periodic boundary conditions with N_{dof} degrees of freedom and N_c interparticle contacts satisfy $N_{\text{dof}} - N_c = d - 1$ [100]. This observation, a consequence of Maxwell-Calladine constraint counting [101], has been used to rationalize the many anomalous mechanical and vibrational properties of jammed solids [102, 103]. However, particles with non-spherical shapes typically jam with more degrees of freedom than interparticle contacts. These *hypostatic* packings gain mechanical stability from higher-order terms in the Taylor expansion of the potential energy [104–106, 22, 107]. Higher-order stability has been observed in jammed packings of a variety of non-spherical particles [22, 107] and even in packings of “breathing” particles that contain size degrees of freedom [108]. Higher-order constraints directly impact the vibrational spectrum [106, 108], shear response [102], and the glass transition at finite temperature [109].

Recent work [110, 111] has proposed that such higher-order rigidity is a generic feature of hypostatic systems with sufficient pre-stress. This phenomenon has been used to explain the rigidity

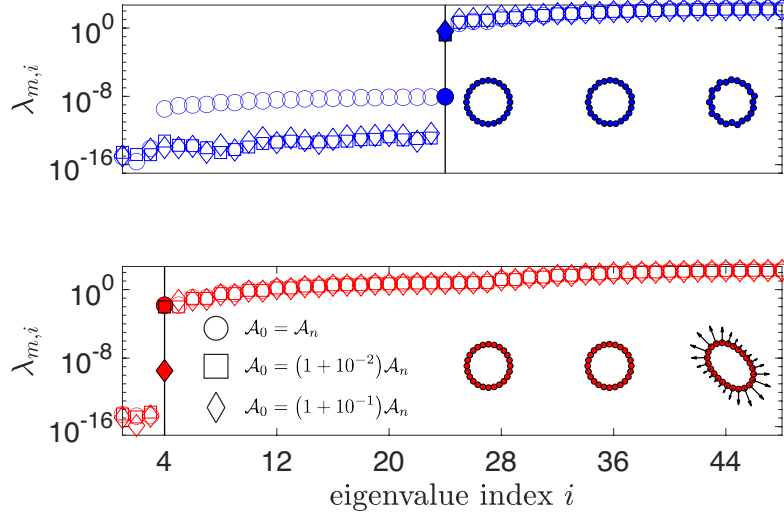


Figure 4.1: Eigenvalues of the single-particle dynamical matrix λ_m for truly deformable particles (DP, top) and deformable particles with bending constraints (DPb, bottom) with $n = 24$ vertices. Symbols correspond to values of the preferred shape parameter \mathcal{A}_0 , and $\mathcal{A}_n = n \tan(\pi/n)/\pi$ is the shape parameter of a regular n -gon. The vertical line at index $i = 24$ ($i = 4$) in the top (bottom) panel correspond to the crossover between zero and non-zero eigenvalues. Energy-minimized shapes are drawn in the insets, with \mathcal{A}_0 increasing from left to right, and the curvature vectors $\vec{\kappa}_i$ defined in Eq. (4.3) are drawn on the buckled DPb particle. In both panels, $K_l = 1$, and $K_b = 10^{-2}$ in the bottom panel.

transition in Vertex models of confluent tissues [112, 113], which can be viewed as dense packings of deformable polygonal cells that are constrained to be confluent. These results suggest that jammed packings of deformable particles might behave similarly, i.e. possess higher-order stability and mechanical and vibrational properties that diverge from those for jammed packings of frictionless, spherical particles. However, are jammed packings of deformable particles identical to those of non-spherical particles such as ellipses? Or does particle deformability lead to fundamentally different mechanical and vibrational response? And how do the properties of jammed packings change as the particles vary from highly deformable to completely rigid?

In this chapter, we study the collective vibrational and mechanical properties of jammed solids composed of particles with varying degrees of deformability. In Sec. 4.2, we introduce a model of deformable polygons. We define deformability through the single-particle vibrational spectra and show that the model can describe truly deformable and rigid-shape particles, as well as *quasi*-deformable particles with characteristics between the two extremes. In Sec. 4.3, we investigate the rigidity, vibrational modes, and shear response in jammed solids composed of deformable particles. Our results emphasize that (a) packings of deformable particles in the rigid-shape-particle limit recover the properties found for jammed packings of soft spherical particles, but that (b) packings

of truly deformable particles do not possess the same vibrational and mechanical properties as those for jammed packings of soft spherical particles. In Sec. 4.4, we conclude with a discussion of the applicability of our results to glassy solids at finite temperature and to several experimental systems. We also include seven appendices, which detail the exact expressions for the forces in the model (Appendix 4.5.1) and for the pressure and shear stress in jammed packings (Appendix 4.5.2), the elements of the dynamical matrix (Appendix 4.5.3), counting effective constraints using the dynamical matrix (Appendix 4.5.4), buckling in single particles with bending energy (Appendix 4.5.5), system-size dependence of the dynamical matrix and shear modulus (Appendix 4.5.6), and identification of the collective shape degrees of freedom (Appendix 4.5.7).

4.2 Methods

Systems of deformable particles in two dimensions are modeled by N distinct polygons, each with n_μ vertices with positions $\vec{r}_{i\mu}$ for $i = 1, \dots, n_\mu$ and $\mu = 1, \dots, N$. Each polygon has an area a_μ and perimeter $p_\mu = \sum_{i=1}^{n_\mu} l_{i\mu}$, where $l_{i\mu}$ is the edge joining vertex i and $i + 1$ on polygon μ . In previous work [47], we studied the deformable polygon (DP) energy,

$$U_{\text{DP}} = \frac{\epsilon_a}{2} \sum_{\mu=1}^N \left(\frac{a_\mu}{a_{0\mu}} - 1 \right)^2 + \frac{\epsilon_l}{2} \sum_{\mu=1}^N \sum_{i=1}^{n_\mu} \left(\frac{l_{i\mu}}{l_{0\mu}} - 1 \right)^2 + U_{\text{int}}, \quad (4.1)$$

where U_{int} is the potential energy between interacting particles, and ϵ_a and ϵ_l are energies controlling area and perimeter fluctuations about the preferred areas $a_{0\mu}$ and edge lengths $l_{0\mu}$, respectively. Interactions between vertices i and j on cells μ and ν are governed by the pair potential v , which we assume depends only on the distance between two vertices, $r_{ij}^{\mu\nu} = |\vec{r}_{i\mu} - \vec{r}_{j\nu}|$. We treat each vertex as a repulsive soft disk, where

$$v(r_{ij}^{\mu\nu}) = \frac{\epsilon_c}{2} \left(1 - \frac{r_{ij}^{\mu\nu}}{\sigma_{\mu\nu}} \right)^2 \Theta \left(1 - \frac{r_{ij}^{\mu\nu}}{\sigma_{\mu\nu}} \right), \quad (4.2)$$

$\sigma_{\mu\nu} = (l_{0\mu} + l_{0\nu})/2$, each vertex has diameter $l_{0\mu}$, ϵ_c controls the strength of the interaction, and Θ is the Heaviside step function to enforce purely repulsive interactions. The total interaction energy is therefore $U_{\text{int}} = \sum_{\nu, \mu} \sum_{i=1}^{n_\mu} \sum_{j=1}^{n_\nu} v(r_{ij}^{\mu\nu})$, though we do not track overlaps between vertices i and $i + 1$ and i and $i - 1$ on the same particle. We measure lengths in units of the square root of the minimum preferred area, $\sqrt{a_0}$, energies in units of ϵ_a , and times in units of $\tau = \sqrt{a_0}/\epsilon_a$, where all vertex masses have been set to 1. The dimensionless *preferred* shape parameter $\mathcal{A}_{0\mu} = (n_\mu l_{0\mu})^2 / (4\pi a_{0\mu})$ measures the amount of excess perimeter above a regular polygon with area $a_{0\mu}$ and thus controls particle deformability [47]. For the DP model, particle shapes depend only on $K_l = \epsilon_l/\epsilon_a$, $K_c = \epsilon_c/\epsilon_a$, and $\mathcal{A}_{0\mu}$.

In Eq. (4.1), we see that the shape of a single DP particle is constrained by $n + 1$ terms given n vertices, but each particle contains $2n$ degrees of freedom. By constraint counting, we expect $2n - (n + 1) = n - 1$ zero energy modes. While each particle contains two translational and one rotational degree of freedom that cannot be constrained by internal forces, DP particles still

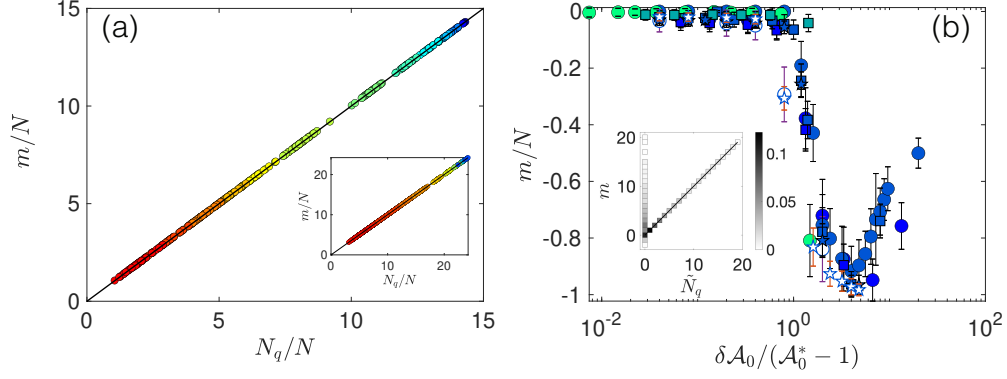


Figure 4.2: (a) Number of missing contacts per particle m/N in packings of $N = 64$ DP particles with $n_S = 16$ (inset, $n_S = 24$) vs. number of quartic modes per particle N_q/N . Black solid line gives $m = N_q$, and colors represent shape parameter values from $\mathcal{A}_0 = 1.0001$ to 1.24, sorted from low (blue) to high (red) values. (b) Missing contacts per particle m/N , where now $m = 3N' - 1 - N_{vv}$ for a system with N' non-rattler particles, in packings of DPb particles plotted vs. $\delta\mathcal{A}_0/(\mathcal{A}_0^* - 1)$. \mathcal{A}_0^* is the particular buckling shape parameter for a given set of parameters, and $\delta\mathcal{A}_0 = \mathcal{A}_0 - 1$. Colors represent K_b (sorted from blue to green), spanning $K_b = 0.005$ to 0.05. The filled symbols are for $n_S = 16$, while white symbols are for $n_S = 24$, and shapes represent different system sizes: $N = 16$ (circles), 32 (squares), and 64 (stars). The inset shows $m = 4N' - 1 - N_{vv}$ for jammed packings of $N = 64$ buckled DPb particles with N' non-rattler particles, and \tilde{N}_q is the number of apparent higher-order modes inferred by the heuristic counting described in the main text. Shape parameters from $\mathcal{A}_0 = 1.04$ to 1.12 are shown, and darker color signifies probability density as denoted by the colorbar. The black line gives $m = \tilde{N}_q$.

contain $n - 4$ non-trivial zero modes. In this sense, DP particles are truly deformable and can change shape with zero energy cost. Example energy-minimized DP particles are shown in the top inset to Fig. 4.1.

To rigidify single DP particles, we add n bending constraints along the particle perimeter [47],

$$U_b = U_{DP} + \frac{k_b}{2} \sum_{i=1}^n \tilde{\kappa}_i^2, \quad \tilde{\kappa}_i = \frac{\vec{l}_i - \vec{l}_{i-1}}{l_0^2}. \quad (4.3)$$

Eq. (4.3) has the additional parameter $K_b = k_b/(\epsilon_a l_0^2)$, which determines the energy cost of bending the particle perimeter. We refer to particles with this additional bending energy term as DPb particles. In addition to single-particle properties, we also study configurations of multiple deformable particles near the onset of jamming. We prepare jammed packings of N bidisperse (50:50 by number) deformable particles in square cells with side length L and periodic boundary conditions. Small (large) particles are given $n_\mu = n_S$ (n_L) vertices with segment lengths $l_{0\mu}$ chosen such that $\mathcal{A}_{0\mu}/\mathcal{A}_n$ is identical for each particle, where \mathcal{A}_n is the shape parameter of a regular n -gon. Therefore, when referring to the shape parameter chosen for a configuration of deformable particles, we will use \mathcal{A}_0

to mean the ratio of $\mathcal{A}_0/\mathcal{A}_n$ for a particle with a given number of vertices. We choose n_L to be the nearest integer to $1.4n_S$ to enforce an approximate 1.4 large-to-small size ratio to avoid crystallization and phase separation [114]. Likewise, large particles are given preferred areas $a_{0\mu} = (1.4)^2 \bar{a}_0$. To create jammed packings, we randomly place particles in the simulation cell at low packing fraction ϕ and isotropically compress the system by increasing the particle size. Compression steps are followed by minimization of the total potential energy U using FIRE. We take configurations as sufficiently near an energy minimum when the total root-mean-square force $< 10^{-12}$. We define the *onset* of jamming, with packing fraction ϕ_J , when the pressure P satisfies $10^{-7} < P < 2 \times 10^{-7}$. Pressure is defined as $P = -dU/dA$, i.e. the total derivative of the potential energy with respect to the area of the simulation domain A , and in Appendix 4.5.2 we derive a formula for the pressure using only the vertex coordinates and domain size. We have confirmed that the results presented below do not depend on the pressure threshold as long as it is sufficiently small. Throughout this work, we will use $K_l = K_c = 1$ unless otherwise stated.

4.3 Results

4.3.1 Rigidity

We first investigate the rigidity of single DP and DPb particles by normal mode analysis. Single-particle normal modes are eigenvectors of the dynamical matrix \mathcal{M} , with block elements defined by $\mathcal{M}_{ij} = \partial^2 U / \partial \vec{r}_i \partial \vec{r}_j$. In Fig. 4.1, we plot the normal mode eigenvalues λ_m for DP and DPb particles with $n = 24$ vertices and varying preferred shape parameters \mathcal{A}_0 . We find DP particles have $n - 1$ near-zero modes ($\lesssim 10^{-15}$) when $\mathcal{A}_0 > 1$, as expected from constraint counting. Interestingly, the DP particle with $\mathcal{A}_0 = 1$ possess $n - 3$ *low* frequency modes significantly above the noise floor. Although the particle shape is underconstrained, particles with $\mathcal{A}_0 = 1$ cannot deform without increasing their perimeter-to-area ratio. These DP particles therefore are stabilized by prestress, a phenomenon found in underconstrained tensegrity structures [101] and disordered spring networks [110].

As DPb particles contain n additional constraints, we expect them to behave as rigid-shape particles (such as frictionless soft disks or ellipses) where any shape deformation costs energy. In Fig. 4.1, we find that there are only 3 near-zero modes, corresponding to the trivial zero modes, for DPb particles with sufficiently small preferred shape parameter and that the particles energy-minimize to regular polygons. However, when $\mathcal{A}_0 = 1.1$, the DP particle is “buckled” with an elliptical shape and an additional low-frequency mode $\lambda_{m,4} \approx 10^{-10}$.

In Appendix 4.5.5, we show that the DPb model contains a buckling transition where energy-minimized shapes elongate from regular polygons and the first non-trivial normal mode eigenvalue $\lambda_{m,4}$ drops from $\sim K_b$ to near zero. The transition point, \mathcal{A}_0^* , varies for different K_b , but the behavior after buckling is similar: $\lambda_{m,4}$ rises from the noise floor with increasing \mathcal{A}_0 , and particles increasingly elongate. The small value of $\lambda_{m,4}$ after buckling suggests that buckled DPb particles gain an extra degree of freedom even though the number of constraints remains constant, a feature

reminiscent of the rigidity transition in vertex models of confluent tissues [112, 110] and topological metamaterials [115, 116].

We then investigate rigidity in jammed packings of DP and DPb particles by calculating the collective vibrational response. In a jammed packing of N' non-rattler DP particles with \bar{n} vertices per particle on average, there are $2N'\bar{n}$ degrees of freedom, $N'(\bar{n} + 1)$ shape constraints and N_{vv} vertex-vertex contacts to constrain the shape degrees of freedom. Isostaticity would dictate $N_{\text{vv}} = N'(\bar{n} - 1) - 1$, but in Fig. 4.2 (a) we show that DP particles at jamming onset are *hypostatic* and seemingly missing the requisite number of interparticle contacts for jamming. The number of missing contacts for jammed DP particles is $m = N'(\bar{n} - 1) - 1 - N_{\text{vv}}$.

Hypostaticity at jamming onset is often observed in packings of non-spherical particles [22]. Recent work has shown that these systems are stabilized by higher-order *quartic* modes of the potential energy [104, 22, 108]. In Appendix 4.5.4, we show that quartic modes can be identified by decomposing the dynamical matrix \mathcal{M} into the stiffness \mathcal{H} and stress \mathcal{S} matrices [105, 106, 113]. As shown in Fig. 4.2 (a), we find that the number of missing contacts in jammed DP packings always matches the number of quartic modes N_q across a wide range of shape parameters.

As DPb particles with n vertices contain n additional constraints, we expect them to behave similarly to packings of rigid-shape bumpy particles that are known to be isostatic at jamming [117]. Indeed, when the particles are regular polygons, i.e. $\mathcal{A}_0 < \mathcal{A}_0^*$, we show in Fig. 4.2 (b) that these packings are isostatic and the number of contacts N_{vv} equals $3N' - 1$, the total number of contacts expected for an isostatic packing of N' non-rattler particles, each with 3 degrees of freedom. However, in Fig. 4.2 (b) we show that near the buckling transition \mathcal{A}_0^* , packings gain contacts and appear to be *hyperstatic* at jamming onset. Hyperstaticity at jamming onset is extremely rare when using athermal protocols [118, 119], so it seems the “buckling mode” with low eigenvalue effectively gives the DPb particles an extra degree of freedom, making these packings actually *hypostatic* with $N_{\text{vv}} < 4N' - 1$.

One might expect to be able to count missing contacts for packings of DPb particles by decomposing the dynamical matrix \mathcal{M} into \mathcal{H} and \mathcal{S} as we did for packings of DP particles. However, we show in Appendix 4.5.4 that all non-trivial eigenvalues of the stiffness matrix \mathcal{H} are nonzero for DPb packings and several orders of magnitude larger than the eigenvalues of the dynamical matrix \mathcal{M} . As discussed in Ref. [110], the presence of positive eigenvalues of the stress matrix \mathcal{S} make counting missing constraints via the dynamical matrix indeterminate. Despite this, we find some evidence of missing contacts in DPb packings using a heuristic approach detailed in Appendix 4.5.4. Briefly, if a packing of DPb particles is missing $m = 4N' - 1 - N_{\text{vv}}$ contacts, we check (i) whether there is a gap between the m th non-trivial eigenvalue λ_m of \mathcal{M} and λ_{m+1} , or (ii) if there is no apparent gap, whether the m th non-trivial mode has a significantly larger participation ratio than mode $m + 1$. We show in the inset of Fig. 4.2 (b) that the heuristic counting largely identifies correctly the \tilde{N}_q higher-order modes that stabilize the missing contacts. However, there are several cases where the counting fails, highlighting the difficulty in determining rigidity in packings of DPb particles with negative pre-stress. Notably, most cases in which the correct number of missing contacts could not be identified in the dynamical matrix eigenvalue spectra or mode structure (i.e., when $m \neq \tilde{N}_q$) occur at $\tilde{N}_q = 0$. That is, whenever there is a notable gap or change in eigenmode participation ratio, we correctly count the number of missing contacts. We reserve a more in-depth analysis of

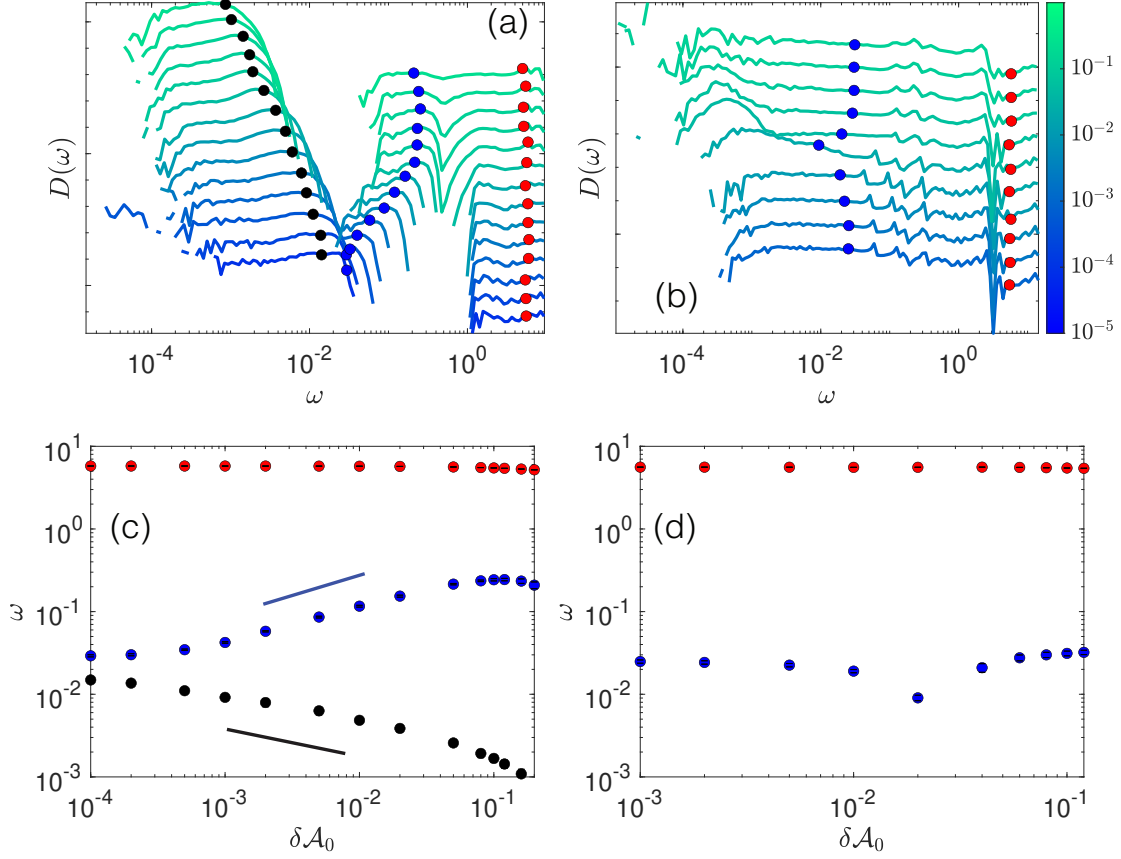


Figure 4.3: Vibrational density of states $D(\omega)$ at jamming onset for the DP and DPb models and a range of shape parameters. (a) $D(\omega)$ for $N = 64$ jammed, bidisperse DP particles ($n_S = 16$) with shape parameter $1.0001 \leq \mathcal{A}_0 \leq 1.20$. (b) $D(\omega)$ for jammed packings of DPb particles ($n_S = 24$) with $K_b = 10^{-2}$ and shape parameters $1.001 \leq \mathcal{A}_0 \leq 1.12$. In both (a) and (b), curves are offset for clarity, the perimeter spring constant $K_l = 1$, and the curve colors shown in the colorbar represent $\delta\mathcal{A}_0 = \mathcal{A}_0 - 1$. (c), (d) Characteristic frequencies ω_0 (black), ω_1 (blue), and ω_2 (red) as a function of $\delta\mathcal{A}_0$ for DP (left) and DPb particles (right). The black and blue lines in (c) represent the scalings $\sim \delta\mathcal{A}_0^{-1/3}$ and $\sim \delta\mathcal{A}_0^{1/2}$, respectively. Dots in (a) and (b) represent the location of each characteristic frequency in $D(\omega)$.

the edge cases where missing contacts were not identified by \mathcal{M} , as well as a predictive theory for the missing contacts as a function of shape parameter, for future work.

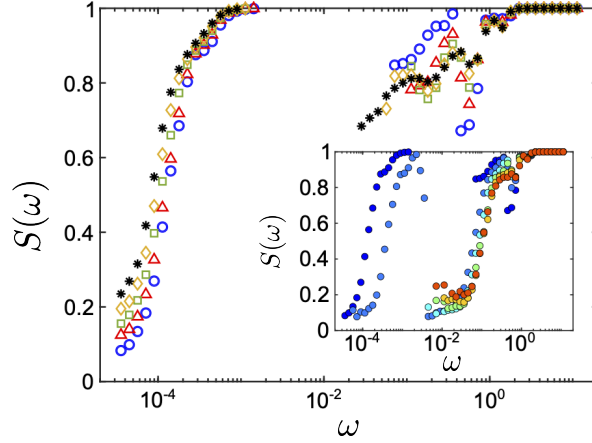


Figure 4.4: Magnitude of mode projection onto the shape degrees of freedom (S) versus the eigenmode frequency ω of the dynamical matrix for $N = 256$ DP packings with $n_S = 16$ and $\mathcal{A}_0 = 1.02$ (circles), 1.06 (triangles), 1.1 (squares), 1.14 (diamonds), and 1.18 (asterisks). Inset: $S(\omega)$ for DP packings with $\mathcal{A}_0 = 1.02$ at several packing fractions from $\phi = \phi_J$ (blue circles) to 0.98 (red circles) in increments of 2×10^{-2} .

4.3.2 Vibrational response

We next study the density of vibrational states $D(\omega)$ for non-trivial vibrational modes with frequency $\omega_i = \sqrt{\lambda_{m,i}}$. In packings of DP particles, we observe three distinct bands of vibrational response in Fig. 4.3 (a) due to quartic modes (with mean frequency ω_0), mid-frequency collective modes (consistently the first $N - 1$ quadratic modes, with mean frequency ω_1), and high-frequency shape modes (with mean frequency ω_2). As shown in Appendix 4.5.6, $D(\omega)$ and the characteristic frequencies do not vary significantly with system size. A similar three-band structure is found in the vibrational response of jammed packings of rigid-shape non-spherical particles [22], although for the DP packings, the second band of modes corresponds to shape fluctuations at particle-particle interfaces rather than particle rotations. Additionally, as shown in Fig 4.3 (c), we find the characteristic scaling $\omega_0 \sim \delta \mathcal{A}_0^{-1/3}$, indicating collective motion becomes less costly as particles become more deformable. We note this behavior differs from jamming of frictionless non-spherical particles with rigid shape [106, 108], where $\omega_0 \sim \mathcal{A}_0^{1/2}$. We find approximate 1/2 scaling with shape parameter in the mid-frequency band ω_1 , although this exponent is ~ 1 in packings of rigid-shape non-spherical particles. The stiff shape mode band with mean frequency ω_2 does not vary with particle shape.

Previous studies have argued that driven and jammed amorphous solids can be described using spherical particles with soft interparticle potentials [77], where particle deformability is modelled by large interparticle overlaps. However, the deviation in $D(\omega)$ for packings of DP particles from that for soft non-spherical particles suggests that explicit shape change plays an important role in determining the vibrational response of soft particles. We further investigate the effect of shape change in the vibrational response by computing the projection of each eigenmode onto collective particle

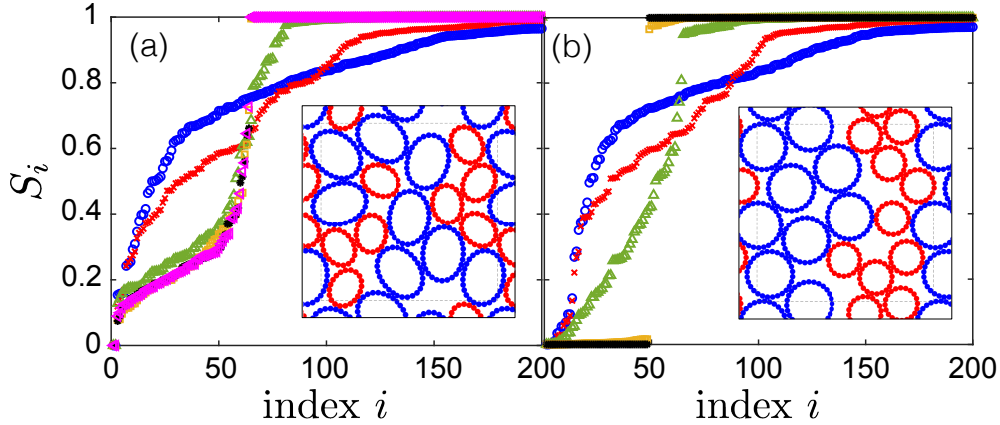


Figure 4.5: (a) Shape degrees of freedom per eigenmode, sorted from smallest to largest, in a packing of $N = 16$ DPb particles (shown in inset) with $n_s = 24$, $K_l = 1$, $K_b = 10^{-2}$, and $\mathcal{A}_0 = 1.04$, which is above the buckling transition $\mathcal{A}_0^* = 1.03$. Curves show changing the interaction parameter K_c from $K_c = 1$ (blue circles) to $K_c = 10^{-5}$ (magenta left triangles) with intermediate values spaced by a factor of 10. (b) Same as (a), but now $\mathcal{A}_0 = 1$ and particles are regular polygons when energy minimized (as shown in inset). K_c is now varied from $K_c = 1$ (blue circles) to $K_c = 10^{-4}$ (black asterisks). The symbols for intermediate values of K_c are the same as in (a).

translations (T), rotations (S), and shape degrees of freedom (S), as described in Appendix 4.5.7. In Fig. 4.4, we show that even the low-frequency modes of jammed DP particles have a significant collective shape projection S across a wide range of shape parameters ($1.02 \leq \mathcal{A}_0 \leq 1.18$). We find that $S(\omega)$ remains > 0 at the lowest frequencies even when the compression is increased well above jamming onset. Explicit shape change is therefore necessary to capture important features of driven soft materials, such as flows of bubbles [80], droplets [120] and emulsions [121, 122].

We also computed $D(\omega)$ for jammed DPb particles as shown in Fig. 4.3 (b). These systems no longer have a distinct band structure in $D(\omega)$, as there are no obvious quartic modes. Here, we define ω_1 as the mean of the first $N - 1$ modes after the trivial zeros in analogy with the DP packings, and ω_2 is the mean of all other modes. For systems with $\mathcal{A}_0 < \mathcal{A}_0^*$, $D(\omega)$ is relatively unchanged as a function of \mathcal{A}_0 . For packings with buckled DPb particles ($\mathcal{A}_0 > \mathcal{A}_0^*$), we observe a higher density of low-frequency modes near the buckling transition and a cusp in ω_1 at \mathcal{A}_0^* as shown in Fig. 4.3 (d). The abundance of low-frequency modes is likely due to the sudden decrease in the magnitude of the single-particle $\lambda_{m,4}$ mode at the buckling transition (see Appendix 4.5.5), and the appearance of modes that can stabilize more than one degree of freedom.

The large density of low frequency modes at the buckling transition for DPb particles raises an important question. Is there a regime where DPb particles will behave as particles with rigid shapes? Or are DPb particles *quasi*-deformable with persistent non-rigid-shape behavior? To address this question, we compute the collective shape degrees of freedom S in individual jammed packings of DPb particles in the rigid-shape limit ($K_c \rightarrow 0$). We show in Fig. 4.5 (b) that non-buckled

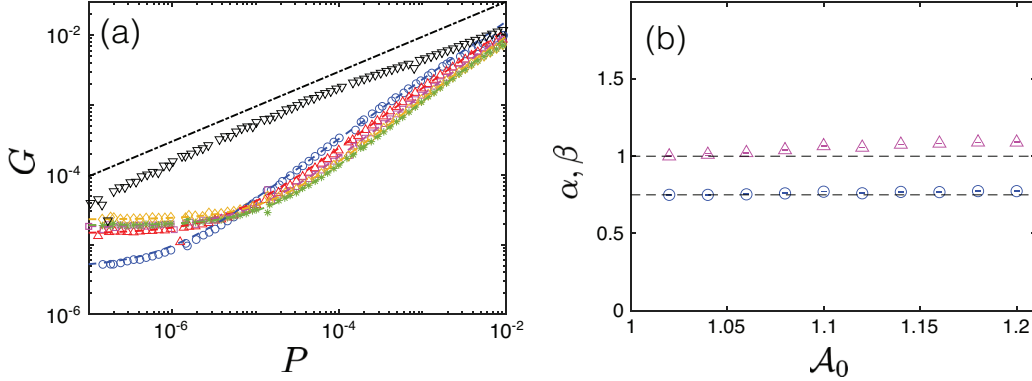


Figure 4.6: (a) Static shear modulus G versus pressure P for $N = 256$ DP packings ($n_S = 16$) with $\mathcal{A}_0 = 1.02$ (circles), 1.06 (triangles), 1.1 (squares), 1.14 (diamonds), and 1.18 (asterisks), and $N = 256$ DPb packings ($n_S = 16$) with $\mathcal{A}_0 = 1$ (downward triangles). The dashed lines are best fits to Eq. (4.4). The dash-dotted line follows the scaling $G \sim P^{1/2}$. (b) Exponents α (triangles) and β (circles) from Eq. (4.4) for the DP packings in (a) versus shape parameter \mathcal{A}_0 . Horizontal lines indicate $\alpha = 1.0$ and $\beta = 0.75$.

DPb particles ($\mathcal{A}_0 < \mathcal{A}_0^*$) eventually reach the rigid-shape limit ($K_c \lesssim 10^{-4}$), where the first $3N$ eigenmodes correspond to purely translational and rotational degrees of freedom and the rest of the spectrum contains only shape degrees of freedom. However, when particles buckle (i.e. $\mathcal{A} > \mathcal{A}_0^*$ in Fig. 4.5 (a)), S is non-zero for the first $3N$ eigenmodes for all values of K_c . We conclude that DPb particles that remain regular polygons are effectively rigid-shape particles, whereas buckled DPb particles are quasi-deformable, and thus the shape degrees of freedom play a key role in their vibrational response.

4.3.3 Shear response

To investigate the effect of particle deformability on bulk mechanical properties, we computed the static shear modulus G for jammed packings of DP and DPb particles. Packings were compressed to a given pressure P , subjected to small, successive simple shear strain steps of size $\Delta\gamma$ with Lees-Edwards boundary conditions [75], and the system was energy-minimized after each step. We measure $G = -d\Sigma/d\gamma$, where Σ is the shear stress and defined in Appendix 4.5.2. We report G averaged over an ensemble of at least 500 configurations. In Fig. 4.6, we show that, although DP packings contain collective low-frequency quartic modes, they possess $G > 0$ at low pressure [104, 47]. In Appendix 4.5.6, we also show characteristic N^{-1} scaling of G in the $P \rightarrow 0$ limit [123]. We find in Fig. 4.6 (a) that $G(P)$ for DP packings over a wide range of \mathcal{A}_0 is well-approximated by the double-power-law functional form [124] used to describe the shear response of

packings of soft frictionless spheres:

$$G = G_0 + \frac{aP^\alpha}{1 + cP^{\alpha-\beta}}, \quad (4.4)$$

where a and c are constants. G_0 is the value in the $P \rightarrow 0$ limit, the exponent α controls the low P response, and the exponent β controls the high P response.

Values of $\alpha \approx 1$ and $\beta \approx 0.5$ have been reported in previous studies of jammed packings of frictionless spherical particles [35, 123], frictional spherical particles [125], and bumpy particles [117]. However, in Fig. 4.6 (b), we find that the large pressure scaling exponent $\beta \approx 0.75$ for DP packings. In Fig. 4.6 (a), we show that $\beta \approx 0.5$ for unbuckled DPb particles with $K_b = 10^{-2}$ and $\mathcal{A}_0 = 1.04$, although we do not observe a plateau at low pressures. This result indicates that the mechanical response for unbuckled DPb particles (with $\mathcal{A}_0 < \mathcal{A}_0^*$) is similar to that for rigid-shape spherical particles.

Note also that $G(P)$ for packings of DPb particles possesses an even smaller scaling exponent at high pressures ($P \sim 10^{-2}$). At these pressures, particles are likely starting to deform from their energy-minimized states to fill in their surrounding Voronoi cell as the packing approaches confluence. However, to fully understand the root cause of these scaling exponents, future work is needed to connect the behavior of single packings to the ensemble average. Prior work on frictionless disks [124] showed that G *decreases* with P for individual packings with fixed contact networks. Only when the contact network changes does the shear modulus increase, leading to a scaling of $P^{1/2}$ when averaging over an ensemble of many configurations with many different contact changes. Understanding the power-law scaling of the ensemble-averaged $G(P)$ for deformable particles requires an analysis of how deformable particles break contacts in response to compression, as well as how the shear modulus varies with pressure when the interparticle contact network does not change [126].

4.3.4 Preliminary studies on non-linear behavior

While measurements of the shear modulus and vibrational density of states have been shown to depend heavily on single-particle mechanical properties, like shape parameter or bending energy, these features only reflect the linear response regime. However, biological systems like dense tissues are inherently out-of-equilibrium and deeply non-linear in almost any context [128, 129]. Therefore, it remains to be seen whether inherently non-linear behavior, like glassy dynamics over long time scales, are affected by single-cell parameters.

To investigate this, we can take a densely-packed configuration of deformable particles and simulate a "crawling" motility strategy that resembles that of single cells on two-dimensional substrates. Each cell μ will be given a preferred crawling director $\hat{n}_\mu = (\cos \psi_\mu, \sin \psi_\mu)$, with dynamics of the crawling angle given by so-called active brownian dynamics [130]

$$\frac{\partial \psi_\mu}{\partial t} = 2D_r \eta_\mu(t) \quad (4.5)$$

where η_μ is Gaussian white noise with zero mean and delta-function spatial and temporal correlations

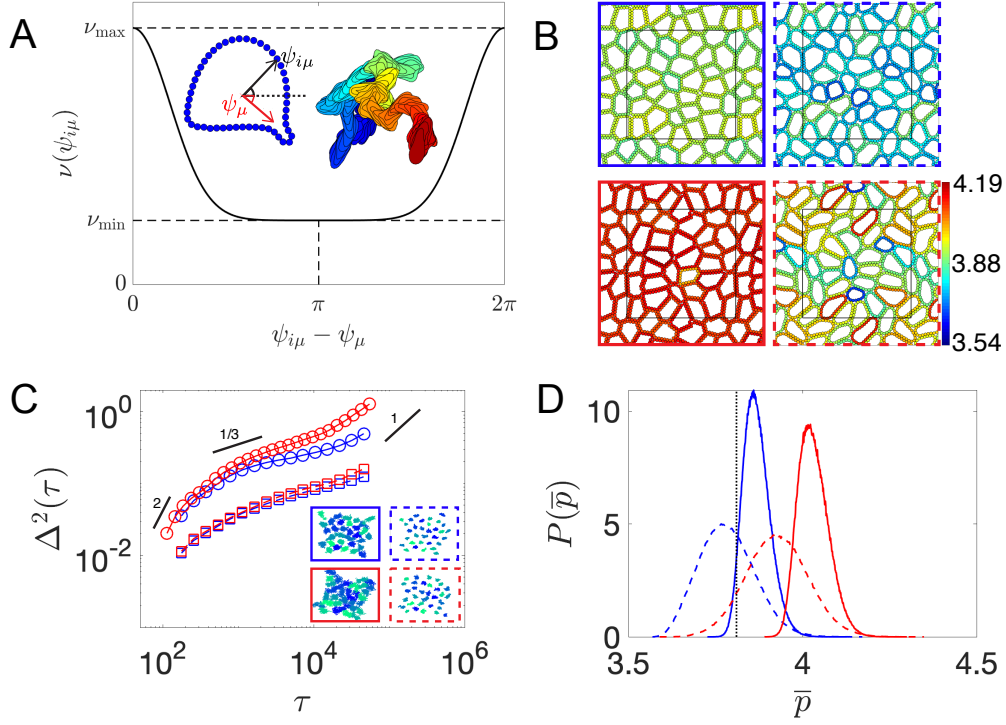


Figure 4.7: (A) Description of the active brownian crawler (ABC) model; each cell μ is given a preferred direction ψ_μ , but each vertex i is given a speed $\nu_{i\mu}$ according to Eq. (4.6), which depends on the angular position $\psi_{i\mu}$ of each vertex. The inset cells with different colors shown a representative time course of a crawling cell, with increasingly warm colors corresponding to later times. (B) Nearly-confluent configurations of either DP (left column) or DPb (right column) polygons, each crawling with ABC dynamics. Rows correspond to preferred shape parameter $\mathcal{A}_0 = 1.1$ (top, blue outline) and $\mathcal{A}_0 = 1.18$ (bottom, red outline). Cell color corresponds to reduced perimeter values $\bar{p} \equiv p/\sqrt{a} = \sqrt{4\pi\mathcal{A}}$; note that the vertex model of confluent tissues [112, 127] predicts fluid-like cells to always have a reduced perimeter of $\bar{p} > 3.81$. Each system has the same crawling speed ν_0 and persistence D_r . (C) Mean-square displacements Δ^2 at lag-time τ for the four systems shown in (B). Circles (squares) correspond to DP (DPb) systems, and blue (red) curves correspond to $\mathcal{A}_0 = 1.1$ ($\mathcal{A}_0 = 1.18$). The inset shows the trajectories of cell centers over time, with the position of each inset matching the position of a given snapshot in panel (B). (D) Stationary probability distribution of shape fluctuations $P(\bar{p})$, which the curve style matching the border of snapshots in (B). The vertical line is drawn at the vertex model prediction for fluid-like cell shapes, $\bar{p} = 3.81$.

$\langle \eta_\mu(t) \eta_\rho(t') \rangle = \delta_{\mu\rho} \delta(t - t')$. Cells will crawl by adding a force to each vertex $\vec{f}_{i\mu} = \nu_{i\mu} \hat{n}_\mu$ along the

crawling director, with the forcing magnitude $\nu_{i\mu}$ given by

$$\nu_{i\mu} = \nu_0 \exp \left[-\frac{(\psi_{i\mu} - \psi_\mu)^2}{2D_\psi^2} \right] \quad (4.6)$$

where ν_0 sets the maximum crawling speed, and $\nu_{i\mu}$ decays with a Gaussian profile (width set by D_ψ) from the vertex with angular position $\psi_{i\mu} = \arctan(y_{i\mu}/x_{i\mu})$ that is closest to the preferred direction of motion ψ_μ . As shown in Fig. 4.7 (A), this drives the cell to crawl through shape change, with the direction motion aligning with the “leading edge” of cell shape deformation. Because cells move similarly to active brownian particles but through a shape-deforming, crawling strategy, we term this model the “Active Brownian Crawling” (ABC) model.

Placing ABC cells together in a dense mixture then allows us to probe the non-equilibrium dynamics of dense cellular configurations. In Fig. 4.7 (B), we show snapshots of small, nearly-confluent systems filled with actively crawling cells. We vary cell preferred shape parameter (either $\mathcal{A}_0 = 1.1$ or 1.18) across the confluence transition of jammed DP particles [47], as well as whether cells have or do not have additional bending constraints to their shape (i.e. DPb or DP, respectively). We track their apparent solid-like or fluid-like behavior by plotting the mean-square displacement Δ^2 of the cell centers (Fig. 4.7 (C)).

We find that, as shown in previous work [131], single-cell mechanical properties play an important role in the non-linear behavior of the system. With no bending energy (i.e. the DP model), preferred shape determines how long it takes for the cells to rearrange their local environment, hop between cages formed by their neighbors, and behave like an equilibrium liquid with $\Delta^2 \sim \tau$. When cells are given bending constraints (i.e. the DPb model), however, $\Delta \sim \tau^{1/3}$ across a broad range of time-scales, indicating the system behaves more like a out-of-equilibrium glass [109]. Here, the preferred shape parameter \mathcal{A}_0 plays less of a role in determining the dynamics. However, actual shape of the ABC cells during dynamics varies significantly, as shown in Fig. 4.7 (D). In particular, note that the red-dashed line in panel (D); this system has a large preferred shape parameter \mathcal{A}_0 and bending energy. We find that the *actual* shape of ABC particles during the simulation are significantly non-spherical even though Δ^2 behaves like that of a glassy solid. Work on the vertex model of confluent tissues [112, 127, 132] has shown that tissue fluidity correlates strongly with cell asphericity. If the mean reduced perimeter of a tissue $\bar{p} > 3.81$ (where $\bar{p} = \sqrt{4\pi\mathcal{A}}$), then the vertex model would predict the tissue to be totally fluid with zero barriers to cell rearrangements. Here, we show a system of strongly-non-spherical deformable particles that can behave solid-like even with apparently fluid-like cell shapes.

Recent work on developing embryos [133] has demonstrated that while the vertex model predictions can capture some behavior of confluent tissues, notable exceptions exist where cells can take on fluid-like shapes yet display little-to-no cellular rearrangements (see e.g. the *bnt* mutant of *Drosophila* in Ref. [133], or the cadherin mutants in Ref. [134]). With ABC deformable particles as model cells, we show that changes to cell rigidity can help explain the complex relationship between cell shape and tissue mechanics. These observations are consistent with what we have seen in the linear response regime (i.e. in vibrational and shear response), reinforcing that studies of the linear response regime still impact non-linear, non-equilibrium properties [135]. However, to understand

these more complex properties in real tissues, much work needs to be done to extend the DP and DPb models to real cells. In particular, cell adhesion and shape change are a fundamentally non-equilibrium processes [88, 94], and the extent to which a purely potential-based model can fully capture the behavior of tissues remains to be seen. However, a purely mechanical model shown in this chapter allows to probe the extent to which these non-equilibrium, fundamentally biological processes affect tissue mechanics, and the extent to which an simpler, effective model can still capture the relevant phenomena.

4.4 Conclusions

In this work, we have studied rigidity, the vibrational density of states, and the mechanical response in athermal, jammed solids composed of particles that can explicitly change shape to varying degrees. We can vary particle deformability by studying the Deformable Polygon (DP) model, where each particle has as many shape degrees of freedom as it has vertices, and the effectively rigid-shape DPb model, which includes bending energy. We also showed that DPb particles can buckle by increasing the preferred shape parameter \mathcal{A}_0 above a characteristic value (\mathcal{A}_0^*), which effectively provides DPb particles with an additional degree of freedom. When studying the rigidity of jammed packings of these particles, we find that DP and DPb particles typically do not jam at a standard isostatic point. Packings of DP particles have too few contacts for collective rigidity, but we find that there are higher-order terms in the potential energy expansion (i.e. quartic modes) that stabilize the packings. Packings of DPb particles below the buckling threshold jam at the expected isostatic point for rigid-shape bumpy particles, but buckled DPb particles jam with *more* contacts than expected and seem to be *hyperstatic*. If we assume that buckled DPb particles have an extra degree of freedom, however, these packings are hypostatic just as in packings of DP particles. Although we cannot reliably count missing contacts from the vibrational spectra for buckled DPb packings, we show that a heuristic counting criterion roughly validates the observation that buckled DPb particles have higher-order rigidity.

Analyzing the vibrational spectra in more detail, we show that the vibrational density of states $D(\omega)$ depends strongly on particle deformability. In particular, we show that the characteristic frequency of quartic modes for jammed DP particles scales *inversely* with particle shape parameter, i.e. $\omega_0 \sim \delta \mathcal{A}_0^{-1/3}$. This result differs from other systems with three vibrational bands, in particular jammed packings of ellipsoids [106] and breathing particles with size degrees of freedom [108]. We also find that collective shape degrees of freedom S play an important role in the low-frequency vibrational response of DP particles across different shapes and with increasing compression. We show that packings of regular-polygon DPb particles in the rigid-shape-particle limit $K_c \rightarrow 0$ do not possess low-frequency collective shape degrees of freedom, whereas $S(\omega) > 0$ for all K_c for packings with buckled DPb particles. We further show that $G \sim P^{3/4}$ over a wide range of pressure for packings of DP particles for all shape parameters studied, which deviates from the power-law scaling for packings of rigid-shape spherical particles. In contrast, packings of regular-polygon DPb particles possess $G \sim P^{1/2}$ scaling.

In all, our results show that explicit particle deformability qualitatively changes the linear response of soft solids. The bulk of these findings can be tested in experiments, either in non-

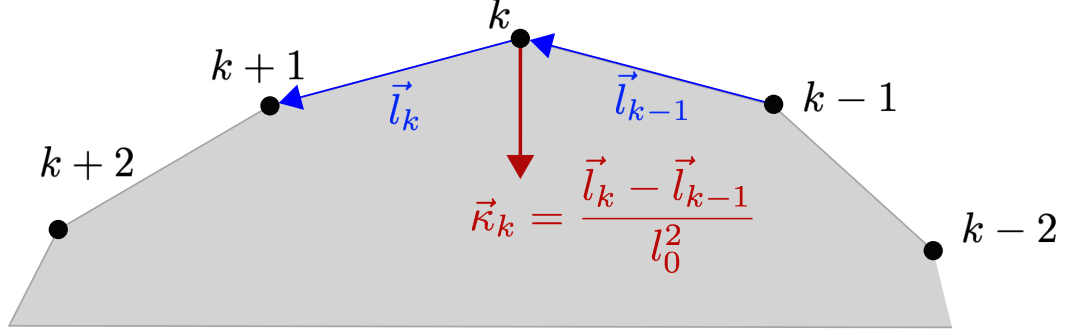


Figure 4.8: Geometry of regions near individual vertices on model cells. Segment vectors $\vec{l}_k = \vec{r}_{k+1} - \vec{r}_k$ drawn in blue define segment lengths $l_k = |\vec{l}_k|$. The curvature at vertex κ_k is defined by the modulus of the vector $\vec{\kappa}_k = l_0^{-2}(\vec{l}_k - \vec{l}_{k-1})$, which always points normal to the cell boundary.

contractile 2D monolayers of epithelial cells (i.e. DP particles) or in soft, quasi-2D packings of hydrogel particles (i.e. regular-polygon DPb particles). The buckling phenomenon observed for DPb particles cannot easily be tested in an experiment, but we plan to carry out further theoretical studies of DPb buckling to gain a deeper understanding of quasi-deformability. Nevertheless, this work lays the foundation for understanding the vibrational and mechanical response in glassy systems of deformable particles, such as hopper flows of emulsion droplets [121, 122] and motile tissues [129].

4.5 Appendices

This section is divided into seven Appendices. In Appendix 4.5.1, we derive the explicit form of the force on each vertex due to the different shape and interaction potential energies. In Appendix 4.5.2, we derive the form of the bulk pressure and shear stress in a given packing of deformable particles as a function of vertex coordinates. In Appendix 4.5.3, we derive the elements of the dynamical matrix as a function of vertex coordinates, and in Appendix 4.5.4 we describe how the dynamical matrix can be used (or how it fails) to count higher-order constraints in jammed packings. In Appendix 4.5.5, we describe buckling of single DPb particles in further detail. In Appendix 4.5.6, we show the effect of finite system size on the vibrational density of states and the static shear modulus. Finally, in Appendix 4.5.7, we demonstrate how normal modes can be decomposed into translational, rotational, and shape contributions.

4.5.1 Force definitions

In this Appendix, we derive the analytical forms of the forces on each vertex k on deformable polygon μ . In general, vertices are acted upon by both shape and interactions forces due to the

area, perimeter, bending and vertex-vertex overlap terms in Eqs. (4.1) and (4.3). We denote the potential of each interaction by U_a , U_l , U_b and U_{int} , respectively. The force on residue i on cell μ is the vector derivative of all potential terms with respect to the coordinate $\vec{r}_{k\mu} = x_{k\mu}\hat{e}_x + y_{k\mu}\hat{e}_y$, i.e.

$$\vec{F}_{k\mu} = -\frac{\partial U}{\partial \vec{r}_{k\mu}} \equiv -\frac{\partial U}{\partial x_{k\mu}}\hat{e}_x - \frac{\partial U}{\partial y_{k\mu}}\hat{e}_y. \quad (4.7)$$

All potential terms

For reference, the four contributions to the total potential energy U for a configuration of N cells, labeled by $\mu = 1, \dots, N$, each with n_μ vertices labeled $i = 1, \dots, n_\mu$, are reproduced below:

$$\begin{aligned} U_a &= \sum_{\mu=1}^N \frac{\epsilon_a}{2} \left(\frac{a_\mu}{a_0} - 1 \right)^2 \\ U_l &= \sum_{\mu=1}^N \sum_{i=1}^{n_\mu} \frac{\epsilon_l}{2} \left(\frac{l_{i\mu}}{l_0} - 1 \right)^2 \\ U_b &= \sum_{\mu=1}^N \sum_{i=1}^{n_\mu} \frac{k_b}{2} \kappa_{i\mu}^2 \\ U_{\text{int}} &= \sum_{\mu=1}^N \sum_{\nu > \mu} \sum_{i=1}^{n_\mu} \sum_{j=1}^{n_\nu} \frac{\epsilon_c}{2} \left(1 - \frac{r_{ij}^{\mu\nu}}{\sigma_{ij}^{\mu\nu}} \right)^2 \Theta \left(1 - \frac{r_{ij}^{\mu\nu}}{\sigma_{ij}^{\mu\nu}} \right) \end{aligned} \quad (4.8)$$

where all geometric quantities are defined in Sec. 4.2.

Area force

The force on vertex k of cell μ due to deviations in the preferred area is given by $\vec{F}_k^a = -\partial U_a / \partial \vec{r}_{k\mu}$. For convenience, we will neglect the index μ when describing purely shape forces, as only vertices on cell μ will be involved. The force is defined as

$$\vec{F}_k^a = -\frac{\partial U_a}{\partial \vec{r}_k} = -\frac{\epsilon_a}{a_0} \left(\frac{a}{a_0} - 1 \right) \frac{\partial a}{\partial \vec{r}_k}. \quad (4.9)$$

The polygonal area a is given by Gauss's "shoelace" formula

$$a = \frac{1}{2} \sum_{i=1}^n x_i y_{i+1} - x_{i+1} y_i, \quad (4.10)$$

where we take the index $n+1$ in the summation to be 1, and the index 0 to be n . Therefore, the derivatives in the x and y directions are not symmetric, rather $\partial a / \partial x_k = (y_{k+1} - y_{k-1})/2$ and $\partial a / \partial y_k = (x_{k-1} - x_{k+1})/2$. Therefore, this gives the force

$$\vec{F}_k^a = \frac{\epsilon_a}{2a_0} \left(\frac{a}{a_0} - 1 \right) [(y_{k-1} - y_{k+1})\hat{e}_x + (x_{k+1} - x_{k-1})\hat{e}_y]. \quad (4.11)$$

Perimeter force

The force on vertex k due to deviations in the segment length l_k (see Fig. 4.8) from its preferred value l_0 is

$$\vec{F}_k^l = -\frac{\partial U_l}{\partial \vec{r}_k} = -\frac{\epsilon_l}{l_0} \left[\left(\frac{l_k}{l_0} - 1 \right) \frac{\partial l_k}{\partial \vec{r}_k} + \left(\frac{l_{k-1}}{l_{0k-1}} - 1 \right) \frac{\partial l_{k-1}}{\partial \vec{r}_k} \right]. \quad (4.12)$$

Only derivatives of l_k with respect to \vec{r}_k and \vec{r}_{k+1} are non-zero. Using the definition of the segment length as $l_k = (l_{kx}^2 + l_{ky}^2)^{1/2}$, the derivatives of l_k and l_{k-1} with respect to \vec{r}_k are

$$\begin{aligned} \frac{\partial l_k}{\partial \vec{r}_k} &= \frac{1}{2l_k} \frac{\partial}{\partial \vec{r}_k} (l_{kx}^2 + l_{ky}^2) = -\hat{l}_k \\ \frac{\partial l_{k-1}}{\partial \vec{r}_k} &= \frac{1}{2l_{k-1}} \frac{\partial}{\partial \vec{r}_k} (l_{k-1x}^2 + l_{k-1y}^2) = \hat{l}_{k-1}, \end{aligned} \quad (4.13)$$

where $\hat{l}_k \equiv \vec{l}_k/l_k$, and the derivatives follow from the definition. Therefore, the force on vertex k due to the perimeter potential is

$$\vec{F}_k^l = \frac{\epsilon_l}{l_0} \left[\left(\frac{l_k}{l_{0k}} - 1 \right) \hat{l}_k - \left(\frac{l_{k-1}}{l_{0k-1}} - 1 \right) \hat{l}_{k-1} \right]. \quad (4.14)$$

Bending force

The force on vertex k due to deviations in the local curvature κ_k from 0 is given by

$$\vec{F}_k^b = -\frac{\partial U_b}{\partial \vec{r}_k} = k_b \sum_{i=1}^n \frac{\partial \kappa_i^2}{\partial \vec{r}_k}. \quad (4.15)$$

In Ref. [47], it was suggested to define the curvature as

$$\kappa_k = \frac{2|\hat{l}_k - \hat{l}_{k-1}|}{l_k + l_{k+1}}. \quad (4.16)$$

which derives from a discretized expression from the curvature of a continuous worm-like chain [136]. In the limit that all segment lengths are their respective rest lengths, this expression reduces to

$$\kappa_k = \frac{|\vec{l}_k - \vec{l}_{k-1}|}{l_0^2} \quad (4.17)$$

which is the magnitude of the vector $\vec{\kappa}_k$ drawn in Figs. 4.1 and 4.8.

To compute the force due to bending, we need $\partial \kappa_i^2 / \partial \vec{r}_k$ for Eq. (4.15). Given the definition of curvature in Eq. (4.17), this derivative is non-zero only when $i = k - 1$, k or $k + 1$. Since

$\kappa_k^2 = l_0^{-4}(\vec{l}_k - \vec{l}_{k-1})^2$, we obtain

$$\begin{aligned}\frac{\partial \kappa_k^2}{\partial \vec{r}_k} &= \frac{4}{l_0^4}(\vec{l}_{k-1} - \vec{l}_k) \\ \frac{\partial \kappa_{k-1}^2}{\partial \vec{r}_k} &= \frac{2}{l_0^4}(\vec{l}_{k-1} - \vec{l}_{k-2}) \\ \frac{\partial \kappa_{k+1}^2}{\partial \vec{r}_k} &= \frac{2}{l_0^4}(\vec{l}_{k+1} - \vec{l}_k)\end{aligned}\tag{4.18}$$

which ultimately leads to the force

$$\vec{F}_k^{\text{b}} = \frac{k_b}{l_0^4} \left[3(\vec{l}_k - \vec{l}_{k-1}) + \vec{l}_{k-2} - \vec{l}_{k+1} \right].\tag{4.19}$$

Interaction forces

The force between two overlapping vertices is defined as

$$\vec{F}_k^{\text{int}} = \sum_{j>i} \frac{\partial v}{\partial r_{ij}} \frac{\partial r_{ij}}{\partial \vec{r}_k}\tag{4.20}$$

where $v(r_{ij})$ is the potential between two overlapping vertices separated by a distance r_{ij} defined in Eq. (4.2). We use a simplified notation here where i and j refer to any vertex on any cell, i.e. we drop the cell-level indices μ and ν , and the sum runs over all pairs of vertices once. Since r_{ij} only depends on the position of vertex k when $i = k$ or $j = k$, we only need the derivative

$$\frac{\partial r_{kj}}{\partial \vec{r}_k} = -\hat{r}_{kj}\tag{4.21}$$

where \hat{r}_{kj} is a unit vector that points from particle k to any other particle j . Therefore, the interaction force is

$$\vec{F}_k^{\text{int}} = - \sum_{j \neq k} \frac{\epsilon_c}{\sigma_{kj}} \left(1 - \frac{r_{kj}}{\sigma_{kj}} \right) \Theta \left(1 - \frac{r_{kj}}{\sigma_{kj}} \right) \hat{r}_{kj}.\tag{4.22}$$

Here, σ_{kj} is the contact distance, or average distance between two vertices at contact. Throughout this chapter, we assumed each vertex has a diameter equal to the preferred length of each segment spring, so throughout $\sigma_{kj} = l_0$. The Θ step function enforces a purely-repulsive interaction potential, as it returns 1 when its argument is positive, and 0 otherwise.

4.5.2 Stress definitions

In this Appendix, we derive the form of the pressure P and shear stress Σ as a function of particle coordinates. Each form of the stress is generally the total derivative of the energy E with respect to a given strain, e.g. volumetric or shear. In this chapter and the subsequent chapter, we will only consider systems in mechanical equilibrium where the kinetic energy is zero, and the only

contribution to E is the potential energy U .

Pressure

For the pressure P , we need to compute the total derivative of the potential energy U with respect to volumetric strains, i.e. to the domain boundary area A . To do this, first note that, for a square boundary with fixed box length L , the pressure can be written $P = -dU/dA = -(2L)^{-1}dU/dL$. The total derivative then becomes

$$P = -\frac{1}{2L} \left(\frac{\partial U}{\partial L} + \sum_{i=1}^{N_v} \sum_{\xi=x,y} \frac{\partial U}{\partial x_{i\xi}} \frac{\partial x_{i\xi}}{\partial L} \right) \quad (4.23)$$

where we sum over all N_v vertices in the system across all cells, and the ξ indicates where $x_{i\xi}$ is either the $\xi = x$ or $\xi = y$ coordinate of vertex i . To measure the pressure, we perform a small deformation to the boundary where $L \rightarrow L + \Delta L$, and an affine transformation to the vertex coordinates $x_{i\xi} \rightarrow x_{i\xi}(L + \Delta L/L)$. In the limit $\Delta L \rightarrow 0$, we have $\partial x_{i\xi}/\partial L = x_{i\xi}/L$, and the pressure becomes

$$P = \frac{1}{2L^2} \sum_{i=1}^{N_v} \vec{F}_i \cdot \vec{r}_i - \frac{1}{2L} \frac{\partial U}{\partial L}. \quad (4.24)$$

The first term is referred to as the system's virial, which captures pressure in thermal systems [75]. The second term captures the purely potential contribution to the pressure, and needs to be derived more explicitly to be computed. We first note that

$$\frac{\partial U}{\partial L} = \sum_{\mu=1}^N \left[\frac{\partial U_a}{\partial a_\mu} \frac{\partial a_\mu}{\partial L} + \sum_{i=1}^{n_\mu} \left(\frac{\partial U_l}{\partial l_{i\mu}} \frac{\partial l_{i\mu}}{\partial L} + \frac{\partial U_b}{\partial \kappa_{i\mu}} \frac{\partial \kappa_{i\mu}}{\partial L} + \sum_{\nu \neq \mu} \sum_{j=1}^{n_\nu} \frac{\partial U_{\text{int}}}{\partial r_{ij\mu\nu}} \frac{\partial r_{ij\mu\nu}}{\partial L} \right) \right]. \quad (4.25)$$

The derivatives of potentials with respect to geometric parameters have already been found in order to compute forces. We then simply need to find how each geometric parameter changes with respect to deformation of the box size L . For lengths, i.e. $l_{i\mu}$ and $r_{ij\mu\nu}$, these objects behave identically to coordinates, therefore

$$\frac{\partial l_{i\mu}}{\partial L} = \frac{l_{i\mu}}{L}, \quad \frac{\partial r_{ij\mu\nu}}{\partial L} = \frac{r_{ij\mu\nu}}{L}. \quad (4.26)$$

For the curvature $\kappa_{i\mu}$, the definition $\kappa_{i\mu}^2 = (\vec{l}_{i+1\mu} - \vec{l}_{i\mu})^2$ makes it more convenient to directly compute the derivative $\partial U_b/\partial L$. We obtain

$$\frac{\partial U_b}{\partial L} = \frac{k_b}{l_0^4} \sum_{\mu=1}^N \sum_{i=1}^{n_\mu} \sum_{\xi=x,y} (l_{i+1\mu,\xi} - l_{i\mu,\xi}) \left(\frac{\partial l_{i+1\mu,\xi}}{\partial L} - \frac{\partial l_{i\mu,\xi}}{\partial L} \right) = \frac{k_b}{l_0^4 L} \sum_{\mu=1}^N \sum_{i=1}^{n_\mu} \kappa_{i\mu}^2 \quad (4.27)$$

where the final expression results from $\partial l_{i\mu,\xi}/\partial L = l_{i\mu,\xi}/L$. We therefore obtain

$$\frac{\partial U_b}{\partial L} = \frac{2U_b}{L} \quad (4.28)$$

for the bending energy contribution to the pressure.

To obtain the change in cell area with respect to box area, we note that the asphericity \mathcal{A} of any given particle will also remain unchanged if the box area is uniformly grown, so therefore $\partial\mathcal{A}/\partial L = 0$. However, using the chain rule, we can show that a cell with area a and perimeter p must obey $\partial a/\partial L = (2a/p) \partial p/\partial L$. Because $\partial p/\partial L = p/L$ given that p is composed of affinely deforming segment lengths, we must have that

$$\frac{\partial a_\mu}{\partial L} = \frac{2a_\mu}{L}. \quad (4.29)$$

Using Eqs. (4.26), (4.28) and (4.29) into Eq. (4.25), we have

$$\frac{\partial U}{\partial L} = \frac{2U_b}{L} + \sum_{\mu=1}^N \left\{ \frac{2a_\mu}{a_{0\mu}L} \left(\frac{a_\mu}{a_{0\mu}} - 1 \right) + \sum_{i=1}^{n_\mu} \left[\frac{\epsilon_l l_{i\mu}}{l_{0i\mu}L} \left(\frac{l_{i\mu}}{l_{0i\mu}} - 1 \right) - \sum_{\nu \neq \mu} \sum_{j=1}^{n_\nu} \frac{\epsilon_{ij\mu\nu} r_{ij\mu\nu}}{\sigma_{ij\mu\nu}L} \left(1 - \frac{r_{ij\mu\nu}}{\sigma_{ij\mu\nu}} \right) \right] \right\}. \quad (4.30)$$

Combining this expression with the virial term in Eq. (4.24) gives a way to compute the instantaneous pressure in a packing given only the vertex coordinates, preferred geometric parameters and the box size.

Shear stress

The shear stress Σ of a packing in mechanical equilibrium is given by the total derivative of the potential energy with respect to a shear strain γ . We can impose the simple shear strain through the deformation

$$x_i \rightarrow x_i + \gamma y_i \quad (4.31)$$

applied to all vertices $i = 1, \dots, N_v$. We then need to compute the total derivative

$$\Sigma = -\frac{dU}{d\gamma} = -\frac{\partial U}{\partial \gamma} - \sum_{i=1}^{N_v} \frac{\partial U}{\partial x_i} y_i. \quad (4.32)$$

In mechanical equilibrium, the second term is zero and we are left solely with the partial derivative of U for our computational of Σ , similarly to the case with computing the pressure above.

Like in the case of pressure, we can rewrite the shear stress equation as

$$\frac{\partial U}{\partial \gamma} = \sum_{\mu=1}^N \left[\frac{\partial U_a}{\partial a_\mu} \frac{\partial a_\mu}{\partial \gamma} + \sum_{i=1}^{n_\mu} \left(\frac{\partial U_l}{\partial l_{i\mu}} \frac{\partial l_{i\mu}}{\partial \gamma} + \frac{\partial U_b}{\partial \kappa_{i\mu}} \frac{\partial \kappa_{i\mu}}{\partial \gamma} + \sum_{\nu \neq \mu} \sum_{j=1}^{n_\nu} \frac{\partial U_{\text{int}}}{\partial r_{ij\mu\nu}} \frac{\partial r_{ij\mu\nu}}{\partial L} \right) \right]. \quad (4.33)$$

First, note that the simple shear strain imposed in Eq. (4.31) is *area-conserving*, i.e. the area of any rectangular boundary is conserved if it is deformed into a simply-sheared parallelogram. Therefore, the individual cell areas a_μ also must be preserved under simple shear, i.e. $\partial a_\mu/\partial \gamma = 0$.

To determine the contributions from other terms, we need to do how general lengths will deform under shear strain, which will impact the interaction, perimeter and bending energy contributions.

As a test case, we will compute $\partial l_{i\mu}/\partial\gamma$ in detail. We have

$$\frac{\partial l_{i\mu}}{\partial\gamma} = \frac{l_{i\mu,x}}{l_{i\mu}} \frac{\partial l_{i\mu,x}}{\partial\gamma}. \quad (4.34)$$

Note that the y -component of the segment vector $l_{i\mu,y}$ does not change under simple shear deformation, and therefore $\partial l_{i\mu,y}/\partial\gamma = 0$. Imposing the simple shear deformation on all coordinates, we must have

$$l_{i\mu,x}(\gamma) = l_{i\mu,x}(0) + \gamma l_{i\mu,y} \quad (4.35)$$

as the affine deformation of the x -component of the segment length vectors under simple shear. Note that this logic can be applied to any length between two coordinates. We therefore have

$$\frac{\partial l_{i\mu,x}}{\partial\gamma} = l_{i\mu,y} \quad (4.36)$$

and

$$\frac{\partial l_{i\mu}}{\partial\gamma} = \frac{l_{i\mu,x} l_{i\mu,y}}{l_{i\mu}}. \quad (4.37)$$

For the inter-vertex distance $r_{ij\mu\nu}$ we similarly can write

$$\frac{\partial r_{ij\mu\nu}}{\partial\gamma} = \frac{x_{ij\mu\nu} y_{ij\mu\nu}}{r_{ij\mu\nu}}. \quad (4.38)$$

For the contribution of the bending energy to the shear stress, it is again more convenient to take the derivative of U_b directly. Using Eq. (4.36) and the chain rule, we obtain

$$\frac{\partial U_b}{\partial\gamma} = \frac{k_b}{l_0^4} \sum_{\mu=1}^N \sum_{i=1}^{n_\mu} (l_{i+1\mu,x} - l_{i\mu,x}) (\hat{l}_{i+1\mu,y} - \hat{l}_{i\mu,y}) \quad (4.39)$$

where $\hat{l}_{i\mu,y}$ is the y -component of the unit vector pointing along the direction of $\vec{l}_{i\mu}$. Combining everything, we have

$$\frac{\partial U}{\partial\gamma} = \sum_{\mu=1}^N \sum_{i=1}^{n_\mu} \frac{\epsilon_l l_{i\mu,x} l_{i\mu,y}}{l_{i\mu} l_{0\mu}} \left(\frac{l_{i\mu}}{l_{0\mu}} - 1 \right) + \frac{k_b (l_{i+1\mu,x} - l_{i\mu,x})}{l_{0\mu}^4} (\hat{l}_{i+1\mu,y} - \hat{l}_{i\mu,y}) - \sum_{\nu \neq \mu} \sum_{j=1}^{n_\nu} \frac{\epsilon_c x_{ij}^{\mu\nu} y_{ij}^{\mu\nu}}{\sigma_{ij}^{\mu\nu} r_{ij}^{\mu\nu}} \left(1 - \frac{r_{ij}^{\mu\nu}}{\sigma_{ij}^{\mu\nu}} \right). \quad (4.40)$$

As in the case of the pressure, we can therefore evaluate the shear stress Σ for any configuration by using the result of Eq. (4.40) in Eq. (4.32).

4.5.3 Dynamical matrix elements

In this Appendix, we derive the elements of the dynamical matrix for packings of DP and DPb particles. Throughout this subsection, we drop the cell-level indices μ and ν for clarity, and consider all vertices i directly for $i = 1, \dots, N_v$, where N_v is the total number of vertices in the configuration. Therefore, the dynamical matrix M of this configuration is defined as

$$M_{i\alpha j\beta} = \frac{\partial^2 U}{\partial \xi_{i\alpha} \partial \xi_{j\beta}} \quad (4.41)$$

where $\xi_{i\alpha}$ corresponds to the coordinate of vertex i in the $\alpha = x$ or y direction. The dynamical matrix M is therefore a symmetric, $2N_v \times 2N_v$ matrix. The total potential energy U takes contributions from individual potentials, i.e. U_a , U_l , U_b and U_{int} . Therefore, the dynamical matrix is the sum of dynamical matrices associated to each term, i.e.

$$M_{i\alpha j\beta} = M_{i\alpha j\beta}^a + M_{i\alpha j\beta}^l + M_{i\alpha j\beta}^b + M_{i\alpha j\beta}^{\text{int}} \quad (4.42)$$

where each dynamical matrix is the matrix of second derivatives of each potential energy terms, i.e. $M_{i\alpha j\beta}^a = \partial^2 U_a / \partial \xi_{i\alpha} \partial \xi_{j\beta}$.

To compute each matrix, note that each potential term depends on the degrees of freedom $\xi_{i\alpha}$ through geometric factors, e.g. the area or particular segment lengths. For example, suppose some term in the potential energy V depends on geometric factor Δ_i , associated to each vertex i , which in turn is a function of an arbitrary number of degrees of freedom. The vector of first derivatives of this potential (i.e. -1 times the vector of all net forces) would be

$$\frac{\partial V}{\partial \xi_{k\alpha}} = \sum_{i=1}^{N_v} \frac{\partial V}{\partial \Delta_i} \frac{\partial \Delta_i}{\partial \xi_{k\alpha}}. \quad (4.43)$$

Using the chain rule, the second derivatives of this potential can be written

$$\frac{\partial^2 V}{\partial \xi_{k\alpha} \partial \xi_{l\beta}} = \sum_{i=1}^{N_v} \left[\frac{\partial^2 V}{\partial \Delta_i^2} \frac{\partial \Delta_i}{\partial \xi_{k\alpha}} \frac{\partial \Delta_i}{\partial \xi_{l\beta}} + \frac{\partial V}{\partial \Delta_i} \frac{\partial^2 \Delta_i}{\partial \xi_{k\alpha} \partial \xi_{l\beta}} \right]. \quad (4.44)$$

While this simplifies our derivative, it also demonstrates two fundamental contributions to the dynamical matrix elements. If the potential V is quadratic in Δ , as all of our potentials are, the first term in Eq. (4.44) must be a constant. Therefore, all contributions from configuration geometry are associated with the first term, and contributions from stress (through the force-like term $\partial V / \partial \Delta_i$) are associated with the second term. Thus, these terms are called the “stiffness”

matrix \mathcal{H} and “stress” matrix \mathcal{S} [105, 106], which are typically defined as

$$\begin{aligned}\mathcal{H}_{k\alpha l\beta} &= \sum_{i=1}^{N_v} \frac{\partial^2 U}{\partial \Delta_i^2} \frac{\partial \Delta_i}{\partial \xi_{k\alpha}} \frac{\partial \Delta_i}{\partial \xi_{l\beta}} \\ \mathcal{S}_{k\alpha l\beta} &= - \sum_{i=1}^{N_v} \frac{\partial U}{\partial \Delta_i} \frac{\partial^2 \Delta_i}{\partial \xi_{k\alpha} \partial \xi_{l\beta}}\end{aligned}\tag{4.45}$$

such that the total dynamical matrix $M = \mathcal{H} - \mathcal{S}$. Prior work has shown that higher-order stability, like that found in systems with quartic modes, derives from dynamical matrices with significant contribution from the stress matrix [110]. We will sketch out the effect of the stiffness and stress matrices in the subsection 4.5.4.

Area term

For the area potential U_a , the relevant geometric quantity is the area strain $\Delta_a = (a/a_0 - 1)$. We therefore need to compute

$$\begin{aligned}\mathcal{H}_{k\alpha l\beta}^a &= \frac{\partial^2 U}{\partial \Delta_a^2} \frac{\partial \Delta_a}{\partial \xi_{k\alpha}} \frac{\partial \Delta_a}{\partial \xi_{l\beta}} \\ \mathcal{S}_{k\alpha l\beta}^a &= - \frac{\partial U}{\partial \Delta_a} \frac{\partial^2 \Delta_a}{\partial \xi_{k\alpha} \partial \xi_{l\beta}}.\end{aligned}\tag{4.46}$$

As a cell’s area only depends on the positions of the vertices associated with that cell, note that we must restrict the elements k and l to be those associated to a particular cell with a given area a . Given the definition of cell area in Eq. (4.10), the first derivatives of Δ_a w.r.t. positions are given by

$$\frac{\partial \Delta_a}{\partial \xi_{k\alpha}} = \frac{1}{2a_0} \begin{cases} y_{k+1} - y_{k-1} & \text{for } \alpha = x \\ x_{k-1} - x_{k+1} & \text{for } \alpha = y \end{cases}\tag{4.47}$$

where k only runs amongst the vertex indices associated with the given cell with area strain Δ_a . This yields the second derivatives

$$\frac{\partial^2 \Delta_a}{\partial \xi_{k\alpha} \partial \xi_{l\beta}} = \frac{1}{2a_0} \begin{cases} 1 & \text{for } \alpha = x, \quad l = k + 1, \quad \beta = y \\ -1 & \text{for } \alpha = x, \quad l = k - 1, \quad \beta = y \\ -1 & \text{for } \alpha = y, \quad l = k + 1, \quad \beta = x \\ 1 & \text{for } \alpha = y, \quad l = k - 1, \quad \beta = x \\ 0 & \text{otherwise.} \end{cases}\tag{4.48}$$

As vertices defined in polygonal rings, also note that $k + 1 = 1$ if $k = n$, and $k - 1 = n$ if $k = 1$. For simplicity, we define the vector $g_{k\alpha} \equiv 2a_0 \partial \Delta_a / \partial \xi_{k\alpha}$ and matrix $G_{k\alpha l\beta} = 2a_0 \partial^2 \Delta_a / \partial \xi_{k\alpha} \partial \xi_{l\beta}$

using equations Eqs. (4.47) and (4.48). Our matrix elements become

$$\begin{aligned}\mathcal{H}_{k\alpha l\beta}^a &= \frac{\epsilon_a}{2a_0^2} g_{k\alpha} g_{l\beta} \\ \mathcal{S}_{k\alpha l\beta}^a &= -\frac{\epsilon_a \Delta_a}{2a_0^2} G_{i\alpha j\beta}\end{aligned}\tag{4.49}$$

which give the full dynamical matrix for this term upon computing $M^a = \mathcal{H}^a - \mathcal{S}^a$.

Segment length term

For the area potential U_1 , the relevant geometric quantity is the strain on the length of segment i , i.e. $\Delta_{li} = (l_i/l_0 - 1)$. We therefore need to compute

$$\begin{aligned}\mathcal{H}_{k\alpha l\beta}^l &= \sum_{i=1}^{N_v} \frac{\partial^2 U}{\partial \Delta_{li}^2} \frac{\partial \Delta_{li}}{\partial \xi_{k\alpha}} \frac{\partial \Delta_{li}}{\partial \xi_{l\beta}} \\ \mathcal{S}_{k\alpha l\beta}^l &= -\sum_{i=1}^{N_v} \frac{\partial U}{\partial \Delta_{li}} \frac{\partial^2 \Delta_{li}}{\partial \xi_{k\alpha} \partial \xi_{l\beta}}.\end{aligned}\tag{4.50}$$

Note that the strain Δ_{li} only depends on $\xi_{k\alpha}$ when $i = k$ or $i = k-1$. The stiffness matrix elements in Eq. (4.50) are therefore

$$\mathcal{H}_{k\alpha l\beta}^l = \begin{cases} \frac{\partial^2 U}{\partial \Delta_{lk-1}^2} \frac{\partial \Delta_{lk-1}}{\partial \xi_{k\alpha}} \frac{\partial \Delta_{lk-1}}{\partial \xi_{k\beta}} + \frac{\partial^2 U}{\partial \Delta_{lk}^2} \frac{\partial \Delta_{lk}}{\partial \xi_{k\alpha}} \frac{\partial \Delta_{lk}}{\partial \xi_{k\beta}} & \text{for } l = k \\ \frac{\partial^2 U}{\partial \Delta_{lk-1}^2} \frac{\partial \Delta_{lk-1}}{\partial \xi_{k\alpha}} \frac{\partial \Delta_{lk-1}}{\partial \xi_{k-1\beta}} & \text{for } l = k-1 \\ 0 & \text{otherwise,} \end{cases}\tag{4.51}$$

and the stress matrix elements are therefore

$$\mathcal{S}_{k\alpha l\beta}^l = \begin{cases} -\frac{\partial U}{\partial \Delta_{lk-1}} \frac{\partial^2 \Delta_{lk-1}}{\partial \xi_{k\alpha} \partial \xi_{k\beta}} - \frac{\partial U}{\partial \Delta_{lk}} \frac{\partial^2 \Delta_{lk}}{\partial \xi_{k\alpha} \partial \xi_{k\beta}} & \text{for } l = k \\ -\frac{\partial U}{\partial \Delta_{lk-1}} \frac{\partial^2 \Delta_{lk-1}}{\partial \xi_{k\alpha} \partial \xi_{k-1\beta}} & \text{for } l = k-1 \\ 0 & \text{otherwise.} \end{cases}\tag{4.52}$$

Note that non-zero derivatives also exist when $l = k+1$, but these derivatives are simply the symmetric pairs of elements in the upper triangle. We will therefore only compute derivatives when $l = k-1$, and fill the lower triangle of the matrix by enforcement of symmetry. The first derivatives we need for Eq. (4.51) above are

$$\begin{aligned}\frac{\partial \Delta_{lk}}{\partial \xi_{k\alpha}} &= \frac{1}{l_0} \frac{\partial l_k}{\partial \xi_{k\alpha}} = -\frac{\hat{l}_{k\alpha}}{l_0} \\ \frac{\partial \Delta_{lk-1}}{\partial \xi_{k\alpha}} &= \frac{1}{l_0} \frac{\partial l_{k-1}}{\partial \xi_{k\alpha}} = \frac{\hat{l}_{k-1\alpha}}{l_0}\end{aligned}\tag{4.53}$$

where $\hat{l}_{k\alpha}$ is the α component of the unit vector pointing along the k th segment vector \vec{l}_k . As our second derivative matrices in Eq. (4.52) come from differentiating this result again, we require the derivative of a unit vector component $\hat{l}_{k\alpha}$ with respect to a degree of freedom $\xi_{j\beta}$, where j must be either k or $k+1$, but β need not be α . When $j = k$ and $\alpha = \beta$, we have

$$\frac{\partial \hat{l}_{k\alpha}}{\partial \xi_{k\alpha}} = \frac{l_k(\partial l_{k\alpha}/\partial \xi_{k\alpha}) - l_{k\alpha}(\partial l_k/\partial \xi_{k\alpha})}{l_k^2} = \frac{\hat{l}_{k\alpha}^2 - 1}{l_k} = -\frac{\hat{l}_{k\alpha'}}{l_k} \quad (4.54)$$

where $\alpha' \neq \alpha$ explicitly in this case, and the final equality comes from the fact that $\hat{l}_{kx}^2 + \hat{l}_{ky}^2 = 1$ by definition. When $\alpha \neq \beta$ we obtain

$$\frac{\partial \hat{l}_{k\alpha}}{\partial \xi_{k\beta}} = -\frac{l_{k\alpha}(\partial l_k/\partial \xi_{k\beta})}{l_k^2} = \frac{\hat{l}_{k\alpha}\hat{l}_{k\beta}}{l_k}. \quad (4.55)$$

Similar expressions can be found easily for the case where $l = k+1$. Therefore, our non-zero second derivatives are

$$\begin{aligned} \frac{\partial^2 \Delta_{lk-1}}{\partial \xi_{k\alpha}^2} &= \frac{\hat{l}_{k-1\alpha'}^2}{l_0 l_{k-1}}, & \alpha' \neq \alpha \\ \frac{\partial^2 \Delta_{lk-1}}{\partial \xi_{k\alpha} \partial \xi_{k\beta}} &= -\frac{\hat{l}_{k-1\alpha}\hat{l}_{k-1\beta}}{l_0 l_{k-1}}, & \beta \neq \alpha \\ \frac{\partial^2 \Delta_{lk}}{\partial \xi_{k\alpha}^2} &= -\frac{\hat{l}_{k\alpha'}^2}{l_0 l_{k-1}}, & \alpha' \neq \alpha \\ \frac{\partial^2 \Delta_{lk}}{\partial \xi_{k\alpha} \partial \xi_{k\beta}} &= \frac{\hat{l}_{k\alpha}\hat{l}_{k\beta}}{l_0 l_k}, & \beta \neq \alpha. \end{aligned} \quad (4.56)$$

Combining Eqs. (4.53) and (4.56), we obtain

$$\begin{aligned} \mathcal{H}_{k\alpha l \beta}^l &= \frac{\epsilon_l}{l_0^2} \begin{cases} \hat{l}_{k-1\alpha}\hat{l}_{k-1\beta} + \hat{l}_{k\alpha}\hat{l}_{k\beta} & \text{for } l = k \\ -\hat{l}_{k\alpha}\hat{l}_{k\beta} & \text{for } l = k+1 \\ 0 & \text{otherwise,} \end{cases} \\ \mathcal{S}_{k\alpha l \beta}^l &= \frac{\epsilon_l}{l_0} \begin{cases} \Delta_{lk} \frac{\hat{l}_{k\alpha'}^2}{l_k} - \Delta_{lk-1} \frac{\hat{l}_{k-1\alpha'}^2}{l_{k-1}} & \text{for } l = k & \beta = \alpha \\ \Delta_{lk-1} \frac{\hat{l}_{k-1\alpha}\hat{l}_{k-1\beta}}{l_{k-1}} - \Delta_{lk} \frac{\hat{l}_{k\alpha}\hat{l}_{k\beta}}{l_k} & \text{for } l = k & \beta \neq \alpha \\ -\Delta_{lk} \frac{\hat{l}_{k\alpha'}^2}{l_k} & \text{for } l = k+1 & \beta = \alpha \\ \Delta_{lk} \frac{\hat{l}_{k\alpha}\hat{l}_{k\beta}}{l_k} & \text{for } l = k+1 & \beta \neq \alpha \\ 0 & \text{otherwise.} \end{cases} \end{aligned} \quad (4.57)$$

In the above expressions, $\alpha' \neq \alpha$ always. Note that the stress matrix elements are all proportional to linear strain terms, i.e. Δ_{lk} , and therefore these terms are only non-zero in the presence of pre-stress (i.e. energy minima that do not fully satisfy all geometric constraints). As we discuss in Sec. 4.5.4, these terms will lead to higher-order rigidity.

Interaction term

To compute the dynamical matrix elements for the interaction term, note that the interaction potential between vertices k and l depends on their mutual overlap $\Delta_{kl} = (1 - r_{kl}/\sigma_{kl})$, where σ_{kl} is the contact distance between the two circular vertices. Note we have dropped the μ and ν indices for convenience from the expression given in Eq. (4.8). Therefore, we can proceed similarly for interaction terms as we did for perimeter terms with few changes. In general, the associated stiffness and stress matrices are given by

$$\begin{aligned}\mathcal{H}_{k\alpha l\beta}^{\text{int}} &= \sum_{i=1}^{N_v} \sum_{j>i}^{N_v} \frac{\partial^2 U}{\partial \Delta_{ij}^2} \frac{\partial \Delta_{ij}}{\partial \xi_{k\alpha}} \frac{\partial \Delta_{ij}}{\partial \xi_{l\beta}} \\ \mathcal{S}_{k\alpha l\beta}^{\text{int}} &= \sum_{i=1}^{N_v} \sum_{j>i}^{N_v} \frac{\partial U}{\partial \Delta_{ij}} \frac{\partial^2 \Delta_{ij}}{\partial \xi_{k\alpha} \partial \xi_{l\beta}}.\end{aligned}\tag{4.58}$$

Note that, as the interaction potential U_{int} in Eq. (4.8) depends on the step function $\Theta(\Delta_{ij})$, all derivatives of U with respect to Δ_{ij} of non-overlapping particles are 0. To simplify our work further, note that $\partial V_{ij}/\partial \xi_{i\alpha} = -\partial V_{ij}/\partial \xi_{j\alpha}$ by reciprocity of forces, where $V_{ij} = (\epsilon_c/2)\Delta_{ij}^2\Theta(\Delta_{ij})$ is the interaction potential between two vertices. Therefore, $\partial^2 V_{ij}/\partial \xi_{i\alpha} \partial \xi_{j\beta} = -\partial^2 V_{ij}/\partial \xi_{i\alpha} \partial \xi_{i\beta}$. As $U_{\text{int}} = \sum_{j>i} V_{ij}$, we must have

$$\frac{\partial^2 U}{\partial \xi_{i\alpha} \partial \xi_{i\beta}} = \sum_{j>i} \frac{\partial^2 U}{\partial \xi_{i\alpha} \partial \xi_{j\beta}}.\tag{4.59}$$

That is, the 2×2 block-diagonal elements of M^{int} are the negative sum of the block-off-diagonal elements purely by reciprocity of force. As $M^{\text{int}} = \mathcal{H}^{\text{int}} - \mathcal{S}^{\text{int}}$, and due to symmetry, we need only compute the terms $\mathcal{H}_{k\alpha l\beta}^{\text{int}}$ and $\mathcal{S}_{k\alpha l\beta}^{\text{int}}$ for the case when $l > k$. Note this holds for any potential described in this Section; the off-diagonal elements also balance the diagonal elements in all matrices shown in Eqs. (4.49), (4.57), and (4.70).

In the case where $l > k$, the first derivatives of Δ_{kl} with respect to relevant coordinates are

$$\begin{aligned}\frac{\partial \Delta_{kl}}{\partial \xi_{k\alpha}} &= \frac{\hat{r}_{kl,\alpha}}{\sigma_{kl}} \\ \frac{\partial \Delta_{kl}}{\partial \xi_{l\alpha}} &= -\frac{\hat{r}_{kl,\alpha}}{\sigma_{kl}}\end{aligned}\tag{4.60}$$

where $\hat{r}_{kl,\alpha}$ is the α component of the unit vector that points from vertex k toward vertex l , which follows from similar logic used to obtain Eq. (4.13). Using the result from Eq. (4.56), our second derivatives are

$$\begin{aligned}\frac{\partial^2 \Delta_{kl}}{\partial \xi_{k\alpha} \partial \xi_{l\alpha}} &= \frac{\hat{r}_{kl,\alpha'}^2}{\sigma_{kl} r_{kl}}, \quad \alpha' \neq \alpha \\ \frac{\partial^2 \Delta_{kl}}{\partial \xi_{k\alpha} \partial \xi_{l\beta}} &= -\frac{\hat{r}_{kl,\alpha} \hat{r}_{kl,\beta}}{\sigma_{kl} r_{kl}}, \quad \beta \neq \alpha.\end{aligned}\tag{4.61}$$

The stiffness and stress matrices associated to the interaction potential are therefore

$$\begin{aligned}\mathcal{H}_{k\alpha l\beta}^{\text{int}} &= -\epsilon_c \frac{\hat{r}_{kl,\alpha} \hat{r}_{kl,\beta}}{\sigma_{kl}^2} \\ \mathcal{S}_{k\alpha l\beta}^{\text{int}} &= \epsilon_c \begin{cases} -(\Delta_{kl} \hat{r}_{kl,\alpha'}^2)/(r_{kl} \sigma_{kl}) & \text{for } \alpha = \beta, \alpha' \neq \alpha \\ (\Delta_{kl} \hat{r}_{kl,\alpha} \hat{r}_{kl,\beta})/(r_{kl} \sigma_{kl}) & \text{for } \alpha \neq \beta \end{cases}\end{aligned}\quad (4.62)$$

for $l > k$. For the entries where $l = k$, we can use reciprocity, and for $l < k$, we can use symmetry to determine the other elements. Of course, to obtain the contribution to the full dynamical matrix due to the vertex-vertex interaction terms, we need simply compute $M^{\text{int}} = \mathcal{H}^{\text{int}} - \mathcal{S}^{\text{int}}$.

Bending term

Now we compute the contribution to the dynamical matrix elements due to bending rigidity. For convenience, let $s_k = |\vec{l}_k - \vec{l}_{k-1}|$ such that we can rewrite the total bending energy

$$U_b = \frac{k_b}{2l_0^4} \sum_{i=1}^{N_v} s_i^2. \quad (4.63)$$

The stiffness and stress matrices due to bending are then

$$\begin{aligned}\mathcal{H}_{k\alpha l\beta}^b &= \sum_{i=1}^{N_v} \frac{\partial^2 U_b}{\partial \kappa_i^2} \frac{\partial s_i}{\partial \xi_{k\alpha}} \frac{\partial s_i}{\partial \xi_{l\beta}} \\ \mathcal{S}_{k\alpha l\beta}^b &= - \sum_{i=1}^{N_v} \frac{\partial U_b}{\partial s_i} \frac{\partial^2 s_i}{\partial \xi_{k\alpha} \partial \xi_{l\beta}}.\end{aligned}\quad (4.64)$$

Notice that $\partial s_i / \partial \xi_{k\alpha}$ is only non-zero when $i = k-1, k$, or $k+1$. Therefore, the stiffness and stress matrices can be written

$$\begin{aligned}\mathcal{H}_{k\alpha l\beta}^b &= \begin{cases} \frac{\partial^2 U_b}{\partial s_{k-1}^2} \frac{\partial s_{k-1}}{\partial \xi_{k\alpha}} \frac{\partial s_{k-1}}{\partial \xi_{k\beta}} + \frac{\partial^2 U_b}{\partial s_k^2} \frac{\partial s_k}{\partial \xi_{k\alpha}} \frac{\partial s_k}{\partial \xi_{k\beta}} + \frac{\partial^2 U_b}{\partial s_{k+1}^2} \frac{\partial s_{k+1}}{\partial \xi_{k\alpha}} \frac{\partial s_{k+1}}{\partial \xi_{k\beta}} & \text{for } l = k \\ \frac{\partial^2 U_b}{\partial s_k^2} \frac{\partial s_k}{\partial \xi_{k\alpha}} \frac{\partial s_k}{\partial \xi_{k+1\beta}} + \frac{\partial^2 U_b}{\partial s_{k+1}^2} \frac{\partial s_{k+1}}{\partial \xi_{k\alpha}} \frac{\partial s_{k+1}}{\partial \xi_{k+1\beta}} & \text{for } l = k+1 \\ \frac{\partial^2 U_b}{\partial s_{k+1}^2} \frac{\partial s_{k+1}}{\partial \xi_{k\alpha}} \frac{\partial s_{k+1}}{\partial \xi_{k+2\beta}} & \text{for } l = k+2 \\ 0 & \text{otherwise.} \end{cases} \\ \mathcal{S}_{k\alpha l\beta}^b &= \begin{cases} \frac{\partial U_b}{\partial s_{k-1}} \frac{\partial^2 s_{k-1}}{\partial \xi_{k\alpha} \partial \xi_{k\beta}} + \frac{\partial U_b}{\partial s_k} \frac{\partial^2 s_k}{\partial \xi_{k\alpha} \partial \xi_{k\beta}} + \frac{\partial U_b}{\partial s_{k+1}} \frac{\partial^2 s_{k+1}}{\partial \xi_{k\alpha} \partial \xi_{k\beta}} & \text{for } l = k \\ \frac{\partial U_b}{\partial s_k} \frac{\partial^2 s_k}{\partial \xi_{k\alpha} \partial \xi_{k+1\beta}} + \frac{\partial U_b}{\partial s_{k+1}} \frac{\partial^2 s_{k+1}}{\partial \xi_{k\alpha} \partial \xi_{k+1\beta}} & \text{for } l = k+1 \\ \frac{\partial U_b}{\partial s_{k+1}} \frac{\partial^2 s_{k+1}}{\partial \xi_{k\alpha} \partial \xi_{k+2\beta}} & \text{for } l = k+2 \\ 0 & \text{otherwise.} \end{cases}\end{aligned}\quad (4.65)$$

The first derivatives we need are given by

$$\begin{aligned}
\frac{\partial s_{k-1}}{\partial \xi_{k\alpha}} &= \frac{l_{k-1\alpha} - l_{k-2\alpha}}{s_{k-1}} = \hat{s}_{k-1\alpha} \\
\frac{\partial s_k}{\partial \xi_{k\alpha}} &= -2 \frac{l_{k\alpha} - l_{k-1\alpha}}{s_k} = -2\hat{s}_{k\alpha} \\
\frac{\partial s_{k+1}}{\partial \xi_{k\alpha}} &= \frac{l_{k+1\alpha} - l_{k\alpha}}{s_{k+1}} = \hat{s}_{k+1\alpha},
\end{aligned} \tag{4.66}$$

where $\hat{s}_{k\alpha}$ is the α -component of the unit vector pointing in the direction of $\vec{s}_k = \vec{l}_k - \vec{l}_{k-1}$. As each first derivative is proportional to a unit vector, the second derivatives take on similar forms to those in Eq. (4.56). We have

$$\frac{\partial^2 s_k}{\partial \xi_{k\alpha}^2} = -2 \frac{\partial \hat{s}_{k\alpha}}{\partial \xi_{k\alpha}} = -2 \frac{s_k (\partial s_{k\alpha} / \partial \xi_{k\alpha}) - s_{k\alpha} (\partial s_k / \partial \xi_{k\alpha})}{s_k^2} = \frac{4 - 4\hat{s}_{k\alpha}^2}{s_k} = 4 \frac{\hat{s}_{k\alpha'}^2}{s_k} \tag{4.67}$$

where, as in Eq. (4.54), $\alpha' \neq \alpha$. When $\beta \neq \alpha$

$$\frac{\partial^2 s_k}{\partial \xi_{k\alpha} \partial \xi_{k\beta}} = -2 \frac{\partial \hat{s}_{k\alpha}}{\partial \xi_{k\beta}} = -2 \frac{s_k (\partial s_{k\alpha} / \partial \xi_{k\beta}) - s_{k\alpha} (\partial s_k / \partial \xi_{k\beta})}{s_k^2} = -4 \frac{\hat{s}_{k\alpha} \hat{s}_{k\beta}}{s_k}. \tag{4.68}$$

Similarly, we also have

$$\begin{aligned}
\frac{\partial^2 s_{k-1}}{\partial \xi_{k\alpha}^2} &= \frac{\partial \hat{s}_{k-1\alpha}}{\partial \xi_{k\alpha}} = \frac{\hat{s}_{k-1\alpha'}}{s_{k-1}} \\
\frac{\partial^2 s_{k-1}}{\partial \xi_{k\alpha} \partial \xi_{k\beta}} &= \frac{\partial \hat{s}_{k-1\alpha}}{\partial \xi_{k\beta}} = -\frac{\hat{s}_{k-1\alpha} \hat{s}_{k-1\beta}}{s_{k-1}} \\
\frac{\partial^2 s_{k+1}}{\partial \xi_{k\alpha}^2} &= \frac{\partial \hat{s}_{k+1\alpha}}{\partial \xi_{k\alpha}} = \frac{\hat{s}_{k+1\alpha'}}{s_{k+1}} \\
\frac{\partial^2 s_{k+1}}{\partial \xi_{k\alpha} \partial \xi_{k\beta}} &= \frac{\partial \hat{s}_{k+1\alpha}}{\partial \xi_{k\beta}} = -\frac{\hat{s}_{k+1\alpha} \hat{s}_{k+1\beta}}{s_{k+1}}
\end{aligned} \tag{4.69}$$

where, again, $\alpha' \neq \alpha$. The stiffness and stress matrix elements are therefore

$$\begin{aligned} \mathcal{H}_{k\alpha l\beta}^b &= \frac{k_b}{l_0^4} \begin{cases} \hat{s}_{k-1\alpha}\hat{s}_{k-1\beta} + 4\hat{s}_{k\alpha}\hat{s}_{k\beta} + \hat{s}_{k+1\alpha}\hat{s}_{k+1\beta} & \text{for } l = k \\ -2\hat{s}_{k\alpha}\hat{s}_{k\beta} - 2\hat{s}_{k+1\alpha}\hat{s}_{k+1\beta} & \text{for } l = k + 1 \\ \hat{s}_{k+1\alpha}\hat{s}_{k+1\beta} & \text{for } l = k + 2 \\ 0 & \text{otherwise.} \end{cases} \\ \mathcal{S}_{k\alpha l\beta}^b &= \frac{k_b}{l_0^4} \begin{cases} -\hat{s}_{k-1\alpha'}^2 - 4\hat{s}_{k\alpha'}^2 - \hat{s}_{k+1\alpha'}^2 & \text{for } l = k & \beta = \alpha \ (\alpha' \neq \alpha) \\ \hat{s}_{k-1\alpha}\hat{s}_{k-1\beta} + 4\hat{s}_{k\alpha}\hat{s}_{k\beta} + \hat{s}_{k+1\alpha}\hat{s}_{k+1\beta} & \text{for } l = k & \beta \neq \alpha \\ 2\hat{s}_{k\alpha'}^2 + 2\hat{s}_{k+1\alpha'}^2 & \text{for } l = k + 1 & \beta = \alpha \ (\alpha' \neq \alpha) \\ -2\hat{s}_{k\alpha}\hat{s}_{k\beta} - 2\hat{s}_{k+1\alpha}\hat{s}_{k+1\beta} & \text{for } l = k + 1 & \beta \neq \alpha \\ -\hat{s}_{k+1\alpha'}^2 & \text{for } l = k + 2 & \beta = \alpha \ (\alpha' \neq \alpha) \\ \hat{s}_{k+1\alpha}\hat{s}_{k+1\beta} & \text{for } l = k + 2 & \beta \neq \alpha \\ 0 & \text{otherwise.} \end{cases} \end{aligned} \quad (4.70)$$

As \mathcal{H} and \mathcal{S} are very similar, we can compute the dynamical matrix elements easily:

$$M_{k\alpha l\beta}^b = \frac{k_b}{l_0^4} \begin{cases} 6 & \text{for } l = k & \beta = \alpha \\ -4 & \text{for } l = k + 1 & \beta = \alpha \\ 1 & \text{for } l = k + 2 & \beta = \alpha \\ 0 & \text{otherwise.} \end{cases} \quad (4.71)$$

This much simpler expression could have been obtained by taking the Jacobian of the bending force in Eq. (4.19), but as we will see in the next section, decomposing the dynamical matrix into stiffness and stress matrices does aid in rationalizing the presence of non-isostatic jammed packings.

4.5.4 Counting degrees of freedom from vibrational modes

In this Appendix, we demonstrate how to count degrees of freedom and identify higher-order rigidity using the vibrational modes gleaned from the dynamical matrix. Prior work has shown that hypostatic systems like jammed non-spherical particles [105, 22, 107] and underconstrained spring networks [110] gain rigidity from higher-order terms in the potential energy that can stabilize multiple degrees of freedom. As noted in prior work [106], these higher-order constraints can be seen as zero modes of the stiffness matrix \mathcal{H} , but non-zero modes of the dynamical matrix \mathcal{M} . In Fig. 4.9 (a), we find that the number of missing contacts exactly equals the number of stiffness matrix eigenvalues λ_h that are significantly smaller than dynamical matrix eigenvalues λ_m . Since $\mathcal{M} = \mathcal{H} - \mathcal{S}$, we expect that small stiffness contribution means that \mathcal{M} is dominated by the stress matrix for these eigenvalues. Indeed, in Fig. 4.9 (a) we also show that we can tune the magnitude of λ_m by maintaining a fixed contact network but increasing the pressure. This pressure dependence of the low frequency eigenvalues of the dynamical matrix is also observed in packings of frictionless non-spherical particles [22].

In hypostatic DP packings, we find that the extra degrees of freedom are constrained by so-called “quartic modes”. Consider an energy-minimized configuration such that the particle coordinates satisfy $\vec{R} = \vec{R}_0$. Perturbations of order δ are then written as $\vec{R} = \vec{R}_0 + \delta\vec{u}$, where \vec{u} is the direction

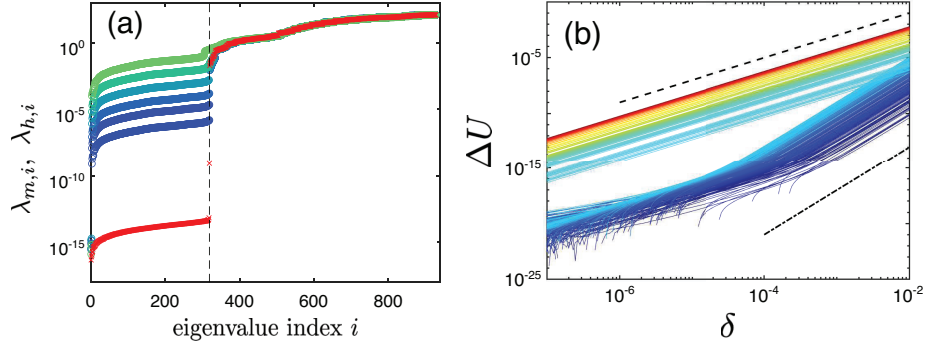


Figure 4.9: (a) Eigenvalues of the dynamical ($\lambda_{m,i}$, circles) and stiffness ($\lambda_{h,i}$, crosses) matrices as a function of eigenvalue index i for a configuration of $N = 16$ DP particles with $\mathcal{A}_0 = 1.02$. $\lambda_{h,i}$ do not depend on pressure, whereas $\lambda_{m,i}$ do. We show $\lambda_{m,i}$ for pressures from $P = 10^{-8}$ (blue) to 10^{-3} (green) separated by factors of 10. The vertical line is placed at $2 + m$. (b) Change in potential energy $\Delta U = U - U_0$ for a system starting at an energy minimum U_0 for perturbations of size δ along the eigenmodes of the dynamical matrix for the systems in (a). Color indicates mode frequency, from blue (lowest) to red (highest). The dot-dashed line represents $\Delta U \sim \delta^4$, whereas the dashed line represents $\Delta U \sim \delta^2$. There is a one-to-one correspondence between the number of modes with quartic δ -dependence in (b) and the pressure-dependent modes that are much larger than the stiffness matrix eigenvalues in (a).

of the perturbation. The potential energy expanded about \vec{R}_0 to fourth order in the perturbation is

$$\begin{aligned}
 U(\vec{R}) = U(\vec{R}_0) &+ \delta \frac{\partial U}{\partial R_i} u_i + \frac{\delta^2}{2} \frac{\partial^2 U}{\partial R_i \partial R_j} u_i u_j + \frac{\delta^3}{6} \frac{\partial^3 U}{\partial R_i \partial R_j \partial R_k} u_i u_j u_k \\
 &+ \frac{\delta^4}{24} \frac{\partial^4 U}{\partial R_i \partial R_j \partial R_k \partial R_l} u_i u_j u_k u_l + \dots,
 \end{aligned} \tag{4.72}$$

where we sum over repeated indices, and all derivatives are evaluated at \vec{R}_0 . The term linear in δ is 0 when the system is at a potential energy minimum. If we choose \vec{u} to be the k th orthonormal eigenvector of the dynamical matrix \mathcal{M} , the potential energy to second order in δ is $U = U_0 + \frac{1}{2} \lambda_{m,k} \delta^2$. However, in Fig. 4.9 (b), we find that the potential energy scales as $U \sim \delta^4$, which is consistent with the observation that $\lambda_{m,k}$ are small, but the quartic terms are non-negligible. Similar behavior is observed in jammed packings of non-spherical particles [104, 106, 22].

While we can easily identify the number of higher-order contacts from the dynamical matrix eigenspectra for packings of DP particles, the same is not true for packings of DPb particles. In Fig. 4.10, we show that the stiffness matrix eigenvalues λ_h are *larger* than the dynamical matrix eigenvalues λ_m for packings of buckled DPb particles. Given that buckled DPb particles have an extra degree of freedom, we investigate whether there is a signature in the eigenmodes with indexes below $m = 4N' - 1 - N_{\text{vv}}$, which corresponds to the number of missing contacts in packings of N' non-rattler particles.

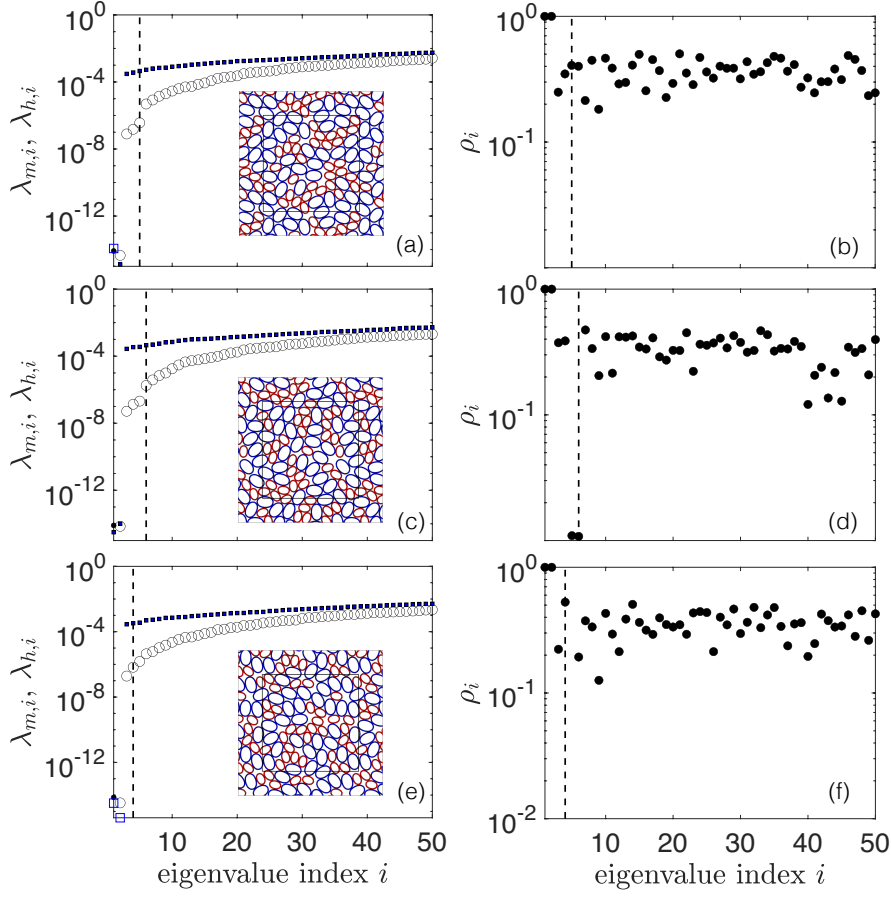


Figure 4.10: (a) The first 50 eigenvalues of the dynamical matrix (λ_m , black circles) and stiffness matrix (λ_h , blue squares) for the jammed packings of $N = 64$ DPb particles ($\mathcal{A}_0 = 1.08 > \mathcal{A}_0^*$, $K_b = 10^{-2}$) shown in the inset. In the inset, small particles with $n_S = 24$ vertices are drawn in red, and large particles with $n_L = 34$ vertices are drawn in blue. Open blue squares (closed black circles) represent negative stiffness (dynamical) matrix eigenvalues. The vertical line drawn at index $i = 6$, which for this system was $2 + m$. Note that all systems considered here have no rattler particles, so there are only 2 trivial zero modes. (b) The participation ratio ρ_i from Eq. (4.73) for each eigenmode of the packing in (a). The vertical line is also drawn at $i = 6$. (c) - (f) Same as (a) and (b), but for different jammed packings. Throughout, the vertical line is drawn at $2 + m$.

We develop the following heuristic for packings of DPb particles: first we check if there is a gap of at least a factor of 10 between the m th non-trivial eigenmode of the dynamical matrix and the next mode, since there is a gap between quartic and quadratic modes for packings of DP particles. If a gap in the eigenspectra of the dynamical matrix is not present, we calculate the participation

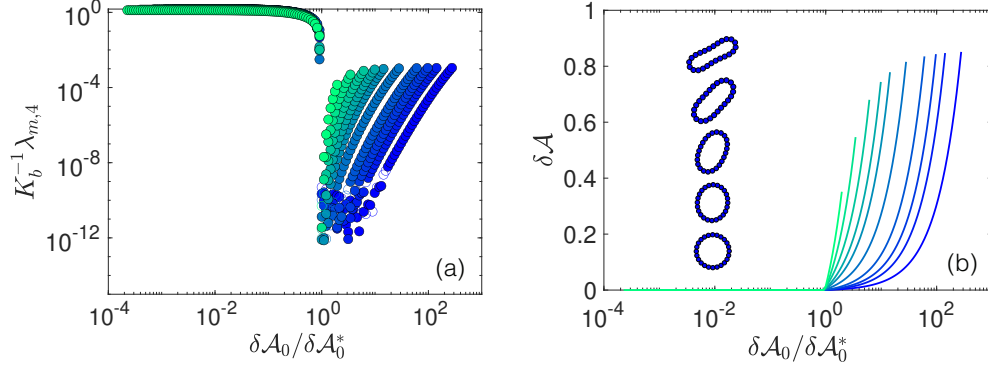


Figure 4.11: (a) The first non-trivial mode $\lambda_{m,4}$, scaled by the bending spring constant K_b , for DPb particles with $n = 24$ vertices as a function of the deviatoric preferred shape parameter $\delta\mathcal{A}_0 = (\mathcal{A}_0 - \mathcal{A}_n)/\mathcal{A}_n$ and different K_b . Closed symbols indicate $\lambda_{m,4} > 0$, and open symbols indicate $\lambda_{m,4} < 0$. Blue colors represent smaller K_b , starting at $K_b = 10^{-3}$, and green colors represent larger K_b , ending with $K_b = 2 \times 10^{-1}$. In both (a) and (b), the x -axis is scaled by $\delta\mathcal{A}_0^* = (\mathcal{A}_0^* - \mathcal{A}_n)/\mathcal{A}_n$. The inset shows $\delta\mathcal{A}_0^*$ vs. K_b ; \mathcal{A}_0^* is defined as the preferred shape parameter when $\lambda_{m,4} < 10^{-8}$. (b) The *true* deviatoric shape parameter $\mathcal{A} = p^2/4\pi a$ (i.e. $\delta\mathcal{A} = (\mathcal{A} - \mathcal{A}_n)/\mathcal{A}_n$) of the buckled DPb particles as a function of $\delta\mathcal{A}_0$. Colors are the same as in (a). The inset shows several representative particle shapes for $\delta\mathcal{A} = 0, 0.01, 0.1, 0.3$, and 0.6 . These shapes are the same for any K_b regardless of $\delta\mathcal{A}$.

ratio of each normal mode k ,

$$\rho_k = \frac{\left(\sum_{\mu=1}^N \sum_{i=1}^{n_\mu} |\vec{V}_{i\mu,k}|^2 \right)^2}{N \sum_{\mu=1}^N \sum_{i=1}^{n_\mu} |\vec{V}_{i\mu,k}|^4}, \quad (4.73)$$

where $\vec{V}_{i\mu,k}$ is the displacement direction of the i th vertex on the μ th particle in mode k . If we observe a gap in the participation ratio of at least a factor of 10 between the m th non-trivial mode and the next mode, we assume that we have identified the m higher-order modes correctly. We use the participation ratio gap because, in general, the participation ratio decreases with increasing eigenmode frequency. However, for DP particles, the highest-frequency quartic mode is usually localized whereas the lowest-frequency quadratic mode is delocalized. If neither of these two conditions is satisfied, we assume that there are no higher-order modes and no missing contacts. We emphasize that these thresholds are *ad hoc*, as the root cause of higher-order stability in jammed packings of buckled DPb particles is still an active area of research.

In Fig. 4.10, we show the outcome of this heuristic counting for several example configurations of $N = 64$ buckled DPb particles at jamming onset. For example, in Fig. 4.10 (a) and (b), we find

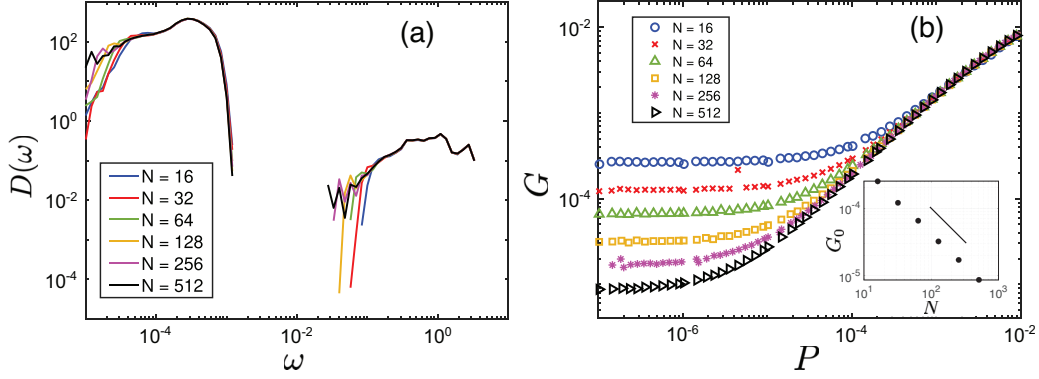


Figure 4.12: (a) Vibrational density of states $D(\omega)$ versus eigenmode frequency ω for jammed packings of bidisperse DP particles with $\mathcal{A}_0 = 1.08$ from $N = 16$ to $N = 512$ with $n_S = 16$. (b) Static shear modulus G versus pressure P for the same systems in (a). The inset shows G_0 (i.e. $G(P \rightarrow 0)$) versus system size N . The black solid line has the form $G_0 \sim N^{-1}$.

that a gap at the m th non-trivial mode appears, allowing us to identify these modes as higher-order constraints in analogy with packings of DP particles. We also can correctly identify higher-order constraints using the participation ratio as shown in Fig. 4.10 (c) and (d). However, we show in Fig. 4.2 (b) and in Fig. 4.10 (e) and (f) that there are several cases with missing contacts $m > 0$, but we cannot identify the missing contacts by analyzing the eigenvalue spectra or participation ratios. The fact that missing contacts cannot be counted consistently underscores the difficulty in identifying higher-order constraints in these jammed systems.

4.5.5 Particle buckling with bending energy

In this Appendix, we demonstrate buckling of DPb particles by increasing the preferred shape parameter \mathcal{A}_0 . In Fig. 4.11 (a), we show that the first non-trivial mode of the single-particle vibrational spectrum of DPb particles decreases by several orders of magnitude at \mathcal{A}_0^* , which depends on K_b . In Fig. 4.11 (b), we show that buckled particles transition from regular polygons with the true shape parameter $\mathcal{A} = \mathcal{A}_n$ to elongated, ellipsoidal particles with $\mathcal{A} > \mathcal{A}_n$ when $\mathcal{A}_0 > \mathcal{A}_0^*$. Note that occasionally, when sufficiently close to buckling, $\lambda_{m,4} < 0$, but $|\lambda_{m,4}|$ is close to numerical precision. These results suggest that, sufficiently close to buckling, DPb particles gain a degree of freedom and $\lambda_{m,4} \approx 0$. However, increasing \mathcal{A}_0 further causes the eigenvalue to grow in magnitude. While $\lambda_{m,4}$ remains significantly smaller than K_b , it is unclear whether DPb particles lose this degree of freedom at higher \mathcal{A}_0 .

4.5.6 System size dependence of the vibrational density of states and shear modulus

In this Appendix, we investigate the system size dependence in the vibrational density of states $D(\omega)$ and static shear modulus G of jammed packings of DP particles. In Fig. 4.12 (a), we show $D(\omega)$ for multiple system sizes spanning $N = 16$ to $N = 512$ with $n_S = 16$ for the preferred shape parameter $\mathcal{A}_0 = 1.08$. We see little change in the structure of $D(\omega)$ except for more low-frequency quartic modes for larger system sizes, though this does not seem to change the peaks in $D(\omega)$. We do however see system-size dependence in the static shear modulus as shown in Fig. 4.12 (b) in the low-pressure limit. At high pressures, G collapses across all system sizes studied, but as $P \rightarrow 0$ we show in the Fig. 4.12 (b) inset that $G(P \rightarrow 0) \sim N^{-1}$, which has been observed in previous work on deformable particles [47].

4.5.7 Mode decomposition

In this Appendix, we will show in detail how to decompose the eigenmodes of the dynamical matrix into contributions from particle translation, rotation, and shape degrees of freedom. We consider a packing of N deformable particles, where each particle μ has a center of mass located at $\vec{c}_\mu = n_\mu^{-1} \sum_{i=1}^{n_\mu} \vec{r}_{i\mu}$. Let \vec{V}^j be the unit vector corresponding to the j th eigenmode of the dynamical matrix \mathcal{M} in Cartesian coordinates. Individual components of \vec{V}^j are arranged such that the components from $2n_{\mu-1}$ to $2n_\mu$ are the n_μ x -coordinates followed by the n_μ y -coordinates for the μ th deformable particle. We can write three unit vectors to describe translation ($\hat{\mathbf{u}}_{\mu,x}$, $\hat{\mathbf{u}}_{\mu,y}$) and rotation ($\hat{\mathbf{u}}_{\mu,r}$) about the center of mass of the μ th particle as follows:

$$\hat{\mathbf{u}}_{\mu,x} = \frac{\vec{u}_{\mu,x}}{|\vec{u}_{\mu,x}|}, \vec{u}_{\mu,x} = (\underbrace{0, \dots, 0}_{1 \text{ to } (\mu-1)}, \underbrace{1, \dots, 1}_{\mu \text{th particle } x}, \underbrace{0, \dots, 0}_{\mu \text{th particle } y}, \underbrace{0, \dots, 0}_{(\mu+1) \text{ to } N}) \quad (4.74)$$

$$\hat{\mathbf{u}}_{\mu,y} = \frac{\vec{u}_{\mu,y}}{|\vec{u}_{\mu,y}|}, \vec{u}_{\mu,y} = (\underbrace{0, \dots, 0}_{1 \text{ to } (\mu-1)}, \underbrace{0, \dots, 0}_{\mu \text{th particle } x}, \underbrace{1, \dots, 1}_{\mu \text{th particle } y}, \underbrace{0, \dots, 0}_{(\mu+1) \text{ to } N}) \quad (4.75)$$

$$\hat{\mathbf{u}}_{\mu,r} = \frac{\vec{u}_{\mu,r}}{|\vec{u}_{\mu,r}|}, \vec{u}_{\mu,r} = (\underbrace{0, \dots, 0}_{1 \text{ to } (\mu-1)}, \underbrace{-(y_{1\mu} - c_{\mu,y}), \dots, -(y_{n_\mu\mu} - c_{\mu,y})}_{\mu \text{th particle } x}, \underbrace{x_{1\mu} - c_{\mu,x}, \dots, x_{n_\mu\mu} - c_{\mu,x}}_{\mu \text{th particle } y}, \underbrace{0, \dots, 0}_{(\mu+1) \text{ to } N}). \quad (4.76)$$

By defining the coefficients,

$$p_{\mu,x}^j = \vec{V}^j \cdot \hat{\mathbf{u}}_{\mu,x} \quad (4.77)$$

$$p_{\mu,y}^j = \vec{V}^j \cdot \hat{\mathbf{u}}_{\mu,y} \quad (4.78)$$

$$p_{\mu,r}^j = \vec{V}^j \cdot \hat{\mathbf{u}}_{\mu,r}, \quad (4.79)$$

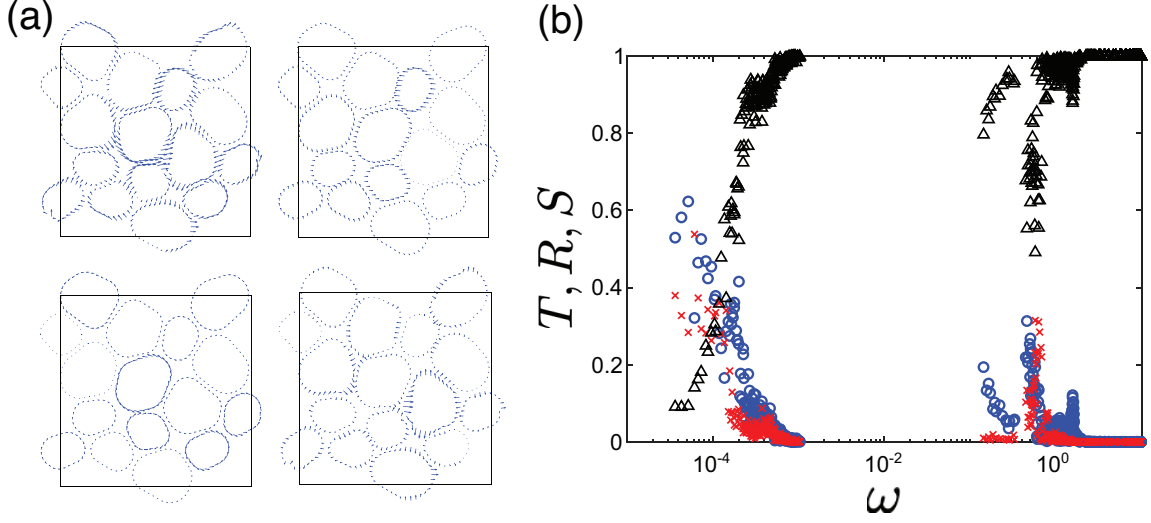


Figure 4.13: (a) Example of an eigenmode of the dynamical matrix (upper left) with frequency $\omega \approx 1.2 \times 10^{-4}$ for a packing of $N = 16$ DP particles ($n_S = 24$) with preferred shape parameter $\mathcal{A}_0 = 1.04$, and its decomposition into translation (upper right), rotation (lower left), and shape (lower right) degrees of freedom. The vectors are rescaled for clarity. (b) Contributions from translation T (blue circles), rotation R (red triangles), and shape S (black squares) degrees of freedom for each eigenmode as a function of the eigenmode frequency ω for the DP particle packing in (a).

we can rewrite the eigenvector \vec{V}^j as

$$\vec{V}^j = \sum_{\mu=1}^N p_{\mu,x}^j \hat{\mathbf{u}}_{\mu,x} + \sum_{\mu=1}^N p_{\mu,y}^j \hat{\mathbf{u}}_{\mu,y} + \sum_{\mu=1}^N p_{\mu,r}^j \hat{\mathbf{u}}_{\mu,r} + \vec{V}_s^j, \quad (4.80)$$

where \vec{V}_s^j is the vector that remains after subtracting the particle translations and rotation out of \vec{V}_j . By applying this decomposition, we can express each eigenmode as the sum of particle translations, rotation, and shape deformations. We show an example of an eigenmode decomposition in Fig. 4.13 (a) for a packing of DP particles.

With these coefficients, we can define the fraction of the translational (T^j) and rotational (R^j)

content in the j th eigenmode of the dynamical matrix as:

$$T^j = \sum_{\mu=1}^N \left[(p_{\mu,x}^j)^2 + (p_{\mu,y}^j)^2 \right] \quad (4.81)$$

$$R^j = \sum_{\mu=1}^N (p_{\mu,r}^j)^2. \quad (4.82)$$

Since we obtain $p_{\mu,x}^j, p_{\mu,y}^j$, and $p_{\mu,r}^j$ from unit vectors, $S^j = 1 - T^j - R^j$ gives the contribution of the shape degrees of freedom to the j th eigenmode. We show T^j, R^j , and S^j for a jammed packing of $N = 16$ DP particles with preferred shape parameter $\mathcal{A}_0 = 1.04$ in Fig. 4.13 (b) as a function of frequency ω , as well as just the S modes in Figs. 4.4 and 4.5. We find that for the quartic modes ($\omega < 10^{-2}$), the shape contribution S increases with ω , while T and R decrease. For the quadratic modes ($\omega > 10^{-2}$), the contribution from S is large, since higher frequency modes tend to deform the particle shape rather than give rise to translation or rotation.

Chapter 5

Beyond packing and jamming: diversity and development of spongy mesophyll

5.1 Introduction

The spongy mesophyll, or the tissue found in the interiors of leaves and flowers, plays a vital role in the life of plants and the health of terrestrial ecosystems (see Fig. 5.1). In leaves, this tissue provides a porous, rigid network of cells that allows carbon dioxide molecules from the atmosphere to diffuse to leaf cell chloroplasts for photosynthesis. As this tissue facilitates the diffusion of carbon dioxide and photosynthetic byproducts, i.e. oxygen, the leaf spongy mesophyll provides key infrastructure for the production of Earth’s atmosphere and carbon sequestration. While the leaf spongy mesophyll has played a vital role in forming our atmosphere, the spongy mesophyll in *flower petals* has also revolutionized the surface of the Earth. Flowering plants, or the angiosperms, are ecologically dominant in almost every terrestrial ecosystem, and are the most diverse group of vascular plants on the planet [137]. While the causes of the diversity and dominance of angiosperms has been a point of debate since the time of Darwin [138], an emerging consensus is that innovations in tissue organization is responsible for enabling heightened rates of metabolism [139] and the development of novel floral structures that are physiologically cheap and rapidly evolvable [140–143].

As flowers evolved from leaves, there is much to be learned in comparing the microstructural features of leaf and flower physiology [139]. However, studying these features is no easy task. In particular, spongy mesophyll is both *porous* and *disordered* in both leaves and flowers in ways that vary from species to species. High tissue porosity presents a developmental mystery, as the progenitor tissues of leaves and flowers begin their development as confluent packings of polyhedral cells [144, 145]. But this porosity varies drastically from species-to-species [146] (see also Fig. 5.1), begging the question of how cells with closely-related genomes can build such vastly different tissue architectures. In particular, data from microcomputed X-ray tomography (i.e. microCT, courtesy

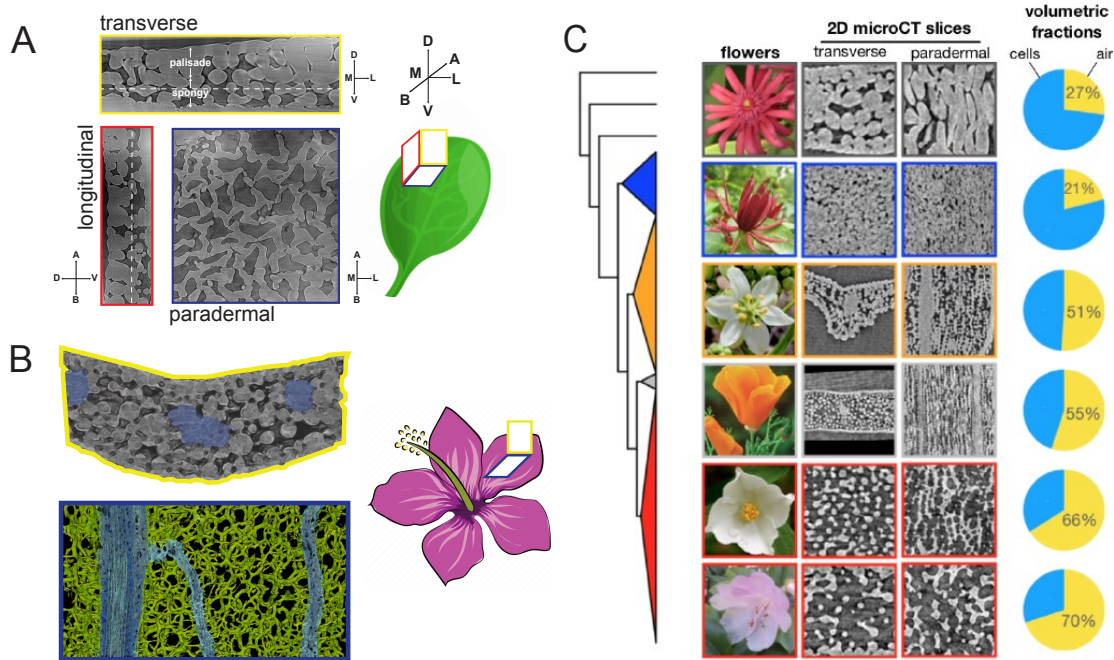


Figure 5.1: (A) Mesophyll tissue in a mature leaf of *Arabidopsis thaliana* via microCT (courtesy of L. Fletcher and C. Brodersen). Cell material is shown as light grey, and background is dark grey. Each perspective (transverse, longitudinal or paradermal) is color-coded as yellow, red and purple, as shown on the schematic leaf. Perspectives lie in one plane of the Dorsal-Ventral (DV, top-bottom), Medial-Lateral (ML, middle-side), Basal-Apical (BA, stem-tip) axes. Mesophyll cells are split into either the palisade or spongy layer; the paradermal slice shows spongy cells taken from a slice located at the dashed-line in the transverse slice. (B) Transverse (paradermal) perspective of the spongy mesophyll of a *Illicium floridanum* (*Cornus florida*) flower petal from the transverse (paradermal) perspective. Here, veins are false-colored in blue, and grey (green) signifies the spongy mesophyll cells. (C) Phylogenetic tree with structural traits for six representative species. From top to bottom, the species are *Illicium floridanum*, *Calycanthus floridus*, *Toxicoscordion fremontii*, *Eschscholzia californica*, *Philadelphus lewisii*, and *Rhododendron cilipense*. Here, box outline color represents species clade, i.e. basal angiosperms (blue), magnoliids (orange), monocots (green) and eudicots (red). Panel (C) courtesy of A. Roddy.

of A. Roddy) show that flower petal spongy mesophyll can vary from highly porous networks that resemble the microstructure of gels [147–150] to dense particulate solids and jammed emulsions [45]. This structural variation spans nearly the entire range of packing fraction accessible to most soft matter systems (e.g. gels, emulsion droplets, and foams) despite having similar developmental starting points, i.e. densely packed confluent tissue.

One hypothesis for how such diversity could emerge is that encoded in different genomes are distinct blueprints for how each tissue should be constructed, with plans for tissue porosity, cell

shape, cell connectivity, and mechanical response explicitly stored in genetic information. However, early lineages of flowering plants underwent a period of genome downsizing compared to the non-flowering plant species (e.g. gymnosperms and ferns) [139] at the same time that angiosperm diversity was increasing [151], which suggests that the loss of information stored in the genome is unrelated to floral morphological diversity. An alternative viewpoint is that these diverse phenotypes are not strictly encoded in the genome, but rather *self-assembled*. In this hypothesis, only a few key biophysical parameters are stored in genomic information and the tissue naturally arises as a result of tunable parameters that control the self-assembly process. For example, in flower petal mesophyll, variation in cell stiffness, cell deformability, cell-cell adhesion, and the difference between cell growth rate and tissue boundary growth rate could lead to a wide variety of structures. This is a highly efficient way to store developmental information in a genome, as opposed to the “genetic blueprint” idea where nearly all features of the developing tissue, such as cell connectivity and tissue porosity are stored and checked during development. While a combination of some genetic blueprinting and self-assembly is likely needed to understand the totality of decisions made during mesophyll development and evolution, studies of how biological components self-assemble using only a few physical mechanisms will yield insight into the design and synthesis of a new generation of smart, biologically-inspired materials [152].

Self-assembly has been studied in a host of diverse systems, particularly in colloidal solutions [153–157], liquid crystals [158], gels [159–162], and active matter [163–166, 3], as well in nanostructures with desirable electronic and mechanical properties [167, 168]. Self-assembly typically gives rise to stable structures [169, 150] with a high degree of tunability [170, 159, 171, 172]. Models of self-assembly have been previously applied to plant development, including the patterning of epidermal cells in leaves [173], the structure of spore wall complexes [174, 175], and pollen grain surfaces [176].

In order to understand the emergence of mesophyll through the lens of self-assembly, however, one must determine how parameters might be tuned to allow for *differences* of form to occur. That is, one must ask how does a given developmental trajectory change over time to give rise to an entirely new phenotype [177, 176]. The focus of this chapter will be to apply a framework of self-assembly to the development of diversity of spongy mesophyll tissue in order to determine how mesophyll microstructure might have been tuned over the course of evolution. In Sec. 5.2, we will evaluate the diversity of mesophyll three-dimensional microstructure across the angiosperm clade and determine how variable this tissue really is. Using structural correlations in Sec. 5.2.2, we will focus on tissue porosity, anisotropy and heterogeneity as key features that allow us to completely describe each tissue. In Sec. 5.2.2 we will compare observed microstructure to computational models of phase-separating liquids, and discuss how quantitative similarities and differences indicate the possible rules guiding mesophyll development. In Sec. 5.3, we discuss more involved computational modelling using deformable particles to simulate the development of spongy mesophyll tissue in two spatial dimensions. As this closely resembles planar spongy mesophyll expansion in developing leaves, we compare our model predictions to the developmental trajectory observed in developing *Arabidopsis thaliana* leaves. This work concludes with some indications to the biophysical origins of the diversity observed in the three-dimensional spongy mesophyll microstructure.

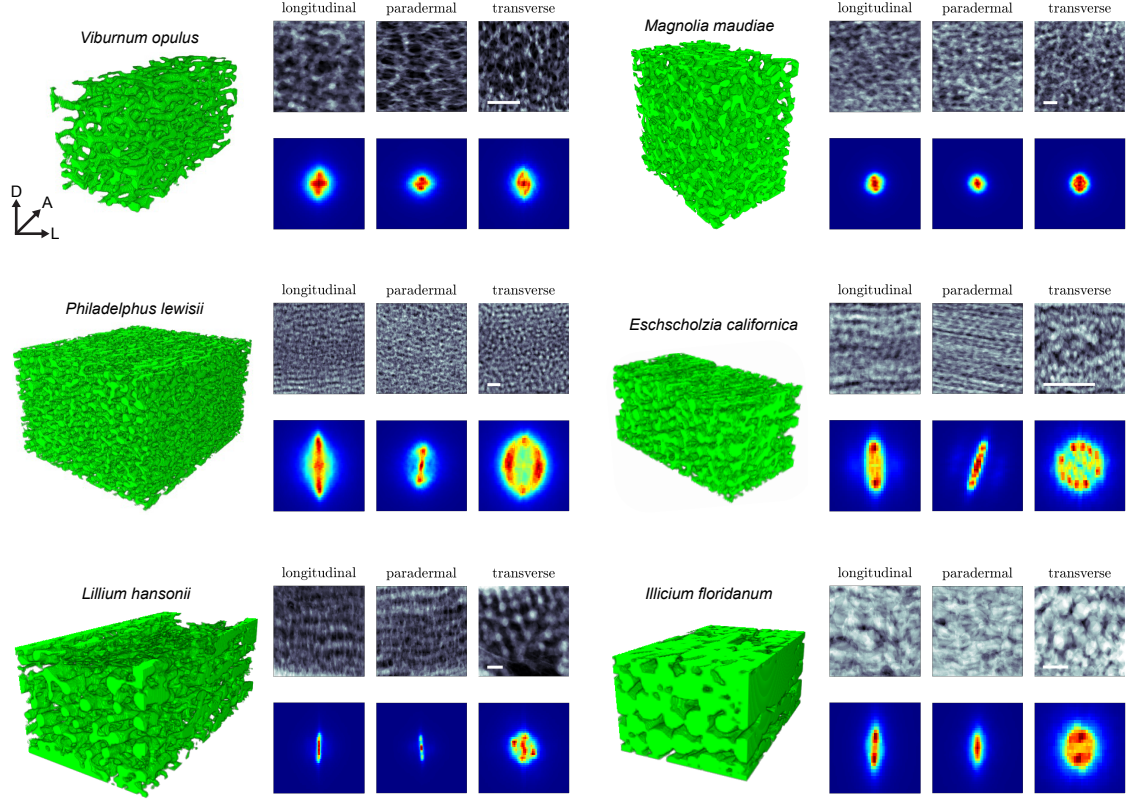


Figure 5.2: MicroCT samples (green) and corresponding $I(\vec{q})$ plots (blue-red panels) for six representative species. Each sample is oriented according to the axes shown at the top-left (A = apical, D = dorsal, and L = lateral). To the right of each microCT sample, we plot the intensity of both the microCT scan $\phi(\vec{x})$ and correlation function $I(\vec{q})$ averaged over a particular direction, leaving a 2D planar slice. See Fig. 5.1 for axis orientation. The flat white bar in each paradermal microCT slice corresponds to 100 μm ; note that the 3D microCT scans are not drawn to scale. In the 2D $I(\vec{q})$ plots, red (blue) corresponds to larger (smaller) intensity in arbitrary units. For all $I(\vec{q})$ plots, the origin $\vec{q} = \mathbf{0}$ is located in the plot's center.

5.2 Spongy mesophyll in flowers

In Fig. 5.1C, we see that the diversity in flower spongy mesophyll microstructure spans a broad range, from tissues denser than the densest packing of spheres, i.e. face-centered cubic packings of spheres (which have a packing fraction $\phi = \pi(3\sqrt{2})^{-1} \approx 0.741$), to tissues with nearly the porosity of hydrogels [149], i.e. $\phi < 0.3$. Interestingly enough, it would appear that flowers that have lower packing fractions, i.e. that are more porous, belong to more recent branches of the angiosperm phylogenetic tree. The flowers of *Calycanthus floridus*, a basal angiosperm more closely related to the angiosperm common ancestor, have mesophyll with packing fractions of nearly $\phi = 0.8$.

Whereas more recently-diverged groups like the eudicots (e.g. *Philadelphus* or *Rhododendron*) typically contain mesophyll at the opposite end of the density spectrum. Moreover, it appears that some species (i.e. *Illicium* and *Eschscholzia*) contain cells that predominantly align with the Basal-Apical axis, whereas the porous *Rhododendron* sample does not. These observations raise the question, do some phenotypes typically correlate with others? And do they also correlate with phylogeny?

5.2.1 Microstructural analysis

To address these questions, we analyze microCT scans of petal spongy mesophyll from nearly 50 angiosperm species, courtesy of A. Roddy and J. Schreel. Each scan is a binary three-dimensional image of the tissue, where voxels labeled as 1 represent regions in the interiors or on the surfaces of cells, whereas voxels labeled 0 represent regions of non-cell void space (i.e. gaseous atmosphere, which fills void space in living plant tissue). Let each image be defined as a function $\phi(\vec{x})$ that reports either 1 or 0 at every voxel location $\vec{x} = (x, y, z)$. We define the two-point autocorrelation function

$$C(\vec{x}) = \langle \phi(\vec{x}_0) \phi(\vec{x}_0 + \vec{x}) \rangle \quad (5.1)$$

where $\langle \rangle$ denotes an average over the origin \vec{x}_0 . If we assume that the image is statistically isotropic, then S measures the probability that two voxels separated by \vec{x} both contain cell material [178]. This function thus contains a wealth of information about how a given microstructure varies in space. In particular, it is useful to consider the Fourier transform $I(\vec{q})$, which takes the form

$$I(\vec{q}) = \frac{\tilde{\phi}(-\vec{q}) \tilde{\phi}(\vec{q})}{V} \quad (5.2)$$

where V is the image volume, $\tilde{\phi}$ is the Fourier transform of the image, and \vec{q} is a wavevector. We can also consider the spherically-averaged spectral correlation

$$I(q) = \frac{1}{4\pi q^2} \int_V d\vec{q} I(\vec{q}) \delta(\vec{q} - |\vec{q}|) \quad (5.3)$$

which measures the correlation of voxels in the Fourier-transformed image separated only by the wavevector distance q . Note that this quantity is typically measured in x-ray scattering experiments [179].

The spectra correlation function I allows us to identify samples with either isotropic or anisotropic microstructures. For example, in Fig. 5.2, we plot the two-dimensional cross-sections of $I(\vec{q})$ in the paradermal, longitudinal and transverse planes for different mesophyll samples. Some samples, e.g. *Magnolia Maudiae*, have a nearly spherical spectral correlation function, reflecting the fact that spatial correlations within the sample have no preferred directionality. Others, such as the *Philadelphus* and *Eschscholzia* samples, have anisotropic I and therefore anisotropic structures. For example, in *Eschscholzia*, the paradermal plane (averaging over the dorsal-ventral axis) shows highly-aligned cells along a near-horizontal direction. The projection of $I\vec{q}$ in this plane there shows correlation intensity perpendicular to the alignment direction, as correlations closer to the q -space

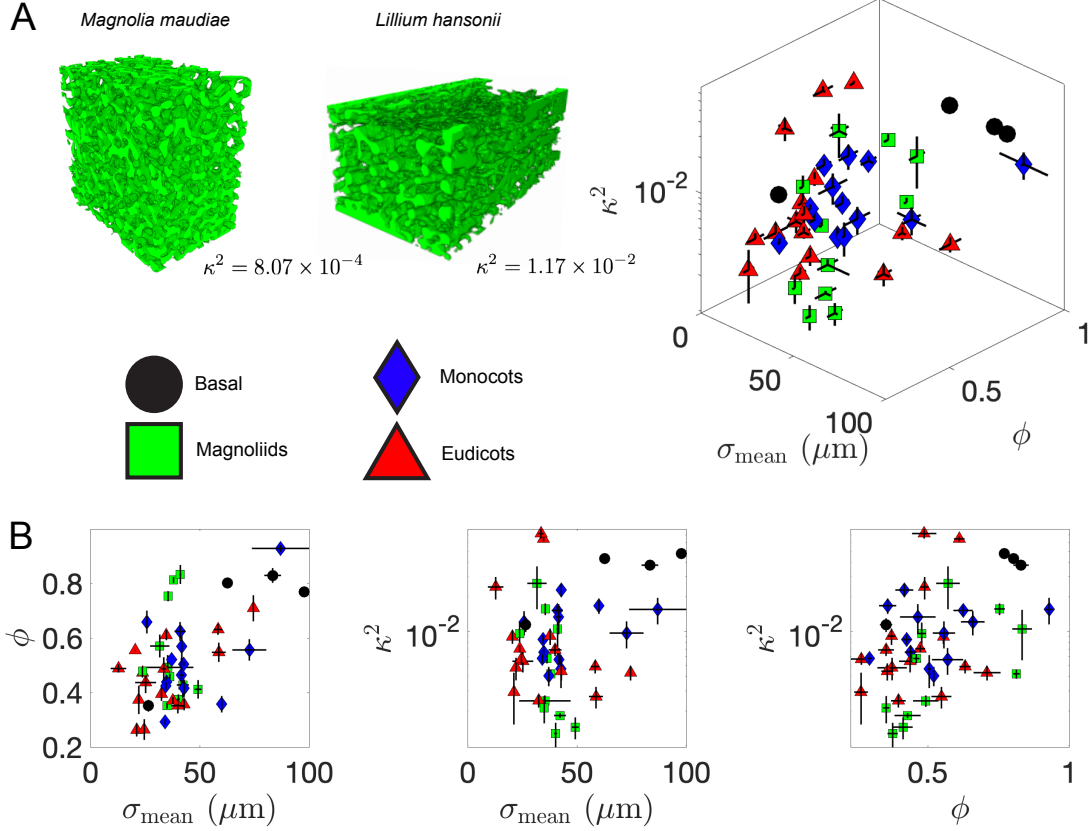


Figure 5.3: (A) Anisotropy κ^2 , packing fraction ϕ and mean cell diameter σ_{mean} plotted for all spongy mesophyll tissues analyzed. Each point represents a given species, with intraspecies variation given by the errorbars. Species clade are given by marker shape and color. We also show two representative species, *Magnolia Maudiae* and *Lilium hansonii* which demarcate the lower- and upper-bounds of anisotropy in our data set despite having similar densities ($\phi = 0.33$ and $\phi = 0.29$, respectively). (B) The same plot as in (A), just shown as three separate two-dimensional planes.

origin reflect correlations on longer scales in real (i.e. \vec{x}) space. Similar results can be found in non-living soft matter, such as nematic liquid crystals [179], and, interesting, copolymer melts that are uniaxially strained after undergoing microphase separation and arrest [180]. Such correlations are even more pronounced in the *Lilium* sample, which reflects the almost tubular structures taken by cells that lie in the basal-apical direction. Looking in the transverse plane, we see a completely different structure, which in the real-space projection looks like crystalline particles. Comparing the extreme anisotropy of the *Lilium* sample with the near-perfect isotropy of the *Magnolia* sample reinforces that the spongy mesophyll of flower petals is a highly tunable tissue that can take on a wide variety of structures. However, note also that several distantly related species like *Illicium*, *Lilium* and *Philadelphus* (which belong to the basal angiosperm, monocot and eudicot clades, respectively)

all have strong cell alignment along the apical-basal direction. That flower spongy mesophyll varies widely, yet can be similar across distantly-related groups, warrants a more detailed look at spongy mesophyll anisotropy across a broader diversity of angiosperms.

In order to quantify anisotropy in the spectral correlation function $I(\vec{q})$, we can define the gyration tensor

$$G_{\alpha\beta} = \frac{\int d\vec{q} q_\alpha q_\beta I(\vec{q})}{\int d\vec{q} I(\vec{q})} \quad (5.4)$$

which is simply the matrix of second moments of I , and is typically used (given a density field in real space) to study structure and fluctuations of polymers [181]. We can use the eigenvalues g_k to define the dimensionless anisotropy

$$\kappa^2 = \frac{3}{2} \frac{g_1^4 + g_2^4 + g_3^4}{(g_1^2 + g_2^2 + g_3^2)^2} - \frac{1}{2} \quad (5.5)$$

which can be understood as follows: if all g_k are equal, i.e. $g_1 = g_2 = g_3 = g$, then no particular direction has a larger variance of correlation than another, and Eq. (5.5) reduces to $\kappa^2 = (3/2)(3/9) - (1/2) = 0$. If only one eigenvalue dominates, i.e. if all correlation is only directed along a single axis, then Eq. (5.5) reduces to $\kappa^2 = 1$. If all correlation is directed in only two directions, i.e. there is a single zero-eigenvalue, $\kappa^2 = 1/4$. In effect, the eigenvalues of G define an ellipsoid that fits the shape of dominant correlations in I , and κ^2 measures the balance between the three ellipse axes.

In Fig. 5.3, we plot the anisotropy parameter κ^2 for each microCT sample of spongy mesophyll, as well as each sample's packing fraction ϕ and mean particle diameter σ_{mean} . Particle sizes were provided courtesy of A. Roddy and J. Schreel, and were obtained by (a rather laborious) manual analysis of each microCT scan. This plot demonstrates several interesting features. First, there seems to be positive correlation between packing fraction and anisotropy, which suggests that preferred cell geometry in denser samples tends to lead to a higher degree of cell alignment. While this correlation is not perfect, and different degrees of anisotropy coexist at intermediate values of ϕ , this hints that there may be physical constraints on anisotropy given a certain packing fraction. Even more surprising, ϕ seems to be well-correlated with mean cell size. Cell size would only *a priori* dictate packing density if flower petal size as a whole is well-conserved. Even so, cell sizes ($\sim 100 \mu\text{m}$) are of course much, much smaller than petals ($\geq 10 \text{ cm}$), which makes this result even more surprising. This raises the question, does cell size dictate packing density in the mesophyll? Or does cell size diversity reflect a different evolutionary innovation that is more important for changing packing density?

Interestingly, cell size and anisotropy seem poorly correlated, with samples varying by an order of magnitude in anisotropy containing cells of nearly the same size. Somehow, different mesophyll tissues can construct microstructures of widely varying anisotropy with cells of the same size, yet the cannot typically do so for tissues of varying porosity. Given that angiosperm cell size typically correlates with genome size, and that the angiosperms experienced a period of genome downsizing during their rise to dominance [139], the correlation in between tissue porosity and cell size may reflect evolutionary progress more porous, more resource-efficient tissues. Early-divergent

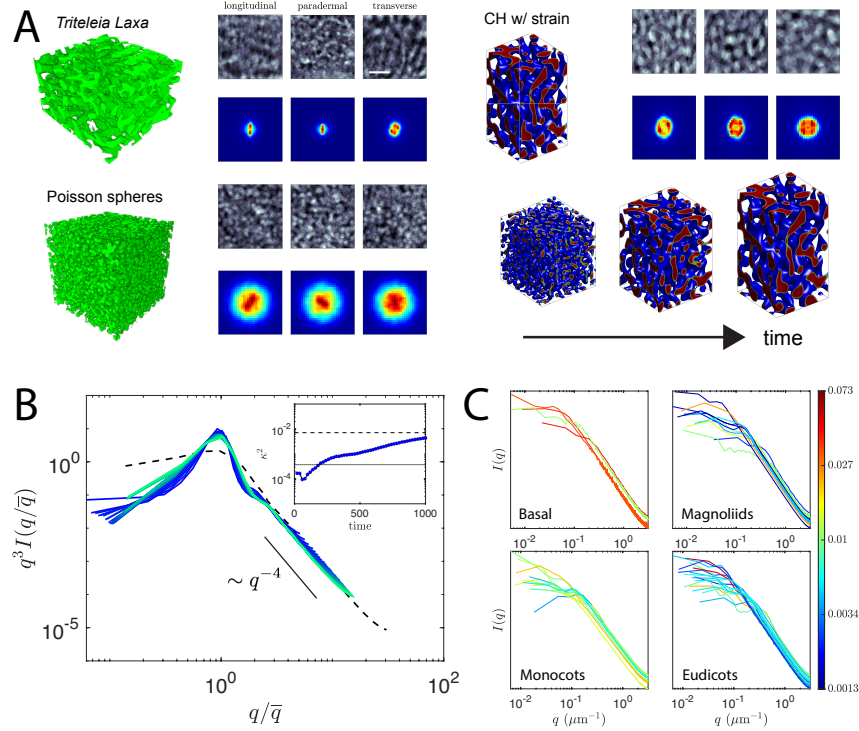


Figure 5.4: (A) Comparing the microstructure of *Triteleia laxa* with that of randomly placed Poisson spheres and the CH model with strain (Eq. (5.6)). (B) The spherically-averaged correlation function $I(q)$ for the CH model with strain shown in panel (A) over time (colored curves, sorted in time from blue to green) compared to $I(q)$ from *Triteleia laxa*. We scale q by the first moment $\bar{q} = \int dq q I(q) / \int dq I(q)$, and $I(q)$ by \bar{q}^3 as predicted by scaling theory [182]. In the inset, we show the anisotropy of the CH model over the course of a single simulation. The dotted line is the long-time anisotropy of the CH model without strain, and the dashed line is the anisotropy of *Triteleia laxa*. (C) $I(q)$ for spongy mesophyll samples averaged over individuals of a given species and sorted by clade, with curve color denoting the species-averaged κ^2 and given by the colorbar on the right.

tissues like the *Illicium* and *Calycanthus* (see Fig. 5.1) are typically dense and anisotropic, whereas more recently-diverged eudicot samples like *Viburnum* (see Fig. 5.2) are more porous and isotropic. However, important exceptions can be found in both cases, and the relationship between anisotropy, packing and evolutionary trajectories still remains unclear.

5.2.2 Computational modelling of anisotropic mesophyll

In order to contextualize these results, and to better understand if there are certain natural constraints on microstructural features, like anisotropy given a certain packing fraction for example, we turn to some simple computational models of complex microstructures. The simplest model for

a porous material is a collection spheres placed by a Poisson point process, which has a trivial (i.e. ideal gas) structure. In Fig. 5.4, we show that this model yields a structure with totally isotropic correlations. To build a model with more structural correlation and tunable anisotropy, we turn to simple dynamical models of phase-separating fluids. For instance, the Cahn-Hilliard (CH) model of phase separation [183] provides a mathematically simple mechanism to generate complex phase-separated domains. To implement this model, we consider a cubic L^3 grid with an initially random field $\psi(\vec{x})$ defined at each grid point \vec{x} . We then evolve this mixture in time using the CH equation

$$\frac{\partial \psi}{\partial t} = \frac{1}{2} \nabla^2 (\psi^3 - \psi - \nabla^2 \psi). \quad (5.6)$$

However, this model is by nature isotropic, and we expect κ^2 to remain small without some broken symmetry. To induce anisotropy, we perform a uniaxial extensional strain on the simulation domain at fixed time intervals τ and fix the boundary volume between strain steps. Strain is imposed in such a way as to preserve the shape of the function ϕ in the direction of strain as much as possible while still conserving the total mass $\bar{\psi} = L^{-3} \int d\vec{x} \psi(\vec{x})$. The addition of an effective strain rate (i.e. the inverse of the strain interval $\gamma = \tau^{-1}$) means we have a single tunable parameter with which to vary the microstructural details of the simulated phase separation. Note that without strain, the CH model is completely parameter-free, which leads to scale-invariant structure during the course of phase separation. It has been shown [182] that the spherical correlation function $I(q)$ has an invariant structure when the wavenumber q is scaled by the normalized first moment of I , i.e.

$$\bar{q} = \frac{\int dq I(q) q}{\int dq I(q)} \quad (5.7)$$

and the correlations are scaled by \bar{q}^d . In Fig. 5.4B, we show that adding a non-zero strain rate leads to slight breakdown of the scaling collapse, as well as emergent anisotropy during phase-separation. We also show that the increase in anisotropy (from $\kappa^2 \sim 10^{-4}$ to nearly 10^{-2}) that occurs during straining becomes on the order of the high degree of most flower mesophyll samples, e.g. *Triteleia laxa*, also shown in Fig. 5.4.

The qualitative agreement between this simple computational model and experimental spongy mesophyll samples (e.g. the sample of *Triteleia laxa* shown in Fig. 5.4) suggests that similar mechanisms may be at play during development of these complex tissues. While mesophyll is certainly not a phase-separating fluid, consider the high concentration phase (shown in blue in Fig. 5.4) as broadly cell material, while the low concentration phase is void space. Early stages of development in young flower petals are akin to well-mixed initial conditions where there is no well-defined void phase. As development progresses, small void domains begin to form and slowly grow until there is clear separation between void and cell in the porous mature state. As we will see in the next section, this mechanism is at play in leaf development, and suggests why the CH model captures some qualitative features of mesophyll structure. Similarly, strain-induced anisotropy may be present in developing flower petals as well. As strain is known to induce alignment and anisotropy in phase-separated materials [180], flower petal growth might provide an effective uniaxial strain to drive mesophyll cells toward aligned states. For example, if flower petal expansion in the lateral direction was somehow stiffer than expansion in the apical direction, cell and boundary growth

could cooperate in positive feedback cycles to drive more elongated, aligned cells.

However, there still exist clear differences between the CH with strain model and mesophyll samples. In particular, in Fig. 5.4B we see that model systems have well-peaked $I(q)$ curves throughout the entire simulation, reflecting consistent domain sizes even in the presence of anisotropic strain. This likely derives from the fact that the CH model has no free parameters to tune the competition between interface surface tension and diffusion. In contrast, almost all mesophyll samples across the angiosperm clades (Fig. 5.4C) have broadly-peaked $I(q)$ curves, indicating cell-rich domains with widely varying sizes. How, then, can these models be adapted to better reflect the complex microstructures observed in experiments? Recent work on randomly end-linked copolymer networks [184] demonstrated that similar broadening of the peak in $I(q)$ occurs in samples with copolymers of larger molecular weights and reflects networks with more structural heterogeneity. While the exact physical remains unknown, it could be that network heterogeneity results from the presence of microscopic disorder, i.e. chain-length disorder in the case of copolymer melts, or cell-size or shape disorder in the case of spongy mesophyll. However, work on explicit inclusion of quenched disorder into the dynamical CH equations [185] has shown that broad classes of quenched disorder merely delays the kinetic of ripening, rather than altering the structure of the phase-separated domains. An alternative origin of network heterogeneity could be the presence of viscoelastic forces during phase-separation. Computational studies of an extension of the CH model with elastic stresses [148] have demonstrated $I(q)$ peak-broadening and other structural effects when phase-separation kinetics are frustrated by elasticity. The inclusion of viscoelastic stresses makes physiological sense, as cell-cell interactions in the spongy mesophyll would certainly deter purely viscous flow of cells during development. After all, the mature mesophyll tissue is undoubtedly more solid-like than liquid-like. The extent to which models of viscoelastic phase-separation capture specific structural features of floral spongy mesophyll samples, let alone features of mesophyll development or diversity, require further study.

Comparisons of computational models to real spongy mesophyll samples in flower petals suggests potential origins for some of their microstructural features, but not all. In particular, the failure of CH models to describe heterogeneity of domain sizes indicates that such a simple model contains too few parameters to fully describe the diverse range of flower petal mesophylls. That is, the effective model that drives spongy mesophyll development in flower petals seems to have multiple intrinsic parameters. Though we have identified that external strain may play a role in driving tissue anisotropy, a more accurate model should be able to capture both domain-size heterogeneity the complex relationship between anisotropy and porosity. An important feature of these tissues, however, is not just their complex structural properties, but more importantly that these diverse structures are formed by the collective actions of individual cells. Whether encoded in the genome or a result of purely physical consequences, cells have to make decisions in order to properly sculpt these microstructures during development. Therefore, we will also explore the ability of computational models of discrete cells to capture mesophyll structural properties. In the next section, we will describe an application of the deformable polygon model described in the previous chapter to studying how dynamics at the single-cell level drive the emergence of complex phenotypes in model spongy mesophyll tissues.

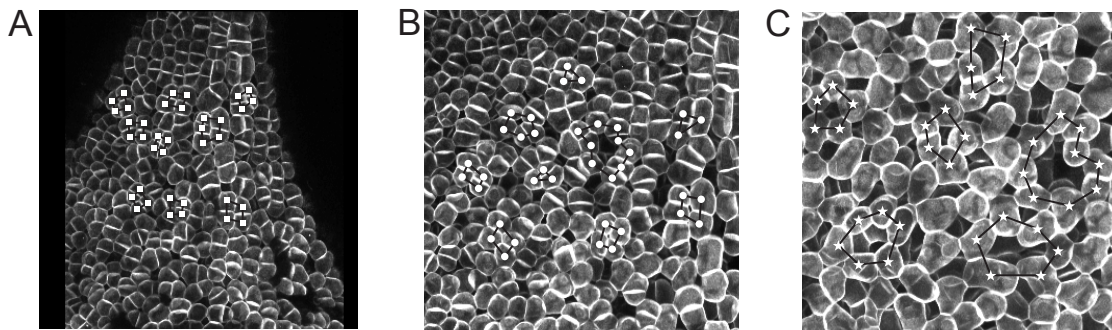


Figure 5.5: (A-C) Confocal microscopy images of the developing spongy mesophyll in *Arabidopsis thaliana* taken at (A) 0, (B) 24, and (C) 72 hours of development (courtesy of C. Ambrose, see Ref. [187] for experimental details). Cells walls are fluorescently labeled. Symbols and black lines are placed to highlight development of void shape from predominantly triangular (A), to rectangular (B) to varying polygonal order (C).

5.3 Spongy mesophyll in leaves

In order to understand the complex microstructure found in the spongy mesophyll, we can turn our attention from analyzing the bulk structural details in mesophyll found in different species to analyzing the cellular-level details of specific developmental trajectories. This perspective is complementary to our previous analyses of mature tissue diversity, as observing which single cell properties most influence tissue morphology is vital to understanding the biophysical origins of tissue tunability. That is, because cells make tissues, understanding tissues require understanding cells. As microCT technology cannot easily resolve borders between cells, technology like confocal microscopy is better suited to analyzing how cells change over the course of development. However, confocal microscopy require more time and perturbation to the tissue under study, making it difficult to study the variation of mesophyll across many different species. It is therefore paramount to utilize data from both experimental paradigms when informing and developing theoretical models. In this section, we will develop a model using deformable polygons to complement confocal microscopy experiments that can resolve the shapes of cells over the course of development. Current state-of-the-art experiments are restricted to observations in the leaves of the model plant *Arabidopsis thaliana*. While leaves have typically distinct mesophyll microstructures from flowers, they also are known to take on quasi-two-dimensional microstructures in maturity [99]. Therefore, we will use leaf development to guide a two-dimensional deformable polygon model to begin to try to study this complex process. In future work, we will use three-dimensional deformable polyhedra [186] to develop more complex three-dimensional developmental processes to mimic flower petal development.

Recent experimental advances in three-dimensional live-cell imaging have shed light on the development of complex cell shapes in both young and mature leaves [95, 96, 187]. As shown in Fig. 5.5, early mesophyll tissue is densely packed and nearly confluent, with each cell taking a convex

polygonal shape when seen in cross-section. However, as the tissue grows, airspaces between cells begin to form and the tissue transitions from confluent to porous. In the mature stage, cells in the spongy mesophyll form tortuous, quasi-two-dimensional (quasi-2D) networks. These pore spaces likely open up due to both cell separation and cell expansion [144], forces known to play a role in the pore space patterning in roots [188]. Precisely how these forces are coordinated, however, remains unclear. While it is known that rearrangement of microtubules near plant cell walls are often correlated with shape change and stress rearrangement during development [189, 190, 187], it is not known whether cells actively coordinate to pattern the pore spaces, or if these cell networks emerge from simple rules of growth and shape change [Cite new pore space rev]. Computational models of discrete, interacting cells can shed light on this question by allowing for the study of growth with few adjustable parameters, while simultaneously providing details on how growth and mechanics at the single cell level impact the microstructure of developing tissues.

To model the development of the spongy mesophyll in two spatial dimensions (2D), we employ numerical simulations of deformable polygons. Similarly to the previous chapter, our simulations contain a collection of N polygons with n vertices each. The dynamics of the polygons, which we will refer to as cells, are governed by forces due to change in shape and forces due to cell-cell interactions. The potential energy of the shape of a cell μ with n_μ vertices is given by

$$U_{\mu,\text{shape}} = \frac{\epsilon_a}{2} \left(\frac{a_\mu}{a_{0\mu}} - 1 \right)^2 + \sum_{i=1}^{n_\mu} \left[\frac{\epsilon_l}{2} \left(\frac{l_{i\mu}}{l_{0i\mu}} - 1 \right)^2 + \frac{\epsilon_b}{2} (\theta_{i\mu} - \theta_{0i\mu})^2 \right], \quad (5.8)$$

where ϵ_a , ϵ_l and ϵ_b control the extent to which geometric parameters can deviate from their preferred values. As there are n_μ vertices on each cell μ , we label each segment length $l_{i\mu}$ and bending angle $\theta_{i\mu}$ with $i = 1, \dots, n_\mu$. Note that Eq. (5.8) is similar to the shape energy in the previous chapter, with the notable exception of the new definition of the bending angle θ . As described below, the preferred bending angle θ_0 will play an integral role in mesophyll development, and the new bending angle definition provides a more convenient way to tune spontaneous curvature. A graphical representation of these quantities are provided in Appendix 5.5.1 as well as derivations of forces due to the new bending potential. The total potential energy U is then given by the shape potential energies plus the contribution from interactions, namely

$$U = \sum_{\mu=1}^N \left[U_{\mu,\text{shape}} + \sum_{\nu > \mu} \sum_{i=1}^{n_\mu} \sum_{j=1}^{n_\nu} \frac{\epsilon_{ij\mu\nu}}{2} \left(1 - \frac{r_{ij\mu\nu}}{\sigma_{ij\mu\nu}} \right)^2 \right]. \quad (5.9)$$

Here, we treat each vertex i on cell μ as a soft disk with diameter $\sigma_{i\mu}$. The parameter $\epsilon_{ij\mu\nu}$ controls the energy required for vertex i on cell μ to change the distance $r_{ij\mu\nu}$ to vertex j on cell ν . The contact distance $\sigma_{ij\mu\nu} = (\sigma_{i\mu} + \sigma_{j\nu})/2$ is the distance at which vertex i is just in contact with vertex j ; if $r_{ij\mu\nu} < \sigma_{ij\mu\nu}$, the vertices overlap, and vice versa. Note that, if $\epsilon_{ij\mu\nu} > 0$, the interacting vertices can either repel or attract. The interaction parameter $\epsilon_{ij\mu\nu}$ is on the order of the energy constant ϵ_c which controls the stiffness of vertex-vertex interactions. We then vary $\epsilon_{ij\mu\nu}$ according to if two vertices are interacting via short-range repulsion due to excluded volume (i.e. $r_{ij\mu\nu} < \sigma_{ij\mu\nu}$), or if two vertices are bonded in order to adhere two cells together. Throughout this

chapter, as in the previous chapter, energies are measured in units of ϵ_a , lengths are measured in units of the smallest square-root of cell area, i.e. $\min_{\mu} \sqrt{a_{0\mu}}$, all vertex masses are 1, and time is measured in units of $a_{0\mu, \min} \epsilon_a^{-1/2}$.

In order to simulate the emergence of cell networks during development, we need to model how intrinsic cell parameters (the individual l_0 , a_0 , etc) vary over time to drive changes to the resultant morphology. Recent experimental work [187] has demonstrated that mesophyll cell shape change during development is primarily driven by changes to the length of *void*-facing cell surfaces. In contrast, cell surfaces in contact with other cells remain nearly the same length during early stages of development. This localized growth of void-facing cell surfaces coincides with increased fluctuations and alignment of microtubules organized in parallel to cell edges near voids. Microtubules are known to play a vital role in driving cell shape in nearby tissues, e.g. the cells on the leaf epidermis [97], which indicates that they are playing a similar role in mesophyll tissue. Therefore, we hypothesize that localized changes to cell shape near perimeter *is sufficient to drive the emergence of complex networks observed in mature mesophyll*.

While this hypothesis is consistent with key observations of *in vivo* mesophyll development, an important question remains: how does the mesophyll tissue gain pores? As shown in both Figs. 5.5 and 5.6, *Arabidopsis* tissue becomes increasingly porous during development. As reported by Ref. [187], the first 120 hours of development from the primordial leaf sees a decrease in packing fraction ϕ from nearly unity to ~ 0.85 . The decrease in density coincides with transition in cell shapes, from initially polygonal at confluence (with $\mathcal{A} \approx 1.15$) to nearly circular (with $\mathcal{A} \approx 1.08$) to increasingly non-spherical ($\mathcal{A} \geq 1.15$). Note the observed asphericity at confluence $\mathcal{A} = 1.15$ is consistent with shapes observed in confluent packings of deformable polygons and polyhedra [47, 186]. As shown in Fig. 5.5, cell boundaries becomes increasing concave as porosity develops. In Fig. 5.6, we also show microCT scans of spongy mesophyll in more mature tissues with increasingly concave and complex shapes. Hand-segmentation of microCT scans demonstrates these cells have asphericity values that are consistent with trends observed earlier in development (see Fig. 5.6 (B) and (C)). Therefore, any model we develop to model this system needs to consistently increase the porosity $1 - \phi$ in the system throughout development.

One way to ensure that mesophyll density constantly decreases during the course of development would be to grow the box size at a constant rate. However, plant cells are known to maintain a constant, positive turgor pressure during development and into maturity [191], and constant strain rate would likely impose a *negative* pressure on the boundary on ultimately on the cells themselves. Therefore, instead of growth by constant *strain rate*, we impose a constant *stress* on the box and perform a growth simulation at minimum *enthalpy* rather than *energy*. We relax the total *enthalpy* $H = U + P_0 A$ for simulation domain area A at constant pressure P_0 and potential energy U . That is, the box area becomes a new degree of freedom that fluctuates in order to maintain that area is grown at a constant rate. In Appendix 5.5.2, we describe a modification of the FIRE algorithm [37, 192] that incorporates the treats the box size dynamically during the minimization protocol. This protocol ensures that we can always dictate the pressure of our mesophyll cell system for a given set of geometric parameters, i.e. a_0, l_{0i}, θ_{0i} , etc. In order to grow both the cells and the amount of void space in the tissue over time, we then modify the shape parameters of each cell that are localized near void space between enthalpy minimization steps. As void-facing surfaces

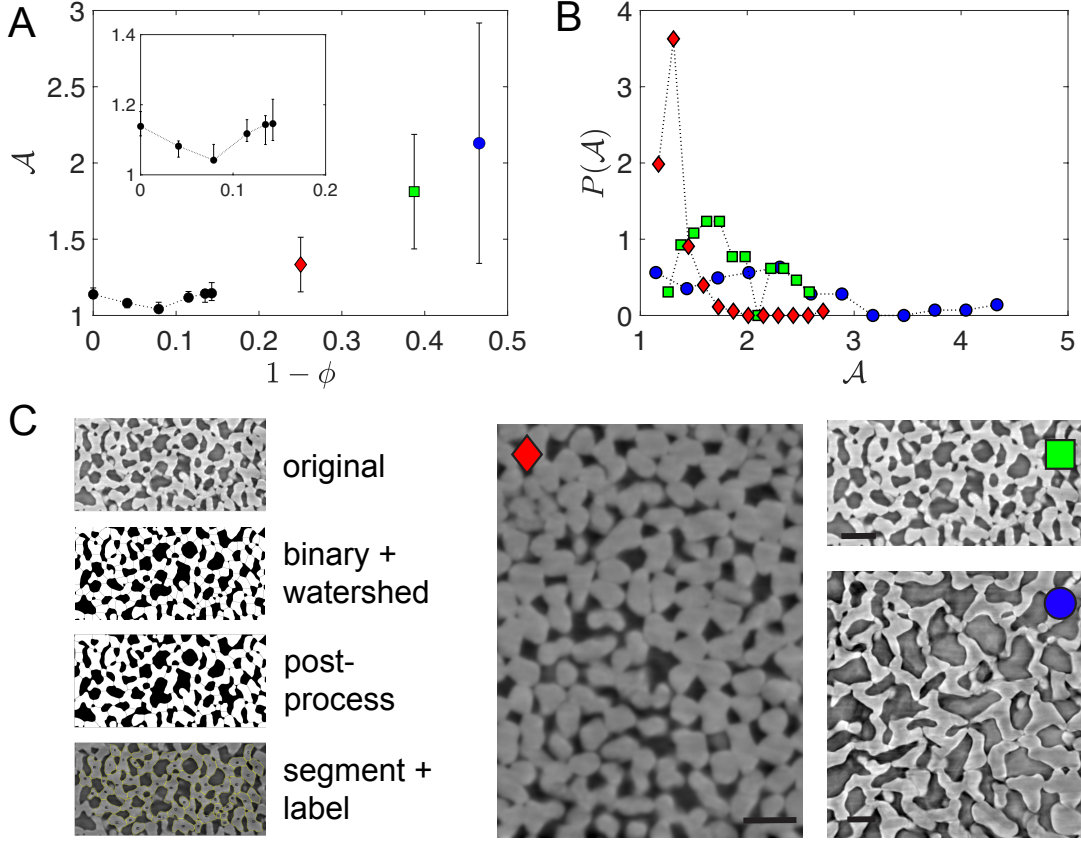


Figure 5.6: (A) Mesophyll cell shape asphericity \mathcal{A} over the course of development. Black symbols are taken from Ref. [187], while symbols with color are taken from microCT scans in panel (D). (B) Distribution of cell \mathcal{A} values from the microCT scans in panel (D). (C) Workflow of hand-segmenting individual mesophyll cell shapes from microCT scans. Using ImageJ, microCT data is binarized and initial cell boundaries are found using a watershed algorithm. Boundaries are then modified to better fit the assumed cell shapes from the microCT data using visual inspection. The microCT scans in (A) and (B) are shown at the right, courtesy of C. Brodersen, A. Roddy and G. Theroux-Rancourt.

grow longer and more concave during development, we grow all perimeter segments i on cells μ that border void space by a fraction δl_0 , and also drive the spontaneous curvature θ_{0i} of void-facing vertices toward a negative minimum value $\theta_{0,\min}$.

One additional feature that is required of any model we construct is the behavior seen in Fig. 5.6, where cells become more circular during early stages of development before increasing in \mathcal{A} . This actually indicates the presence of constant volumetric (in 2D, areal) growth of the cells in these early stages. When cells are densely-packed, there is very little void-facing perimeter for the cell

to grow. If cells are constantly growing their areas, however, the ratio $\mathcal{A} = p^2/4\pi a$ is therefore only *decreasing* in this stage, which drives cells to take on more circular shapes. Polygons can tile the plane, but circles cannot; this geometric incompatibility would then naturally increase the pressure and drive *boundary* growth until the cells form a tissue reminiscent of jammed circular particles. Once in this state, the circular cells have more void-facing perimeter, and the growth mechanism can kick in to drive the cells back toward more non-circular shapes. Therefore, between each enthalpy minimization step, we modify each preferred cell area $a_{0\mu}$, preferred segment length $l_{0i\mu}$ and preferred curvature $\theta_{0i\mu}$ using

$$\begin{aligned} a_{0\mu} &\rightarrow a_{0\mu}(1 + c\delta l_0)^2 \\ l_{0i\mu} &\rightarrow l_{0i\mu}(1 + \delta l_0) \quad \text{for } i \text{ void-facing} \\ \theta_{0i\mu} &\rightarrow \theta_{0i\mu} - \delta\theta_0 \quad \text{for } i \text{ void-facing.} \end{aligned} \tag{5.10}$$

We choose a form of the areal growth rate such that $c = 1$ corresponds to a constant shape parameter \mathcal{A} if all perimeter segments are grown in a given step, and the preferred angle does not change.

In Fig. 5.7, we show the results of a small simulation of four mesophyll cells growing in a small square boundary with rigid walls. Vertices interact both with their initial contact point on the wall and with each other via fixed, bonded spring potential given by the interaction term in Eq. (5.9). Bonds are modeled to have a fractional fracture length λ ; if the distance r_{ij} between two vertices i and j grows to be larger than $\lambda\sigma_{ij}$, where σ_{ij} is the mean of the radii of the two vertices, we disengage the bond. We set $\lambda = 1.1$ here to ensure that bonds can be maintain some extensional stress, but not so much that all bonds remain. We set all bond energies $\epsilon_{ij\mu\nu} = \epsilon_c$, and $\epsilon_c = \epsilon_l = 1$, as ϵ_a sets the unit of energy for our model. We set $\epsilon_b = 10^{-1}$, which is small enough to allow for deviations in the bending angle from the preferred curvature, but large enough to drive the system to more porous states. For growth rates, we set $\delta l_0 = \delta\theta_0 = 10^{-2}$ and $c = 0.4$, and we set a net pressure of $P_0 = 10^{-4}$ at the boundary, where P_0 is in units of ϵ_a/a_0 , and cells initially have preferred areas set to 1. We set the minimum bending angle $\theta_{0,\min} = -\pi/8$.

We find that our simple model of cell growth at constant pressure is sufficient to generate a qualitatively similar developmental trajectory to that which has been observed in experiments. Cells initially begin as convex polygons due to packing constraints, but quickly gain more circular shapes due to the constant increase in cell area without the presence of void-facing perimeter. These shapes eventually bottom out, however, and once there exists sufficient void-facing perimeter the cells begin to grow their asphericity. As shown in Fig. 5.7, the drive toward negative curvature pushes circular cells toward increasingly non-circular shapes as the bending energy stabilizes bending fluctuations and push out on the confining boundary. Note that cell sizes do not decrease during this, rather the cell sizes are grown larger given the positive value of the area scaling constant c . Therefore, the boundary grows in response to local changes to cell shape, effectively decreasing the packing fraction and increasing the amount of void space.

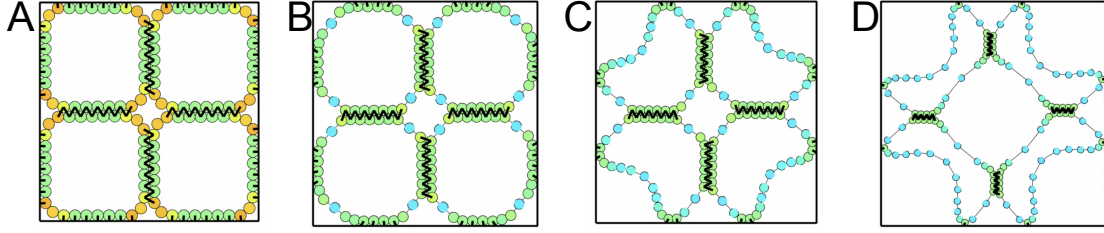


Figure 5.7: (A)-(D) progression of cell shape and contacts during the simulated developmental model described in the text. Vertex color corresponds to the local preferred bending angle $\theta_{0i\mu}$, where warmer colors correspond to larger values. Vertices outlined with dotted lines correspond to values of $\theta_{0i\mu} < \theta_{0,\min} = -\pi/8$. Black lines correspond to unbroken bonds between vertices or between a vertex and a pinned location on the wall. Each simulation is scaled to show an equal-sized boundary.

5.4 Conclusion

The development of mesophyll tissue in leaves and flowers is a complex process involving changes to cell shape and tissue density, variables which have been rarely studied together in biological contexts. In this chapter, we first analyzed the consequences of development (i.e maturity) across a wide range of mesophyll samples taken from flower petals. This demonstrated that whatever developmental process exists to form these tissues, it is tunable. Biophysical parameters at the single-cell level somehow can vary from species to species in order to generate variable tissue morphologies. We found a key morphological feature that can be tuned is tissue anisotropy. Each sample we analyzed are generally anisotropic, but that anisotropy varies greatly both across clades, and within the species members of a certain clade. Additionally, we found that tissue density was well-correlated with anisotropy and mean cell size, but that there are many tissues with the same size cells with differing anisotropies. This indicates that the biophysical mechanisms that generate these morphologies in development are potentially constrained in some ways, i.e. only larger cells typically make dense, anisotropic tissues, but are free to vary in others. Are these mechanisms constrained by the physical constraints of development? Or do these structures represent distinct pockets of fitness optima, and the developmental process can in principle generate a far greater diversity of structures with lower fitness? While the analyses shown here begin to shed light on these questions, more detailed knowledge of the developmental process is required in order to fully understand them.

Therefore, we introduced computational models that sought to understand the diversity and formation of these structures. We first introduced the strained Cahn-Hilliard model, which we showed was sufficient to generate structural anisotropy, but not domain-size heterogeneity. This indicates that boundary-driven stress is likely the origin of tissue anisotropy, but there must be additional free parameters to generate heterogenous structures. Seeing as these additional parameters are likely variable mechanical properties and active processes that occur at the single-cell level, we also introduced a computational model of mesophyll development using deformable polygons

in two spatial dimensions. While this two-dimensional model is insufficient to describe the full complexity of three-dimensional flower mesophylls, it provides both a simpler model in which we can develop intuition, and, fortuitously, a model of the quasi-two-dimensional development of leaf spongy mesophyll in the paradermal plane. We show that we can qualitatively capture developmental dynamics with deformable polygons that (a) grow at minimum enthalpy and (b) balance the rates of areal and local perimeter growth. This model helps explain the microscopic origins of pore space development in leaf spongy mesophyll, which until now have been a mystery [193].

Even more, the hypothesis that mesophyll tissues drive the development of pore spaces by localized growth opens up a new frontier in studying how individual cells can modify developmental trajectories to build different tissues. The next step for this work is to better understand the space of parameters described here in larger systems with more complex boundaries, i.e. in periodic boundaries to simulate a large bulk system, or with a deformable boundary mimicking a two-dimensional epidermis. But once the two-dimensional case is fully understood, the main test of this model would be to extend it to fully three-dimensional deformable polyhedra [186]. One could then study how changes to biophysical parameters (like shape growth rates, mechanical constants, and confining pressure) leads to anisotropy and domain-size heterogeneity. Are domains heterogeneous simply because of cell size polydispersity, or is there mechanical heterogeneity that leads to heterotypic branches in the mature cell network? Additionally, one potential origin of tissue anisotropy may be anisotropic boundary stiffness, i.e. some domain boundaries are easier to move than others. For example, in the example posed above, what if both the x and y boundaries were free to move independent of one another? Even more, one could set different effective masses for each of the confining boundary directions, making one more difficult to grow than the other. Would this be sufficient to generate anisotropic cell networks? Addressing these questions will allow us to understand how the space of possible single-cell biophysical parameters maps to the space of possible tissue morphologies. Such understanding broadens our ability to span broad scales of lengths and time and understanding the biophysical origins of complex biological phenomena like the robustness of certain developmental phenomena or even processes like evolution itself.

5.5 Appendices

5.5.1 Mesophyll bending force

The force on vertex k due to deviations in the local bending angle θ_k from its preferred value, i.e. $\delta\theta_k = \theta_k - \theta_{0k}$, is given by

$$\vec{F}_k^{\text{b}} = -\frac{\partial U_{\text{b}}}{\partial \vec{r}_k} = -\epsilon_b \left[\delta\theta_{k-1} \frac{\partial \theta_{k-1}}{\partial \vec{r}_k} + \delta\theta_k \frac{\partial \theta_k}{\partial \vec{r}_k} + \delta\theta_{k+1} \frac{\partial \theta_{k+1}}{\partial \vec{r}_k} \right]. \quad (5.11)$$

As shown in Fig. 5.8, we define the local bending angle at vertex k to be the angle between the vectors \vec{l}_{k-1} and \vec{l}_k . This angle can be calculated using the definitions $\sin \theta_k = \left| \hat{l}_{k-1} \times \hat{l}_k \right|$,

$\cos \theta_k = \hat{l}_k \cdot \hat{l}_{k-1}$, and $\theta_k = \tan^{-1} (\sin \theta_k / \cos \theta_k)$, i.e.

$$\theta_k = \tan^{-1} \left(\frac{l_{k-1,x} l_{k,y} - l_{k,x} l_{k-1,y}}{l_{k-1,x} l_{k-1,x} + l_{k,x} l_{k,y}} \right), \quad (5.12)$$

where $l_{k,\xi}$ is the $\xi = x$ or y component of the segment vector \vec{l}_k . We define the temporary variables

$$s_k = l_{k-1,x} l_{k,y} - l_{k,x} l_{k-1,y}; \quad c_k = l_{k-1,x} l_{k-1,x} + l_{k,x} l_{k,y}, \quad (5.13)$$

which gives

$$\frac{\partial \theta_k}{\partial \vec{r}_k} = \frac{c_k \frac{\partial s_k}{\partial \vec{r}_k} - s_k \frac{\partial c_k}{\partial \vec{r}_k}}{c_k^2 + s_k^2} \quad (5.14)$$

where we have used $\partial \tan^{-1}(x)/\partial x = (1+x^2)^{-1}$. The derivatives in the numerator of Eq. (5.14) are

$$\begin{aligned} \frac{\partial s_k}{\partial \vec{r}_k} &= (l_{k,y} + l_{k-1,y}) \hat{e}_x - (l_{k,x} + l_{k-1,x}) \hat{e}_y \\ \frac{\partial c_k}{\partial \vec{r}_k} &= (l_{k,y} - l_{k-1,y}) \hat{e}_x + (l_{k,x} - l_{k-1,x}) \hat{e}_y, \end{aligned} \quad (5.15)$$

and we note that $c_k^2 + s_k^2 = l_k^2 l_{k-1}^2$, given the definition of c_k and s_k in Eq. (5.13). Incorporating Eq. (5.15) into Eq. (5.14) yields

$$\frac{\partial \theta_k}{\partial \vec{r}_k} = \frac{\hat{n}_k}{l_k} + \frac{\hat{n}_{k-1}}{l_{k-1}}, \quad (5.16)$$

where $\vec{n}_k = l_{k,y} \hat{e}_x - l_{k,x} \hat{e}_y$ is the vector normal to \vec{l}_k , and is drawn in red as the unit vector $\hat{n}_k = \vec{n}_k / l_k$ in Fig. 5.8.

Because each polygonal cell forms a closed loop, it is always true that $\sum_{i=1}^n \theta_i = 2\pi$, and thus $\theta_k = 2\pi - \sum_{i \neq k} \theta_i$. Therefore,

$$\frac{\partial \theta_k}{\partial \vec{r}_k} = -\frac{\partial \theta_{k-1}}{\partial \vec{r}_k} - \frac{\partial \theta_{k+1}}{\partial \vec{r}_k}, \quad (5.17)$$

as θ_k only depends on \vec{r}_{k-1} , \vec{r}_k and \vec{r}_{k+1} . Comparing coefficients with Eq. (5.16), we then have the derivatives of the other two angles with respect to \vec{r}_k :

$$\frac{\partial \theta_{k-1}}{\partial \vec{r}_k} = -\frac{\hat{n}_{k-1}}{l_{k-1}}, \quad \frac{\partial \theta_{k+1}}{\partial \vec{r}_k} = -\frac{\hat{n}_k}{l_k}. \quad (5.18)$$

Note this can also be obtained by taking each derivative directly. We can then substitute the forms of the derivatives of each angle with respect to \vec{r}_k into Eq. (5.11) to obtain

$$\vec{F}_k^b = \epsilon_b \left[\left(\frac{\delta \theta_{k-1} - \delta \theta_k}{l_{k-1}} \right) \hat{n}_{k-1} + \left(\frac{\delta \theta_{k+1} - \delta \theta_k}{l_k} \right) \hat{n}_k \right]. \quad (5.19)$$

Bending forces therefore typically point normal to the cell surface.

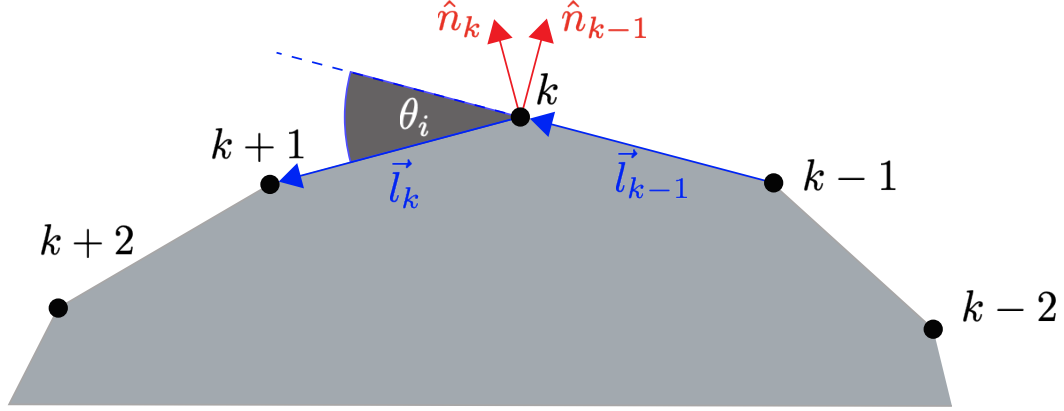


Figure 5.8: Geometry of regions near individual vertices on model cells. The bending angle θ_k is defined to be the angle between \vec{l}_{k-1} and \vec{l}_k . The unit vectors \hat{n}_k and \hat{n}_{k-1} drawn in red are normal to \vec{l}_k and \vec{l}_{k-1} , respectively.

5.5.2 Derivation of enthalpy minimization

Our simulations will minimize a particular configuration's enthalpy $H = U + P_0 L_x L_y$, where U is the total potential energy described in Eq. (5.9), P_0 is the constant constraint pressure, and L_x and L_y are the side lengths of our simulation domain in the horizontal (x) and vertical (y) directions of our simulation, respectively. By configuration, we mean the set of coordinates $\vec{r}_{i\mu}$ for all vertices $i = 1, \dots, n_\mu$ on all cells $\mu = 1, \dots, N$. In this work, we only consider a square domain that can change size isotropically, i.e. $L_x = L_y = L$. However, this protocol can be extended to domains where L_x and L_y may vary independently.

To minimize a given configuration's enthalpy H , we combine dynamics in the isoenthalpic-isobaric ensemble first described by Andersen [194] with the minimization dynamics of the FIRE algorithm [37]. Isoenthalpic-isobaric dynamics (or NPH dynamics, as opposed to classical NVE dynamics with constant particle number, domain volume, and total internal energy) are obtained by treating the total domain size $A = L^2$ as an additional dynamical quantity with inertial mass M . In addition to momenta $\vec{p}_{i\mu}$ and masses $m_{i\mu}$ for all vertices i on cells μ , we define the domain "momentum" $\Pi = M \partial A / \partial t$. As described by Andersen, one obtains the following equations of

motion for NPH dynamics in two spatial dimensions:

$$\begin{aligned}
\frac{\partial \vec{r}_{i\mu}}{\partial t} &= \frac{\vec{p}_{i\mu}}{m_{i\mu}} + \frac{\Pi}{2MA} \vec{r}_{i\mu} \\
\frac{\partial A}{\partial t} &= \frac{\Pi}{M} \\
\frac{\partial \vec{p}_{i\mu}}{\partial t} &= -\frac{\partial U}{\partial \vec{r}_{i\mu}} - \frac{\Pi}{2MA} \vec{p}_{i\mu} \\
\frac{\partial \Pi}{\partial t} &= P(t) - P_0
\end{aligned} \tag{5.20}$$

where $P(t)$ is the instantaneous pressure at a given time t . If truly simulating a dynamical system at constant enthalpy and pressure, the pressure P must take into account both kinetic and potential contributions [75]. However, as our growth simulations are athermal and quasistatic, we only consider the potential contribution to the pressure and define $P = -dU/dA$, or the total derivative of the potential energy with respect to the domain area. We use the form of the pressure described in Eq. (4.24) in the previous chapter, except note that in this chapter, we must compute the partial derivative $\partial\theta_k/\partial L$. Because θ_k is non-dimensional and does *not* change with change of the simulation domain area, there is no bending energy contribution to the pressure.

While the dynamics in Eq. (5.20) lead to dynamic sampling of the NPH ensemble, we require that each step in our mesophyll development simulation be at a local minimum of enthalpy. Therefore, we adapt these NPH dynamics into the dynamics of the FIRE algorithm. Standard FIRE minimization amounts to standard NVE dynamics, where conservative forces $\vec{F}_{i\mu} = -\partial U/\partial \vec{r}_{i\mu}$ drive the system to minimize the potential energy, along with a force that increases momentum if the net force is decreasing in magnitude, but decreases momentum if the force magnitude increases. The dynamical equations for systems in domains of fixed size are

$$\begin{aligned}
\frac{\partial \vec{r}_{i\mu}}{\partial t} &= \frac{\vec{p}_{i\mu}}{m_{i\mu}} \\
\frac{\partial \vec{p}_{i\mu}}{\partial t} &= \vec{F}_{i\mu} - \gamma(t) \left(\vec{p}_{i\mu} - \frac{\pi}{\Phi} \vec{F}_{i\mu} \right),
\end{aligned} \tag{5.21}$$

where π and Φ are the magnitudes of $\vec{\pi}$ and $\vec{\Phi}$, which are the vectors

$$\begin{aligned}
\vec{\pi} &= (p_{11,x} \quad p_{11,y} \quad p_{21,x} \quad \cdots \quad p_{n_N N,x} \quad p_{n_N N,y}) \\
\vec{\Phi} &= (F_{11,x} \quad F_{11,y} \quad F_{21,x} \quad \cdots \quad F_{n_N N,x} \quad F_{n_N N,y}),
\end{aligned} \tag{5.22}$$

i.e. the vectors of all momenta and net forces for all degrees of freedom, respectively. The variable $\gamma(t)$ is updated at each time to tune whether the direction of $\vec{\pi}$ increases Φ (i.e. $\gamma > 0$) or decreases Φ (i.e. $\gamma < 0$).

To introduce dynamics that minimize the enthalpy, we include the domain size momentum Π into the definition $\vec{\pi}$, and the domain boundary force $P(t) - P_0$ into the definition of $\vec{\Phi}$. We also use the domain-adjusted forces $\vec{F}'_{i\mu} = -\partial U/\partial \vec{r}_{i\mu} - (\Pi/2MA)\vec{p}_{i\mu}$ in the global force vector $\vec{\Phi}$. Our

final dynamical equations are then

$$\begin{aligned}
\frac{\partial \vec{r}_{i\mu}}{\partial t} &= \frac{\vec{p}_{i\mu}}{m_{i\mu}} + \frac{\Pi}{2MA} \vec{r}_{i\mu} \\
\frac{\partial A}{\partial t} &= \frac{\Pi}{M} \\
\frac{\partial \vec{p}_{i\mu}}{\partial t} &= \vec{F}'_{i\mu} - \gamma(t) \left(\vec{p}_{i\mu} - \frac{\pi}{\Phi} \vec{F}'_{i\mu} \right) \\
\frac{\partial \Pi}{\partial t} &= P(t) - P_0 - \gamma(t) \left[\Pi - \frac{\pi}{\Phi} (P(t) - P_0) \right].
\end{aligned} \tag{5.23}$$

To solve these equations, we use an implementation and associated parameter values described in a recent assessment of FIRE [192] using a two-step velocity-Verlet algorithm [75]. We terminate the enthalpy-minimization process when $\vec{\Phi}$ reaches a sufficiently small magnitude, which corresponds to both small magnitudes for all net forces between degrees of freedom, as well as a small difference between $P(t)$ and P_0 . Throughout our simulations, we set the mass of each vertex and the domain boundary mass to be equal to 1.

Chapter 6

Conclusions

6.1 Thesis summary

In this thesis, I have outlined several computational models that have been inspired by the physics of glassy and jammed materials to systems of biological relevance. In Chapters 2 and 3, I outline how jammed packings of amino acid-shaped particles can aid our understanding of core packing of well-folded protein molecules. I showed that the cores of protein structures resolved by X-ray crystallography (xtal) pack almost identically, in both bulk packing fraction and distribution of void space and local particle volumes, to packings of amino acid-shaped particles with purely rigid shapes and purely repulsive interaction potentials. Proteins resolved by nuclear magnetic resonance (NMR) imaging, however, were shown to pack slightly more densely than crystal structures, which was rationalized by the presence a difference in preparation protocol with greater thermalization. Given that NMR structures are resolved in dilute aqueous environments, whereas xtal structures are resolved in densely-packed crystals, the fact that thermalized packing protocols better match NMR structures may suggest that the crystallization process “traps” protein cores in less well-packed configurations. NMR structures, on the other hand, may allow for more thermalization during folding, allowing more well-packed protein cores.

In Chapter 4, I introduced a different computational model, now of deformable polygons that can change shape in response to applied stress. This model provides a natural starting point for simulating and studying a host of materials composed of deformable constituents like foams, emulsions, and, perhaps most importantly in the context of this thesis, cells. However, before describing possible applications to biological systems, I addressed a purely physical question; how can deformable objects collectively create rigid, solid-like materials? That is, a packing of infinitely-deformable objects presumably never forms solids that can resist applied load, as any external driving will result in particles sliding past each other as if in a fluid. Therefore, deformable particles must form solids when they particles are driven to their *limit* of deformability, but to what extent? And does jamming under this scenario differ from more familiar situations like, for example, jamming of hard granular materials?

To address this question, I introduced the deformable polygon (DP) model and first addressed

rigidity of individual particles using constraint counting and the dynamical matrix. The simplest DP model was found to produce particles with an abundance of floppy modes, and in jammed packings were found to rigidify through the use of quartic modes. Though quartic modes have been observed in other jammed packings of non-spherical particles, we find distinct differences in jammed packings of truly deformable particles. In particular, soft modes with characteristic vibrational frequency ω_0 scale as $\sim \mathcal{A}^{-1/3}$ with particle asphericity, whereas only $\sim \mathcal{A}^{1/2}$ has been previously observed in packings of other non-spherical particles. In addition, I showed that the static shear modulus $G = -d^2U/d\gamma^2$ had an anomalous scaling with pressure, satisfying $G \sim P^{3/4}$ in the high-pressure limit. As spherical particles have a scaling exponent $\sim P^{1/2}$, and other non-spherical particles typically have a scaling exponent of $G \sim P$, the vibrational and mechanical response of jammed packings of deformable particles indicate that these jammed materials may be members of new universality class.

This thesis concludes with the discussion of two different potential applications of the DP model to biological systems. At the end of Chapter 4, I demonstrate that actively crawling deformable polygons can display complex glassy behavior that can vary significantly depending on the mechanical properties of individual cells. In particular, I showed that the addition of bending energy along the cell's exterior can lead to solid-like arrested states even though cell shapes are highly non-spherical and similar to those found typically in flowing states. This result may help explain recent, puzzling observations that some epithelial tissues resist flow even with highly-deformed cells [133]. In Chapter 4, I then described the spongy mesophyll tissue in leaves and flowers, a biological system where packing constraints drive both development *and* diversity of the tissue. I first described mechanisms to quantitatively describe the microstructure of these complex tissues, in particular how to quantify mesophyll anisotropy across a wide variety of microCT scans of flower petals. Preliminary computational modelling suggests that anisotropy is driven by boundary-driven stress, but I showed that it is difficult to coax simple models to adopt the complex microstructures of real spongy mesophylls. I then motivated the need for more complex models that may be able to capture this microstructural complexity in future work, and described a simple test case for how deformable polygons may provide the foundation for such models. In particular, this model focuses specifically on how local growth and shape change coupled to a constant-stress boundary condition can drive the development of porous networks of highly non-spherical cells. While this model requires additional work to allow for more comprehensive study of mesophyll morphogenesis, it provides additional insight into how cell packing in early developmental stages helps pattern the ultimate microstructure in the mature tissue.

Throughout this thesis, a primary concern has been how physical constraints of packing and jamming affect biological systems. However, the biology in question have been purely structural, e.g. the packing fraction of protein cores, or the typical cell shape during mesophyll development. Future work can and should broaden the focus of these studies to not just structural, but *functional* aspects of biological systems. That is, future studies should ask, how do physical constraints inform the important biological functions that regulate living systems, lead to disease, and ultimately drive diversity and evolution of species? These questions yet lay beyond the reach of biological physics, but in the next section I lay out several directions that begin to address these questions.

6.2 Future Directions

As mentioned in Chapter 5, future work must continue to adapt the deformable polygon model for the simulation of spongy mesophyll development. This work can progress in several ways. First, larger systems deformable polygons in either (a) periodic boundary conditions or (b) deformable boundaries with enthalpy minimization in order to simulate physiologically-relevant system sizes and boundaries. In either of these boundary conditions, an important study will be to understand the parameter space of the model. While the coarse-grained model described in this thesis has considerably fewer adjustable parameters than the actual biological system, the various mechanisms of cell growth and adhesion provide a substantive parameter space. Future studies should seek to understand how parameters might be combined to reduce the size of the space in order to compute a minimal "phase diagram" that incorporates the most structural variability with the fewest number of parameters. By reducing parameters and fully exploring space, we will learn what physical mechanisms most sensitively affect development. In addition, three-dimensional computational modelling suggests that development in flower petals is driven in part by anisotropic boundary-driven stress. We will test the origins of this anisotropy by imposing anisotropic pressure on network-forming deformable polygons. For example, if the x -direction is easier to expand (with a lower imposed stress P_x) than the y -direction (with a larger imposed stress $P_y > P_x$), is this sufficient to generate anisotropic cell networks? Addressing this question is an important future direction, which I will do in both two- and three-dimensional [186] systems of deformable particles.

At the end of Chapter 4, I demonstrated that the dynamical behavior of densely-packed deformable polygons depends heavily on the mechanical properties of individual particles. Understanding how single-cell mechanics impact tissue solidification and fluidization is vital for our understanding of a host of biological functions such as morphogenesis, homeostasis or tumor invasion [132, 131, 134]. Therefore, another important future direction of this work is to further develop the DP model for the study of flowing and developing tissues. In particular, the addition of biologically-relevant adhesion mechanisms are of the utmost importance. While adhesion was incorporated into the simple DP model in Chapter 5, adhesion (particularly in developing epithelia) is inherently a non-equilibrium process [88, 94]. Incorporating processes like active tension fluctuations, binding kinetics of cadherin molecules and cytoskeletal-cadherin coupling into the DP model will allow for more realistic models of multicellular systems. However, the addition of biological detail comes at the cost of specificity and an increased number of parameters. Therefore, future work must strive to develop *minimal* models that incorporate these effects. Recent work on active foam-like particles [94] provides an excellent starting point for minimal models with biological relevance. Future work will then address how rigidity and adhesion at the single-cell level can sculpt and steer the *function* of tissues.

These future directions will investigate how physical constraints, such as dense cell packing, impact the function of specific biological systems, such as developing multicellular systems in animals and plants. However, a key future direction for biological physics in general is to seek out universal physical principles that form the foundation of many, if not all, biological phenomena. That is, rather than looking at a particular biological phenomenon and asking what physical features are at play, can we begin to derive and predict biological behavior purely from physical principles? It

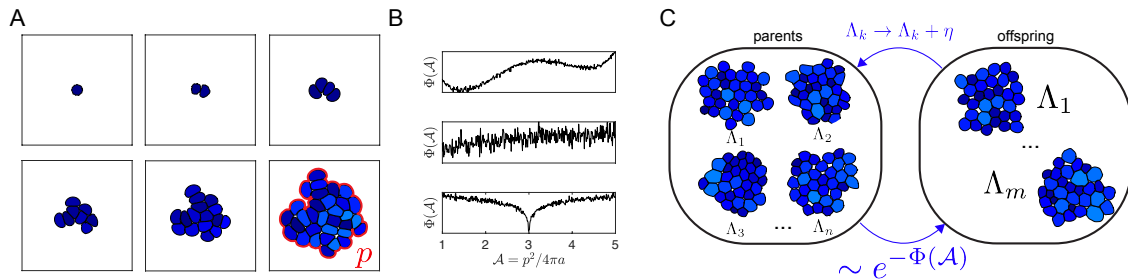


Figure 6.1: (A) The growth of a monolayer proceeds from the division of a single cell (color shows relative cell size). When the monolayer reaches a certain population, growth slows and the monolayer interface (with perimeter p , shown in red) becomes a trait under selection. (B) Schematic fitness function Φ of the monolayer’s surface-to-area ratio $\mathcal{A} = p^2/4\pi a$ for a monolayer of total area a . Possible landscape classes include smooth (top), rough (middle) and funnelled (bottom). (C) Evolution through a walk across the fitness landscape. A “parent” ensemble of n cell monolayers is generated, with each monolayer k obtained via a certain set of biophysical parameters Λ_k . “Offspring” ($m \leq n$) are selected with a probability $\sim \exp(-\Phi)$ and are given genetic noise η each generation.

seems a daunting task, but coarse-grained computational models might aid in the establishment of the “Rules of Life” that are present at any length scale, at any time, and in any region in the phylogenetic Tree of Life.

For example, in Fig. 6.1, a schematic evolutionary simulation is described that can model dynamics of cellular traits on a fitness landscape in a model multicellular system. Consider the developmental growth of a two-dimensional, multicellular “organism” composed of deformable polygons governed by a set of coarse-grained biophysical parameters Λ . These parameters, e.g. adhesion strength, division rates, etc., combine to influence the overall monolayer shape parameter $\mathcal{A} = p^2/4\pi a$ for a monolayer with perimeter p and area a . We could then study the coupling between evolutionary and developmental dynamics given the fitness Φ granted by the “mature” \mathcal{A} of a particular configuration. For example, given a particular Φ (Fig. 6.1B), one could track dynamics in Λ -space and monitor the emergence of stable points and limit cycles (niches and tradeoffs [177]), or diffusion (neutral speciation [195]).

One key benefit of this approach is that the coarse-grained parameters Λ may be drawn randomly from any initial distribution, but will naturally evolve toward steady states according to the selected fitness landscape. Therefore, this approach facilitates the study of purely *evolutionary* parameters (i.e. Φ and η); that is, evolution itself. This work can lend insight to fundamental problems to the dynamics of evolution, such as if stable phenotypes can emerge without strong pressure from the environment. For example, smooth circular interfaces are natural in systems with high surface tension, and if the fitness landscape Φ does not have clear optima (such as the “smooth” and “funnelled” cases in Fig. 6.1B), large genetic drift η might naturally create more circular phenotypes ($\mathcal{A} \rightarrow 1$) if a larger number of possible configurations simply possess a high interfacial tension. In this case, a purely physical principle, that interfacial tension drives bulk morphology, could drive

evolutionary forces absent of clear selective pressure. Although this “neutral” form of evolution has been suggested as an explanation for several observed evolutionary trends [196], the computational approach described here can systematically study and control how this evolutionary dynamic can lead to the emergence of complexity and diversity.

Such a simulation scheme would aid our understanding of the coupling between evolutionary dynamics and physical constraints, as well as how the evolution of subcellular components impact multicellular behavior. A key challenge to studying evolutionary dynamics is precisely the nearly impossible separation of scales at play; evolution occurs at the level of genes and proteins, yet we observe its effects over generations at the scale of organisms and species. The approach outlined above therefore allows the interrogation of evolutionary dynamics at biology’s mesoscale, i.e. cells and tissues, which might help bridge the gap between molecules and populations. However, the proposed research direction leaves several key questions unanswered *prima facie*. For example, what sets the fitness landscape? Are there ways of generating Φ , or methodologies systematically studying all possible Φ ? Also, how will this approach be experimentally validated? Although it would be possible to validate the effect of a local set of parameters on global tissue morphology (i.e. of a given Λ on the monolayer \mathcal{A}), measurement of evolution or even the properties of the fitness landscape would be considerably more difficult. Nevertheless, I believe that coarse-grained computational models inspired by biological systems are uniquely suited to bridge the gap between biological phenomena and physical principles. Only with concerted combination of computational modelling, theoretical exploration and experimental validation can we find what unites the disparate arms of biology and begin to understand Life as a purely physical phenomenon.

Bibliography

- [1] Tamás Vicsek, András Czirók, Eshel Ben-Jacob, Inon Cohen, and Ofer Shochet. Novel type of phase transition in a system of self-driven particles. *Phys. Rev. Lett.*, 75:1226–1229, 1995.
- [2] Christopher Dombrowski, Luis Cisneros, Sunita Chatkaew, Raymond E. Goldstein, and John O. Kessler. Self-concentration and large-scale coherence in bacterial dynamics. *Phys. Rev. Lett.*, 93, 2004.
- [3] Guannan Liu, Adam Patch, Fatmagül Bahar, David Yllanes, Roy D. Welch, M. Cristina Marchetti, Shashi Thutupalli, and Joshua W. Shaevitz. Self-driven phase transitions drive myxococcus xanthus fruiting body formation. *Phys. Rev. Lett.*, 122:248102, 2019.
- [4] Dipjyoti Das, Veena Chatti, Thierry Emonet, and Scott A. Holley. Patterned disordered cell motion ensures vertebral column symmetry. *Developmental Cell*, 42(2):170 – 180.e5, 2017.
- [5] Andrea Cavagna, Alessio Cimorelli, Irene Giardina, Giorgio Parisi, Raffaele Santagati, Fabio Stefanini, and Massimiliano Viale. Scale-free correlations in starling flocks. *Proceedings of the National Academy of Sciences*, 107(26):11865–11870, 2010.
- [6] Thierry Mora and William Bialek. Are biological systems poised at criticality? *Journal of Statistical Physics*, 144(2):268–302, 2011.
- [7] John D. Treado, Zhe Mei, Lynne Regan, and Corey S. O’Hern. Void distributions reveal structural link between jammed packings and protein cores. *Phys. Rev. E*, 99:022416, 2019.
- [8] Zhe Mei, John D. Treado, Alex T. Grigas, Zachary A. Levine, Lynne Regan, and Corey S. O’Hern. Analyses of protein cores reveal fundamental differences between solution and crystal structures. *Proteins*, 88(9):1154–1161, 2020.
- [9] John D. Treado, Dong Wang, Arman Boromand, Michael P. Murrell, Mark D. Shattuck, and Corey S. O’Hern. Bridging particle deformability and collective response in soft solids. *Phys. Rev. Materials*, 5:055605, 2021.
- [10] Ken A. Dill. Dominant forces in protein folding. *Biochemistry*, 29(31):7133–7155, 1990.
- [11] Jie Liang and Ken A. Dill. Are proteins well-packed? *Biophysical Journal*, 81(2):751 – 766, 2001.
- [12] Julien Roche, Jose A. Caro, Douglas R. Norberto, Philippe Barthe, Christian Roumestand, Jamie L. Schlessman, Angel E. Garcia, Bertrand García-Moreno E., and Catherine A. Royer.

- Cavities determine the pressure unfolding of proteins. *Proceedings of the National Academy of Sciences*, 109(18):6945–6950, 2012.
- [13] Nathaniel V. Nucci, Brian Fuglestad, Evangelia A. Athanasoula, and A. Joshua Wand. Role of cavities and hydration in the pressure unfolding of t4 lysozyme. *Proceedings of the National Academy of Sciences*, 111(38):13846–13851, 2014.
 - [14] Michael T. Lerch, Carlos J. López, Zhongyu Yang, Margaux J. Kreitman, Joseph Horwitz, and Wayne L. Hubbell. Structure-relaxation mechanism for the response of t4 lysozyme cavity mutants to hydrostatic pressure. *Proceedings of the National Academy of Sciences*, 112(19):E2437–E2446, 2015.
 - [15] Benjamin Borgo and James J. Havranek. Automated selection of stabilizing mutations in designed and natural proteins. *Proceedings of the National Academy of Sciences*, 109(5):1494–1499, 2012.
 - [16] A. E. Eriksson, W. A. Baase, X.-J. Zhang, D. W. Heinz, M. Blaber, E. P. Baldwin, and B. W. Matthews. Response of a protein structure to cavity-creating mutations and its relation to the hydrophobic effect. *Science*, 255(5041):178–183, 1992.
 - [17] Will Sheffler and David Baker. Rosettaholes: Rapid assessment of protein core packing for structure prediction, refinement, design, and validation. *Protein Science*, 18(1):229–239, 2008.
 - [18] Sarel J. Fleishman and *et. al.* Community-wide assessment of protein-interface modeling suggests improvements to design methodology. *Journal of Molecular Biology*, 414(2):289 – 302, 2011.
 - [19] D. Caballero, A. Virrueta, C.S. O’Hern, and L. Regan. Steric interactions determine side-chain conformations in protein cores. *Protein Engineering, Design and Selection*, 29(9):367–376, 2016.
 - [20] Jennifer C. Gaines, W. Wendell Smith, Lynne Regan, and Corey S. O’Hern. Random close packing in protein cores. *Phys. Rev. E*, 93:032415, 2016.
 - [21] J.C. Gaines, A. Virrueta, D.A. Buch, S.J. Fleishman, C.S. O’Hern, and L. Regan. Collective repacking reveals that the structures of protein cores are uniquely specified by steric repulsive interactions. *Protein Engineering, Design and Selection*, 30(5):387–394, 2017.
 - [22] Kyle VanderWerf, Weiwei Jin, Mark D. Shattuck, and Corey S. O’Hern. Hypostatic jammed packings of frictionless nonspherical particles. *Phys. Rev. E*, 97:012909, 2018.
 - [23] I. E. T. Iben, D. Braunstein, W. Doster, H. Frauenfelder, M. K. Hong, J. B. Johnson, S. Luck, P. Ormos, A. Schulte, P. J. Steinbach, A. H. Xie, and R. D. Young. Glassy behavior of a protein. *Phys. Rev. Lett.*, 62:1916–1919, 1989.
 - [24] D. L. Stein. A model of protein conformational substates. *Proceedings of the National Academy of Sciences*, 82(11):3670–3672, 1985.
 - [25] J D Bryngelson and P G Wolynes. Spin glasses and the statistical mechanics of protein folding. *Proceedings of the National Academy of Sciences*, 84(21):7524–7528, 1987.

- [26] Richard J. Loncharich and Bernard R. Brooks. Temperature dependence of dynamics of hydrated myoglobin: Comparison of force field calculations with neutron scattering data. *Journal of Molecular Biology*, 215(3):439 – 455, 1990.
- [27] Dagmar Ringe and Gregory A. Petsko. The ‘glass transition’ in protein dynamics: what it is, why it occurs, and how to exploit it. *Biophysical Chemistry*, 105(2):667 – 680, 2003.
- [28] Kertész, J. Percolation of holes between overlapping spheres : Monte carlo calculation of the critical volume fraction. *J. Physique Lett.*, 42(17):393–395, 1981.
- [29] Alison L. Cuff and Andrew C. R. Martin. Analysis of void volumes in proteins and application to stability of the p53 tumour suppressor protein. *Journal of Molecular Biology*, 344(5):1199–1209, 2004.
- [30] Guoli Wang and Roland L. Dunbrack, Jr. PISCES: A protein sequence culling server. *Bioinformatics*, 19(12):1589–1591, 2003.
- [31] Guoli Wang and Roland L. Dunbrack, Jr. PISCES: Recent improvements to a PDB sequence culling server. *Nucleic Acids Research*, 33:W94–8, 2005.
- [32] J. Michael Word, Simon C. Lovell, Jane S. Richardson, and David C. Richardson. Asparagine and glutamine: Using hydrogen atom contacts in the choice of side-chain amide orientation. *Journal of Molecular Biology*, 285(4):1735 – 1747, 1999.
- [33] S. J. Hubbard and J. M. Thornton. Naccess, 1993.
- [34] J. C. Gaines, S. Acebes, A. Virrueta, M. Butler, L. Regan, and C. S. O’Hern. Comparing side chain packing in soluble proteins, protein-protein interfaces and transmembrane proteins. *Proteins: Structure, Function, and Bioinformatics*, 86(5):581–591, 2018.
- [35] Corey S. O’Hern, Leonardo E. Silbert, Andrea J. Liu, and Sidney R. Nagel. Jamming at zero temperature and zero applied stress: The epitome of disorder. *Phys. Rev. E*, 68:011306, 2003.
- [36] J C Gaines, A H Clark, L Regan, and C S O’Hern. Packing in protein cores. *Journal of Physics: Condensed Matter*, 29(29):293001, 2017.
- [37] Erik Bitzek, Pekka Koskinen, Franz Gähler, Michael Moseler, and Peter Gumbsch. Structural relaxation made simple. *Phys. Rev. Lett.*, 97:170201, 2006.
- [38] Dmitri Rozmanov and Peter G. Kusalik. Robust rotational-velocity-Verlet integration methods. *Physical Review E*, 81(5):056706, 2010.
- [39] Fabian M. Schaller, Sebastian Kapfer, Myfanwy Evans, Matthias J. F. Hoffmann, Tomaso Aste, Mohammad Saadatfar, Klaus Mecke, Gary W. Delaney, and Gerd Schröder-Turk. Set Voronoi diagrams of 3D assemblies of aspherical particles. *Philosophical Magazine*, 93:3993–4017, 2013.
- [40] Simon Weis, Philipp W. A. Schönhöfer, Fabian M. Schaller, Matthias Schröter, and Gerd E. Schröder-Turk. Pomelo, A tool for computing generic set Voronoi diagrams of aspherical particles of arbitrary shape. *EPJ Web Conf.*, 140:06007, 2017.

- [41] M. E. J. Newman and R. M. Ziff. Fast monte carlo algorithm for site or bond percolation. *Phys. Rev. E*, 64:016706, 2001.
- [42] T. Aste and T. Di Matteo. Emergence of gamma distributions in granular materials and packing models. *Phys. Rev. E*, 77:021309, 2008.
- [43] T. Aste, T. Di Matteo, M. Saadatfar, T. J. Senden, Matthias Schröter, and Harry L. Swinney. An invariant distribution in static granular media. *Europhysics Letters*, 79(2):24003, 2007.
- [44] Fabian M. Schaller, Robert F. B. Weigel, and Sebastian C. Kapfer. Densest local structures of uniaxial ellipsoids. *Phys. Rev. X*, 6:041032, 2016.
- [45] Ivane Jorjadze, Lea-Laetitia Pontani, Katherine A. Newhall, and Jasna Brujić. Attractive emulsion droplets probe the phase diagram of jammed granular matter. *Proceedings of the National Academy of Sciences*, 108(11):4286–4291, 2011.
- [46] Jindong Li, Yixin Cao, Chengjie Xia, Binqun Kou, Xianghui Xiao, Kamel Fezzaa, and Yujie Wang. Similarity of wet granular packing to gels. *Nature Communications*, 5:5014, 2014.
- [47] Arman Boromand, Alexandra Signoriello, Fangfu Ye, Corey S. O’Hern, and Mark D. Shattuck. Jamming of deformable polygons. *Phys. Rev. Lett.*, 121:248003, 2018.
- [48] Eloi Pineda, Pere Bruna, and Daniel Crespo. Cell size distribution in random tessellations of space. *Phys. Rev. E*, 70:066119, 2004.
- [49] A R Kerstein. Equivalence of the void percolation problem for overlapping spheres and a network problem. *Journal of Physics A: Mathematical and General*, 16(13):3071, 1983.
- [50] M. D. Rintoul. Precise determination of the void percolation threshold for two distributions of overlapping spheres. *Phys. Rev. E*, 62:68–72, 2000.
- [51] M D Rintoul and S Torquato. Precise determination of the critical threshold and exponents in a three-dimensional continuum percolation model. *Journal of Physics A: Mathematical and General*, 30(16):L585, 1997.
- [52] Srdjan Ostojic, Ellák Somfai, and Bernard Nienhuis. Scale invariance and universality of force networks in static granular matter. *Nature*, 439:828, 2006.
- [53] Tianqi Shen, Corey S. O’Hern, and M. D. Shattuck. Contact percolation transition in athermal particulate systems. *Phys. Rev. E*, 85:011308, 2012.
- [54] Gregg Lois, Jerzy Blawdziewicz, and Corey S. O’Hern. Jamming transition and new percolation universality classes in particulate systems with attraction. *Phys. Rev. Lett.*, 100:028001, 2008.
- [55] Y. B. Yi. Void percolation and conduction of overlapping ellipsoids. *Phys. Rev. E*, 74:031112, 2006.
- [56] Dietrich Stauffer and Amnon Aharony. *Introduction to Percolation Theory*. CRC Press, 1994.
- [57] Geoffrey Grimmet. *Percolation*, volume 321 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag Berlin Heidelberg, 2 edition, 1999.

- [58] Richard J. Loncharich and Bernard R. Brooks. Temperature dependence of dynamics of hydrated myoglobin: Comparison of force field calculations with neutron scattering data. *Journal of Molecular Biology*, 215(3):439 – 455, 1990.
- [59] John Moult, Krzysztof Fidelis, Andriy Kryshchak, Torsten Schwede, and Anna Tramontano. Critical assessment of methods of protein structure prediction (casp)—round xii. *Proteins: Structure, Function, and Bioinformatics*, 86(S1):7–15, 2018.
- [60] Ladislav Hovan, Vladimiras Oleinikovas, Havva Yalinca, Andriy Kryshchak, Giorgio Saladino, and Francesco Luigi Gervasio. Assessment of the model refinement category in casp12. *Proteins: Structure, Function, and Bioinformatics*, 86(S1):152–167, 2018.
- [61] Sergiy O. Garbuzynskiy, Bogdan S. Melnik, Michail Yu. Lobanov, Alexei V. Finkelstein, and Oxana V. Galzitskaya. Comparison of x-ray and nmr structures: Is there a systematic difference in residue contacts between x-ray- and nmr-resolved protein structures? *Proteins: Structure, Function, and Bioinformatics*, 60(1):139–147, 2005.
- [62] Michael Schneider, Xiaoran Fu, and Amy E. Keating. X-ray vs. nmr structures as templates for computational protein design. *Proteins: Structure, Function, and Bioinformatics*, 77(1):97–110, 2009.
- [63] Binchen Mao, Roberto Tejero, David Baker, and Gaetano T. Montelione. Protein nmr structures refined with rosetta have higher accuracy relative to corresponding x-ray crystal structures. *Journal of the American Chemical Society*, 136(5):1893–1906, 2014.
- [64] Julia Koehler Leman, Andrew R. D’Avino, Yash Bhatnagar, and Jeffrey J. Gray. Comparison of nmr and crystal structures of membrane proteins and computational refinement to improve model quality. *Proteins: Structure, Function, and Bioinformatics*, 86(1):57–74, 2018.
- [65] Robert B. Best, Kresten Lindorff-Larsen, Mark A. DePristo, and Michele Vendruscolo. Relation between native ensembles and experimental structures of proteins. *Proceedings of the National Academy of Sciences*, 103(29):10901–10906, 2006.
- [66] Lee-Wei Yang, Eran Eyal, Chakra Chennubhotla, JunGoo Jee, Angela M. Gronenborn, and Ivet Bahar. Insights into equilibrium dynamics of proteins from comparison of nmr and x-ray data with computational predictions. *Structure*, 15(6):741 – 749, 2007.
- [67] Kresimir Sikic, Sanja Tomic, and Oliviero Carugo. Systematic comparison of crystal and nmr protein structures deposited in the protein data bank. *The Open Biochemistry Journal*, 4(83-95):83–95, 2010.
- [68] John K. Everett, Roberto Tejero, Sarath B. K. Murthy, Thomas B. Acton, James M. Aramini, Michael C. Baran, Jordi Benach, John R. Cort, Alexander Eletsky, Farhad Forouhar, Rongjin Guan, Alexandre P. Kuzin, Hsiau-Wei Lee, Gaohua Liu, Rajeswari Mani, Binchen Mao, Jeffrey L. Mills, Alexander F. Montelione, Kari Pederson, Robert Powers, Theresa Ramelot, Paolo Rossi, Jayaraman Seetharaman, David Snyder, G. V. T. Swapna, Sergey M. Vorobiev, Yibing Wu, Rong Xiao, Yunhuang Yang, Cheryl H. Arrowsmith, John F. Hunt, Michael A. Kennedy, James H. Prestegard, Thomas Szyperski, Liang Tong, and Gaetano T. Montelione. A community resource of experimental data for nmr / x-ray crystal structure pairs. *Protein Science*, 25(1):30–45, 2016.

- [69] B. Lee and F.M. Richards. The interpretation of protein structures: Estimation of static accessibility. *Journal of Molecular Biology*, 55(3):379–400, 1971.
- [70] D. M. Endres and J. E. Schindelin. A new metric for probability distributions. *IEEE Transactions on Information Theory*, 49(7):1858–1860, 2003.
- [71] Bertil Halle. Biomolecular cryocrystallography: Structural changes during flash-cooling. *Proceedings of the National Academy of Sciences*, 101(14):4793–4798, 2004.
- [72] James S. Fraser, Henry van den Bedem, Avi J. Samelson, P. Therese Lang, James M. Holton, Nathaniel Echols, and Tom Alber. Accessing protein conformational ensembles using room-temperature x-ray crystallography. *Proceedings of the National Academy of Sciences*, 108(39):16247–16252, 2011.
- [73] Alice Qinhua Zhou, Corey S. O’Hern, and Lynne Regan. Predicting the side-chain dihedral angle distributions of nonpolar, aromatic, and polar amino acids using hard sphere models. *Proteins: Structure, Function, and Bioinformatics*, 82(10):2574–2584, 2014.
- [74] Alexander Miguel Monzon, Cristian Oscar Rohr, María Silvina Fornasari, and Gustavo Parisi. CoDNaS 2.0: a comprehensive database of protein conformational diversity in the native state. *Database*, 2016, 2016.
- [75] Michael P. Allen and Dominic J. Tildesley. *Computer Simulation of Liquids*. Oxford University Press, 2nd edition, 2017.
- [76] Ning Xu, Jerzy Blawdziewicz, and Corey S. O’Hern. Random close packing revisited: Ways to pack frictionless disks. *Phys. Rev. E*, 71:061306, 2005.
- [77] D. J. Durian. Foam mechanics at the bubble scale. *Phys. Rev. Lett.*, 75:4780–4783, 1995.
- [78] M van Hecke. Jamming of soft particles: geometry, mechanics, scaling and isostaticity. *J. Phys.: Condens. Matter*, 22(3):033101, 2009.
- [79] F. Bolton and D. Weaire. Rigidity loss transition in a disordered 2d froth. *Phys. Rev. Lett.*, 65:3449–3451, 1990.
- [80] Yann Bertho, Christophe Becco, and Nicolas Vandewalle. Dense bubble flow in a silo: An unusual flow of a dispersed medium. *Phys. Rev. E*, 73:056309, 2006.
- [81] H.M Princen. Rheology of foams and highly concentrated emulsions: I. elastic properties and yield stress of a cylindrical model system. *J. Colloid. Interf Sci.*, 91(1):160 – 175, 1983.
- [82] Arman Boromand, Alexandra Signoriello, Janna Lowensohn, Carlos S. Orellana, Eric R. Weeks, Fangfu Ye, Mark D. Shattuck, and Corey S. O’Hern. The role of deformability in determining the structural and mechanical properties of bubbles and emulsions. *Soft Matter*, 15:5854–5865, 2019.
- [83] William P. J. Smith, Yohan Davit, James M. Osborne, Wook Kim, Kevin R. Foster, and Joe M. Pitt-Francis. Cell morphology drives spatial patterning in microbial communities. *Proc. Natl. Acad. Sci. USA*, 114(3):E280–E286, 2017.

- [84] Farzan Beroz, Jing Yan, Yigal Meir, Benedikt Sabass, Howard A. Stone, Bonnie L. Bassler, and Ned S. Wingreen. Verticalization of bacterial biofilms. *Nat. Phys.*, 14(9):954–960, 2018.
- [85] Jin-Ah Park, Jae Hun Kim, Dapeng Bi, Jennifer A. Mitchel, Nader Taheri Qazvini, Kelan Tantisira, Chan Young Park, Maureen McGill, Sae-Hoon Kim, Bomi Gweon, Jacob Notbohm, Robert Steward Jr, Stephanie Burger, Scott H. Randell, Alvin T. Kho, Dhananjay T. Tambe, Corey Hardin, Stephanie A. Shore, Elliot Israel, David A. Weitz, Daniel J. Tschumperlin, Elizabeth P. Henske, Scott T. Weiss, M. Lisa Manning, James P. Butler, Jeffrey M. Drazen, and Jeffrey J. Fredberg. Unjamming and cell shape in the asthmatic airway epithelium. *Nat. Mater.*, 14(10):1040–1048, 2015.
- [86] Xavier Trepast and Erik Sahai. Mesoscale physical principles of collective cell organization. *Nat. Phys.*, 14(7):671–682, 2018.
- [87] Thomas Lecuit and Pierre-François Lenne. Cell surface mechanics and the control of cell shape, tissue patterns and morphogenesis. *Nature Reviews Molecular Cell Biology*, 8(8):633–644, 2007.
- [88] Michael Murrell, Patrick W. Oakes, Martin Lenz, and Margaret L. Gardel. Forcing cells into shape: the mechanics of actomyosin contractility. *Nat. Rev. Mol. Cell Bio.*, 16(8):486–498, 2015.
- [89] Mohit Kumar Jolly, Marcelo Boareto, Bin Huang, Dongya Jia, Mingyang Lu, Eshel Ben-Jacob, José N. Onuchic, and Herbert Levine. Implications of the hybrid epithelial/mesenchymal phenotype in metastasis. *Front. Oncol.*, 5:155, 2015.
- [90] Patrick McMillen, Veena Chatti, Dörthe Jülich, and Scott A. Holley. A sawtooth pattern of cadherin 2 stability mechanically regulates somite morphogenesis. *Curr. Biol.*, 26(4):542 – 549, 2016.
- [91] Linda Oswald, Steffen Grosser, David M Smith, and Josef A Käs. Jamming transitions in cancer. *J. Phys. D Appl. Phys.*, 50(48):483001, 2017.
- [92] Alessandro Mongera, Payam Rowghanian, Hannah J. Gustafson, Elijah Shelton, David A. Kealhofer, Emmet K. Carn, Friedhelm Serwane, Adam A. Lucio, James Giammona, and Otger Campàs. A fluid-to-solid jamming transition underlies vertebrate body axis elongation. *Nature*, 561(7723):401–405, 2018.
- [93] Olga Ilina, Pavlo G. Gritsenko, Simon Syga, Jürgen Lippoldt, Caterina A. M. La Porta, Oleksandr Chepizhko, Steffen Grosser, Manon Vullings, Gert-Jan Bakker, Jörn Starrau, Peter Bult, Stefano Zapperi, Josef A. Käs, Andreas Deutsch, and Peter Friedl. Cell-cell adhesion and 3d matrix confinement determine jamming transitions in breast cancer invasion. *Nat. Cell Biol.*, 22(9):1103–1115, 2020.
- [94] Sangwoo Kim, Marie Pochitaloff, Georgina A. Stooke-Vaughan, and Otger Campàs. Embryonic tissues as active foams. *Nature Physics*, 17(7):859–866, 2021.
- [95] Nathalie Wuyts, Jean-Christophe Palauqui, Geneviève Conejero, Jean-Luc Verdeil, Christine Granier, and Catherine Massonnet. High-contrast three-dimensional imaging of the arabidopsis leaf enables the analysis of cell dimensions in the epidermis and mesophyll. *Plant Methods*, 6(1):17, 2010.

- [96] Shweta Kalve, Joanna Fotschki, Tom Beeckman, Kris Vissenberg, and Gerrit T. S. Beemster. Three-dimensional patterns of cell division and expansion throughout the development of *Arabidopsis thaliana* leaves. *J. Exp. Bot.*, 65(22):6385–6397, 2014.
- [97] Aleksandra Sapala, Adam Runions, Anne-Lise Routier-Kierzkowska, Mainak Das Gupta, Lilan Hong, Hugo Hofhuis, Stéphane Verger, Gabriella Mosca, Chun-Biu Li, Angela Hay, Olivier Hamant, Adrienne HK Roeder, Miltos Tsiantis, Przemyslaw Prusinkiewicz, and Richard S Smith. Why plants make puzzle cells, and how their shape emerges. *eLife*, 7: e32794, 2018.
- [98] Pablo Martinez, Lindy A. Allsman, Kenneth A. Brakke, Christopher Hoyt, Jordan Hayes, Hong Liang, Wesley Neher, Yue Rui, Allyson M. Roberts, Amir Moradifam, Bob Goldstein, Charles T. Anderson, and Carolyn G. Rasmussen. Predicting division planes of three-dimensional cells by soap-film minimization. *Plant Cell*, 30(10):2255–2266, 2018.
- [99] Aleca M. Borsuk, Adam B. Roddy, Guillaume Thérout-Rancourt, and Craig R. Brodersen. Structural organization of the spongy mesophyll. *New Phytologist*, 2022.
- [100] Alexei V. Tkachenko and Thomas A. Witten. Stress propagation through frictionless granular material. *Phys. Rev. E*, 60:687–696, 1999.
- [101] S. Pellegrino and C.R. Calladine. Matrix analysis of statically and kinematically indeterminate frameworks. *International Journal of Solids and Structures*, 22(4):409–428, 1986.
- [102] Carl F. Schreck, Thibault Bertrand, Corey S. O’Hern, and M. D. Shattuck. Repulsive contact interactions make jammed particulate systems inherently nonharmonic. *Phys. Rev. Lett.*, 107: 078301, 2011.
- [103] Carl P. Goodrich, Andrea J. Liu, and Sidney R. Nagel. Solids between the mechanical extremes of order and disorder. *Nat. Phys.*, 10(8):578–581, 2014.
- [104] Mitch Mailman, Carl F. Schreck, Corey S. O’Hern, and Bulbul Chakraborty. Jamming in systems composed of frictionless ellipse-shaped particles. *Phys. Rev. Lett.*, 102:255501, 2009.
- [105] Aleksandar Donev, Robert Connelly, Frank H. Stillinger, and Salvatore Torquato. Underconstrained jammed packings of nonspherical hard particles: Ellipses and ellipsoids. *Phys. Rev. E*, 75:051304, 2007.
- [106] Carl F. Schreck, Mitch Mailman, Bulbul Chakraborty, and Corey S. O’Hern. Constraints and vibrations in static packings of ellipsoidal particles. *Phys. Rev. E*, 85:061305, 2012.
- [107] Ye Yuan, Kyle VanderWerf, Mark D. Shattuck, and Corey S. O’Hern. Jammed packings of 3d superellipsoids with tunable packing fraction, coordination number, and ordering. *Soft Matter*, 15:9751–9761, 2019.
- [108] Carolina Brito, Harukuni Ikeda, Pierfrancesco Urbani, Matthieu Wyart, and Francesco Zamponi. Universality of jamming of nonspherical particles. *Proc. Natl. Acad. Sci. USA*, 115(46): 11736–11741, 2018.
- [109] Tianqi Shen, Carl Schreck, Bulbul Chakraborty, Denise E. Freed, and Corey S. O’Hern. Structural relaxation in dense liquids composed of anisotropic particles. *Phys. Rev. E*, 86: 041303, 2012.

- [110] Ojan Khatib Damavandi, Varda F. Hagh, Christian D. Santangelo, and M. Lisa Manning. Energetic rigidity. i. a unifying theory of mechanical stability. *Phys. Rev. E*, 105, 2022.
- [111] Ojan Khatib Damavandi, Varda F. Hagh, Christian D. Santangelo, and M. Lisa Manning. Energetic rigidity. ii. applications in examples of biological and underconstrained materials. *Phys. Rev. E*, 105, 2022.
- [112] Dapeng Bi, J. H. Lopez, J. M. Schwarz, and M. Lisa Manning. A density-independent rigidity transition in biological tissues. *Nature Physics*, 11:1074, 2015.
- [113] Le Yan and Dapeng Bi. Multicellular rosettes drive fluid-solid transition in epithelial tissues. *Phys. Rev. X*, 9:011029, 2019.
- [114] Kai Zhang, W. Wendell Smith, Minglei Wang, Yanhui Liu, Jan Schroers, Mark D. Shattuck, and Corey S. O’Hern. Connection between the packing efficiency of binary hard spheres and the glass-forming ability of bulk metallic glasses. *Phys. Rev. E*, 90:032311, 2014.
- [115] C. L. Kane and T. C. Lubensky. Topological boundary modes in isostatic lattices. *Nature Physics*, 10(1):39–45, 2014.
- [116] Bryan Gin-ge Chen, Nitin Upadhyaya, and Vincenzo Vitelli. Nonlinear conduction via solitons in a topological mechanical insulator. *Proceedings of the National Academy of Sciences*, 111(36):13004–13009, 2014.
- [117] Stefanos Papanikolaou, Corey S. O’Hern, and Mark D. Shattuck. Isostaticity at frictional jamming. *Phys. Rev. Lett.*, 110:198002, 2013.
- [118] Carl F. Schreck, Corey S. O’Hern, and Leonardo E. Silbert. Tuning jammed frictionless disk packings from isostatic to hyperstatic. *Phys. Rev. E*, 84:011305, 2011.
- [119] Philip J. Tuckman, Kyle VanderWerf, Ye Yuan, Shiyun Zhang, Jerry Zhang, Mark D. Shattuck, and Corey S. O’Hern. Contact network changes in ordered and disordered disk packings. *Soft Matter*, 16:9443–9455, 2020.
- [120] M. Foglino, A. N. Morozov, O. Henrich, and D. Marenduzzo. Flow of deformable droplets: Discontinuous shear thinning and velocity oscillations. *Phys. Rev. Lett.*, 119:208002, 2017.
- [121] Xia Hong, Meghan Kohne, Mia Morrell, Haoran Wang, and Eric R. Weeks. Clogging of soft particles in two-dimensional hoppers. *Phys. Rev. E*, 96:062605, 2017.
- [122] Iaroslava Golovkova, Lorraine Montel, Elie Wandersman, Thibault Bertrand, Alexis Michel Prevost, and Lea-Laetitia Pontani. Depletion attraction impairs the plasticity of emulsions flowing in a constriction. *Soft Matter*, 16:3294–3302, 2020.
- [123] Carl P. Goodrich, Andrea J. Liu, and Sidney R. Nagel. Finite-size scaling at the jamming transition. *Phys. Rev. Lett.*, 109:095704, 2012.
- [124] Kyle VanderWerf, Arman Boromand, Mark D. Shattuck, and Corey S. O’Hern. Pressure dependent shear response of jammed packings of frictionless spherical particles. *Phys. Rev. Lett.*, 124:038004, 2020.

- [125] Ellák Somfai, Martin van Hecke, Wouter G. Ellenbroek, Kostya Shundyak, and Wim van Saarloos. Critical and noncritical jamming of frictional grains. *Phys. Rev. E*, 75:020301(R), 2007.
- [126] Philip Wang, Shiyun Zhang, Philip Tuckman, Nicholas T. Ouellette, Mark D. Shattuck, and Corey S. O’Hern. Shear response of granular packings compressed above jamming onset. *Phys. Rev. E*, 103:022902, 2021.
- [127] Dapeng Bi, Xingbo Yang, M. Cristina Marchetti, and M. Lisa Manning. Motility-driven glass and jamming transitions in biological tissues. *Phys. Rev. X*, 6:021011, 2016.
- [128] Georgina A Stooke-Vaughan and Otger Campàs. Physical control of tissue morphogenesis across scales. *Current Opinion in Genetics & Development*, 51:111 – 119, 2018. Developmental mechanisms, patterning and evolution.
- [129] Visar Ajeti, A. Pasha Tabatabai, Andrew J. Fleszar, Michael F. Staddon, Daniel S. Seara, Cristian Suarez, M. Sulaiman Yousafzai, Dapeng Bi, David R. Kovar, Shiladitya Banerjee, and Michael P. Murrell. Wound healing coordinates actin architectures to regulate mechanical work. *Nat. Phys.*, 15(7):696–705, 2019.
- [130] Yaouen Fily and M. Cristina Marchetti. Athermal phase separation of self-propelled particles with no alignment. *Phys. Rev. Lett.*, 108:235702, 2012.
- [131] Xinzhi Li, Amit Das, and Dapeng Bi. Mechanical heterogeneity in tissues promotes rigidity and controls cellular invasion. *Phys. Rev. Lett.*, 123, 2019.
- [132] Lior Atia, Dapeng Bi, Yasha Sharma, Jennifer A. Mitchel, Bomi Gweon, Stephan A. Koehler, Stephen J. DeCamp, Bo Lan, Jae Hun Kim, Rebecca Hirsch, Adrian F. Pegoraro, Kyu Ha Lee, Jacqueline R. Starr, David A. Weitz, Adam C. Martin, Jin-Ah Park, James P. Butler, and Jeffrey J. Fredberg. Geometric constraints during epithelial jamming. *Nature Physics*, 14(6):613–620, 2018.
- [133] Xun Wang, Matthias Merkel, Leo B. Sutter, Gonca Erdemci-Tandogan, M. Lisa Manning, and Karen E. Kasza. Anisotropy links cell shapes to tissue flow during convergent extension. *Proc. Natl. Acad. Sci. USA*, 117(24):13541–13551, 2020.
- [134] Nabila Founounou, Reza Farhadifar, Giovanna M. Collu, Ursula Weber, Michael J. Shelley, and Marek Mlodzik. Tissue fluidity mediated by adherens junction dynamics promotes planar cell polarity-driven ommatidial rotation. *Nature Communications*, 12(1):6974, 2021.
- [135] Asaph Widmer-Cooper, Heidi Perry, Peter Harrowell, and David R. Reichman. Irreversible reorganization in a supercooled liquid originates from localized soft modes. *Nature Physics*, 4(9):711–715, 2008.
- [136] Shay M. Rappaport and Yitzhak Rabin. Bending affects entropy of semiflexible polymers: Application to protein-dna complexes. *Phys. Rev. E*, 80:052801, 2009.
- [137] Frank Berendse and Marten Scheffer. The angiosperm radiation revisited, an ecological explanation for darwin’s ‘abominable mystery’. *Ecology Letters*, 12(9):865–872, 2009.

- [138] William E. Friedman. The meaning of darwin’s “abominable mystery”. *American Journal of Botany*, 96(1):5–21, 2009.
- [139] Kevin A. Simonin and Adam B. Roddy. Genome downsizing, physiological novelty, and the global dominance of flowering plants. *PLOS Biology*, 16(1):1–15, 2018.
- [140] Charles C. Davis, Maribeth Latvis, Daniel L. Nickrent, Kenneth J. Wurdack, and David A. Baum. Floral gigantism in rafflesiaceae. *Science*, 315(5820):1812–1812, 2007.
- [141] Adam B. Roddy, Craig R. Brodersen, and Todd E. Dawson. Hydraulic conductance and the maintenance of water balance in flowers. *Plant, Cell & Environment*, 39(10):2123–2132, 2016.
- [142] Adam B. Roddy, Guo-Feng Jiang, Kunfang Cao, Kevin A. Simonin, and Craig R. Brodersen. Hydraulic traits are more diverse in flowers than in leaves. *New Phytologist*, 223(1):193–203, 2019.
- [143] Adam B. Roddy, Guillaume Th  roux-Rancourt, Tito Abbo, Joseph W. Benedetti, Craig R. Brodersen, Mariana Castro, Silvia Castro, Austin B. Gilbride, Brook Jensen, Guo-Feng Jiang, John A. Perkins, Sally D. Perkins, Jo  o Loureiro, Zuhah Syed, R. Alexander Thompson, Sara E. Kuebbing, and Kevin A. Simonin. The scaling of genome size and cell size limits maximum rates of photosynthesis with implications for ecological strategies. *International Journal of Plant Sciences*, 181(1):75–87, 2020.
- [144] Psaras G. K. and S. Rhizopoulou. Mesophyll structure during leaf development in ballota acetabulosa. *New Phytologist*, 131(3):303–309, 1995.
- [145] S. Rhizopoulou and Psaras G. K. Development and Structure of Drought-tolerant Leaves of the Mediterranean Shrub Capparis spinosa L. *Annals of Botany*, 92(3):377–383, 2003.
- [146] Guillaume Th  roux-Rancourt, Adam B. Roddy, J. Mason Earles, Matthew E. Gilbert, Maciej A. Zwieniecki, C. Kevin Boyce, Danny Tholen, Andrew J. McElrone, Kevin A. Simonin, and Craig R. Brodersen. Maximum co2 diffusion inside leaves is limited by the scaling of cell size and genome size. *Proceedings of the Royal Society B: Biological Sciences*, 288(1945):20203145, 2021.
- [147] A. Onuki and S. Puri. Spinodal decomposition in gels. *Phys. Rev. E*, 59:R1331–R1334, 1999.
- [148] Hajime Tanaka. Viscoelastic phase separation. *Journal of Physics: Condensed Matter*, 12(15):R207–R264, 2000.
- [149] O. Okay. *General Properties of Hydrogels*, pages 1–14. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.
- [150] N. Mahmoudi and A. Stradner. Structural arrest and dynamic localization in biocolloidal gels. *Soft Matter*, 13:4629–4635, 2017.
- [151] William L. Crepet and Karl J. Niklas. Darwin’s second “abominable mystery”: Why are there so many angiosperm species? *American Journal of Botany*, 96(1):366–381, 2009.
- [152] Peter Fratzl. Biomimetic materials research: what can we really learn from nature’s structural materials? *Journal of The Royal Society Interface*, 4(15):637–642, 2007.

- [153] Anthony D Dinsmore, John C Crocker, and Arjun G Yodh. Self-assembly of colloidal crystals. *Current Opinion in Colloid & Interface Science*, 3(1):5 – 11, 1998.
- [154] Zhenli Zhang and Sharon C. Glotzer. Self-assembly of patchy particles. *Nano Letters*, 4(8): 1407–1413, 2004.
- [155] Sharon C. Glotzer and Michael J. Solomon. Anisotropy of building blocks and their assembly into complex structures. *Nature Materials*, 6(8):557–562, 2007.
- [156] Jason D. Forster, Jin-Gyu Park, Manish Mittal, Heeso Noh, Carl F. Schreck, Corey S. O’Hern, Hui Cao, Eric M. Furst, and Eric R. Dufresne. Assembly of optical-scale dumbbells into dense photonic crystals. *ACS Nano*, 5(8):6695–6700, 2011.
- [157] Katerina Ioannidou, Konrad J. Krakowiak, Mathieu Bauchy, Christian G. Hoover, Enrico Masoero, Sidney Yip, Franz-Josef Ulm, Pierre Levitz, Roland J.-M. Pellenq, and Emanuela Del Gado. Mesoscale texture of cement hydrates. *Proceedings of the National Academy of Sciences*, 113(8):2029–2034, 2016.
- [158] Wei-Shao Wei, Yu Xia, Sophie Ettinger, Shu Yang, and A. G. Yodh. Molecular heterogeneity drives reconfigurable nematic liquid crystal drops. *Nature*, 576(7787):433–436, 2019.
- [159] Zhang, Mark A. Horsch, Monica H. Lamm, and Sharon C. Glotzer. Tethered nano building blocks: Toward a conceptual framework for nanoparticle self-assembly. *Nano Letters*, 3(10): 1341–1346, 2003.
- [160] Jennifer J. McManus, Patrick Charbonneau, Emanuela Zaccarelli, and Neer Asherie. The physics of protein self-assembly. *Current Opinion in Colloid & Interface Science*, 22:73 – 79, 2016.
- [161] Zhanpeng Zhang, Ryan L. Marson, Zhishen Ge, Sharon C. Glotzer, and Peter X. Ma. Simultaneous nano- and microscale control of nanofibrous microspheres self-assembled from star-shaped polymers. *Advanced Materials*, 27(26):3947–3952, 2015.
- [162] Fei Li, Jinsong Han, Tian Cao, William Lam, Baoer Fan, Wen Tang, Sijie Chen, Kin Lam Fok, and Linxian Li. Design of self-assembly dipeptide hydrogels and machine learning via their chemical features. *Proceedings of the National Academy of Sciences*, 116(23):11259–11264, 2019.
- [163] Clifford P. Brangwynne. Soft active aggregates: mechanics, dynamics and self-assembly of liquid-like intracellular protein bodies. *Soft Matter*, 7:3052–3059, 2011.
- [164] Jeremie Palacci, Stefano Sacanna, Asher Preska Steinberg, David J. Pine, and Paul M. Chaikin. Living crystals of light-activated colloidal surfers. *Science*, 339(6122):936–940, 2013.
- [165] Ian Linsmeier, Shiladitya Banerjee, Patrick W. Oakes, Wonyeong Jung, Taeyoon Kim, and Michael P. Murrell. Disordered actomyosin networks are sufficient to produce cooperative and telescopic contractility. *Nature Communications*, 7(1):12615, 2016.
- [166] Daniel S. Seara, Vikrant Yadav, Ian Linsmeier, A. Pasha Tabatabai, Patrick W. Oakes, S. M. Ali Tabei, Shiladitya Banerjee, and Michael P. Murrell. Entropy production rate is maximized in non-contractile actomyosin. *Nature Communications*, 9(1):4948, 2018.

- [167] Witor Wolf, Sebastian A. Kube, Sungwoo Sohn, Yujun Xie, Judy J. Cha, B. Ellen Scanley, Claudio S. Kiminami, Claudemiro Bolfarini, Walter J. Botta, and Jan Schroers. Formation and stability of complex metallic phases including quasicrystals explored through combinatorial methods. *Scientific Reports*, 9(1):7136, 2019.
- [168] Naijia Liu, Yujun Xie, Guannan Liu, Sungwoo Sohn, Arindam Raj, Guoxing Han, Bozhao Wu, Judy J. Cha, Ze Liu, and Jan Schroers. General nanomolding of ordered phases. *Phys. Rev. Lett.*, 124:036102, 2020.
- [169] Yongxiang Gao, Juntae Kim, and Matthew E. Helgeson. Microdynamics and arrest of coarsening during spinodal decomposition in thermoreversible colloidal gels. *Soft Matter*, 11:6360–6370, 2015.
- [170] Yiyong Mai and Adi Eisenberg. Self-assembly of block copolymers. *Chem. Soc. Rev.*, 41:5969–5985, 2012.
- [171] Hugues Chaté, Francesco Ginelli, Guillaume Grégoire, and Franck Raynaud. Collective motion of self-propelled particles interacting without cohesion. *Phys. Rev. E*, 77:046113, 2008.
- [172] Jordan Hoffmann, Seth Donoughe, Kathy Li, Mary K. Salcedo, and Chris H. Rycroft. A simple developmental model recapitulates complex insect wing venation patterns. *Proceedings of the National Academy of Sciences*, 115(40):9905–9910, 2018.
- [173] Aleksandra Sapala, Adam Runions, Anne-Lise Routier-Kierzkowska, Mainak Das Gupta, Lilan Hong, Hugo Hofhuis, Stéphane Verger, Gabriella Mosca, Chun-Biu Li, Angela Hay, Olivier Hamant, Adrienne HK Roeder, Miltos Tsiantis, Przemyslaw Prusinkiewicz, and Richard S Smith. Why plants make puzzle cells, and how their shape emerges. *eLife*, 7:e32794, 2018.
- [174] Nina Gabarayeva, Valentina Grigorjeva, John R. Rowley, and Alan R. Hemsley. Sporoderm development in *trevesia burckii* (araliaceae): II. post-tetrad period: Further evidence for the participation of self-assembly processes. *Review of Palaeobotany and Palynology*, 156(1):233 – 247, 2009.
- [175] Nina I. Gabarayeva, Valentina V. Grigorjeva, and Maxim O. Lavrentovich. Artificial pollen walls simulated by the tandem processes of phase separation and self-assembly in vitro. *New Phytologist*, 2019.
- [176] Asja Radja, Eric M. Horsley, Maxim O. Lavrentovich, and Alison M. Sweeney. Pollen cell wall patterns form from modulated phases. *Cell*, 176(4):856 – 868, 2019.
- [177] K J Niklas. Morphological evolution through complex domains of fitness. *Proceedings of the National Academy of Sciences*, 91(15):6772–6779, 1994.
- [178] James G. Berryman. Measurement of spatial correlation functions using image processing techniques. *Journal of Applied Physics*, 57(7):2374–2384, 1985.
- [179] P. M. Chaikin and T. C. Lubensky. *Principles of Condensed Matter Physics*. Cambridge University Press, 1995.

- [180] Di Zeng, Alexander Ribbe, and Ryan C. Hayward. Anisotropic and interconnected nanoporous materials from randomly end-linked copolymer networks. *Macromolecules*, 50(12):4668–4676, 2017.
- [181] Doros N. Theodorou and Ulrich W. Suter. Shape of unperturbed linear polymers: polypropylene. *Macromolecules*, 18(6):1206–1214, 1985.
- [182] Tsuyoshi Koga and Kyozi Kawasaki. Spinodal decomposition in binary fluids: Effects of hydrodynamic interactions. *Phys. Rev. A*, 44, 1991.
- [183] John W. Cahn and John E. Hilliard. Free energy of a nonuniform system. i. interfacial free energy. *The Journal of Chemical Physics*, 28(2):258–267, 1958.
- [184] Adrian-Marie Philippe, Luca Cipelletti, and Domenico Larobina. Mucus as an arrested phase separation gel. *Macromolecules*, 50(20):8221–8230, 2017.
- [185] Mark F. Gyure, Stephen T. Harrington, Richard Strilka, and H. Eugene Stanley. Scaling in late stage spinodal decomposition with quenched disorder. *Phys. Rev. E*, 52, 1995.
- [186] Dong Wang, John D. Treado, Arman Boromand, Blake Norwick, Michael P. Murrell, Mark D. Shattuck, and Corey S. O’Hern. The structural, vibrational, and mechanical properties of jammed packings of deformable particles in three dimensions. *Soft Matter*, 17:9901–9915, 2021.
- [187] Liyong Zhang, Delanie McEvoy, Yen Le, and Chris Ambrose. Live imaging of microtubule organization, cell expansion, and intercellular space formation in Arabidopsis leaf spongy mesophyll cells. *The Plant Cell*, 33(3):623–641, 2020.
- [188] Jr Seago, James L., Leland C. Marsh, Kevin J. Stevens, Aleš Soukup, Olga Votrubová, and Daryl E. Enstone. A re-examination of the root cortex in wetland flowering plants with respect to aerenchyma. *Annals of Botany*, 96(4):565–579, 2005.
- [189] Hamant Olivier, Heisler Marcus G., Jönsson Henrik, Krupinski Pawel, Uyttewaal Magalie, Bokov Plamen, Corson Francis, Sahlin Patrik, Boudaoud Arezki, Meyerowitz Elliot M., Couder Yves, and Traas Jan. Developmental patterning by mechanical signals in arabidopsis. *Science*, 322(5908):1650–1655, 2008.
- [190] Ryan Christopher Eng and Arun Sampathkumar. Getting into shape: the mechanics behind plant morphogenesis. *Current Opinion in Plant Biology*, 46:25 – 31, 2018. Cell Biology.
- [191] U. Zimmermann, D. Hüskens, and E. D. Schulze. Direct turgor pressure measurements in individual leaf cells of tradescantia virginiana. *Planta*, 149(5):445–453, 1980.
- [192] Julien Guénolé, Wolfram G. Nöhring, Aviral Vaid, Frédéric Houllé, Zhuocheng Xie, Aruna Prakash, and Erik Bitzek. Assessment and optimization of the fast inertial relaxation engine (fire) for energy minimization in atomistic simulations and its implementation in lammmps. *Computational Materials Science*, 175:109584, 2020.
- [193] Christopher D. Whitewoods. Riddled with holes: Understanding air space formation in plant leaves. *PLOS Biology*, 19(12):1–13, 2021.

- [194] Hans C. Andersen. Molecular dynamics simulations at constant pressure and/or temperature. *The Journal of Chemical Physics*, 72(4):2384–2393, 1980.
- [195] Nicolas Galtier and Laurent Duret. Adaptation or biased gene conversion? extending the null hypothesis of molecular evolution. *Trends in Genetics*, 23(6):273–277, 2007.
- [196] Michael Lynch, Mark C. Field, Holly V. Goodson, Harmit S. Malik, José B. Pereira-Leal, David S. Roos, Aaron P. Turkewitz, and Shelley Sazer. Evolutionary cell biology: Two origins, one objective. *Proceedings of the National Academy of Sciences*, 111(48):16990–16994, 2014.