Computational and Structural Biotechnology Journal 18 (2020) 2774–2788

COMPUTATIONAL AND STRUCTURAL BIOTECHNOLOGY JOURNAL

journal homepage: www.elsevier.com/locate/csbj

A NanoFE simulation-based surrogate machine learning model to predict mechanical functionality of protein networks from live confocal imaging

Pouyan Asgharzadeh [a,b,*,1], Annette I. Birkhold [a,b,1,2], Zubin Trivedi [a], Bugra Özdemir [c,d], Ralf Reski [c,d,e], Oliver Röhrle [a,b]

[a] Institute for Modelling and Simulation of Biomechanical Systems, University of Stuttgart, Stuttgart, Germany
[b] Stuttgart Center for Simulation Science (SC SimTech), Stuttgart, Germany
[c] Plant Biotechnology, Faculty of Biology, University of Freiburg, Freiburg, Germany
[d] Signalling Research Centres BIOSS and CIBSS, Freiburg, Germany
[e] Cluster of Excellence livMatS @ FIT – Freiburg Centre for Interactive Materials and Bioinspired Technologies, Freiburg, Germany

ARTICLE INFO

Article history:
Received 6 July 2020
Received in revised form 12 September 2020
Accepted 13 September 2020
Available online 24 September 2020

Keywords:
Structure-function relationship
Protein network
Machine learning
Finite element analysis
Confocal imaging

ABSTRACT

Sub-cellular mechanics plays a crucial role in a variety of biological functions and dysfunctions. Due to the strong structure-function relationship in cytoskeletal protein networks, light can be shed on their mechanical functionality by investigating their structures. Here, we present a data-driven approach employing a combination of confocal live imaging of fluorescent tagged protein networks, in silico mechanical experiments and machine learning to investigate this relationship. Our designed image processing and nanoFE mechanical simulation framework resolves the structure and mechanical behaviour of cytoskeletal networks and the developed gradient boosting surrogate models linking network structure to its functionality. In this study, for the first time, we perform mechanical simulations of Filamentous Temperature Sensitive Z (FtsZ) complex protein networks with realistic network geometry depicting its skeletal functionality inside organelles (here, chloroplasts) of the moss *Physcomitrella patens*. Training on synthetically produced simulation data enables predicting the mechanical characteristics of FtsZ network purely based on its structural features ($R^2 \geqslant 0.93$), therefore allowing to extract structural principles enabling specific mechanical traits of FtsZ, such as load bearing and resistance to buckling failure in case of large network deformation.

1. Introduction

Bio-polymer networks are pervasive as key promoters of strength, support and integrity. This is true irrespectively of their scale, i.e., from the nano-scale of the cytoskeleton to the macro-scale of connective tissues. As cells sense external physical signals and translate them into a cellular responses, cellular mechanics has been proven to be crucial for a wide range of biological functions and dysfunctions. In particular cytoskeletal protein networks exhibit strong structure-function relationships, e.g., the role of microtubule network during mitosis [1], cell movement with the help of actin assembly/disassembly [2] or utilizing intermediate filament networks for stabilizing mechanical stresses [3]. In all these processes protein networks transmit besides biochemical also biophysical cues from the cell microenvironment that trigger and regulate cell behaviors. Investigating the structure of protein networks allows deeper insights into cellular functionality and dysfunctionality and may further elucidate many critical cell responses observed in vivo. Understanding these cellular responses in 3D may further lead to the development of functional and biomimetic materials [4] for engineering the 3D cell microenvironment able to control cell behaviors in 3D and may advance the fields of tissue regeneration and in vitro tissue models.

In the recent decades, taking advantage of new methodological developments in experimental and computational physics and applying them to biological systems allowed substantial progress in elucidating particular mechanical phenomena to biological function. It have shown that mechanical processes convey biochemical signals and are therefore crucial for cell functions including proliferation, polarity, migration and differentiation. Further,

* Corresponding author at: Institute for Modelling and Simulation of Biomechanical Systems, University of Stuttgart, Stuttgart, Germany.
E-mail address: asgharzadeh@imsb.uni-stuttgart.de (P. Asgharzadeh).
URL: http://bit.ly/2Tqx_PA (P. Asgharzadeh).
[1] These authors contributed equally to this work.
[2] Currently an employee of Siemens Healthcare GmbH.

https://doi.org/10.1016/j.csbj.2020.09.024

connections between the mechanical properties of cells and initiation and pathological progression of cancer were established [5,6]. Mendez et al. and Liu et al. showed that the epithelial-to-mensenchymal transition (EMT) leading to cancer metastasis are linked to changes in mechanical characteristics of the cytoskeleton influencing the vimentin network [7,8] as well as the polarity of the cell [9]. Moreover, cancer cells are typically found to be softer than normal cells. A decrease in the level of actin in the cytoskeleton of cancerous cells was linked to changes in the mechanical properties of the cell [10]. Such research underpins the importance of linking molecular changes within the cytoskeleton to structural and functional changes of the entire cell and therefore changes to the tissue.

In summary, in-depth knowledge of cellular and sub-cellular mechanics might allow the identification and classification of cells at different physiological and patho-physiological stages. However, to do so, new approaches need to be developed that are capable of simultaneously performing structural and mechanical analysis of sub-cellular structures in a (semi-) automated way. Mechanical stability and its contribution to shaping processes on the molecular scale are far from being completely understood. Further, it is not clear, if mechanical processes, besides conveying biochemical signals, also purely convey mechanical signals to invoke structural changes. The concept of the cytoskeleton as a shape-determining scaffold for the cell is well established [11], however, the tight coupling of actin, microtubule and intermediate filament networks impedes a separate analysis. To date, computer models of cytoskeletal biopolymer networks are based on models that represent the geometry in a (strongly) simplified way [12–16]. In depth analysis of structure-function relationships, however, require detailed structural and functional modelling.

Development of such models requires a protein network with similar structural functionality to cytoskeletal networks while being structurally less complicated, therefore allowing a detailed validation of the derived results. Proteins homologous to tubulin, which is part of the eukaryote cytoskeleton, such as the Filamentous Temperature Sensitive Z (FtsZ) protein family in the chloroplasts of the moss *Physcomitrella patens* are excellent examples for this purpose. FtsZ generates complex polymer networks, showing striking similarity to the cytoskeleton, and hence were named plastoskeleton [17]. In bacteria, FtsZ is a part of the bacterial cytoskeleton providing a scaffold for cell division [18–20].

The genome of *Physcomitrella patens* encodes three subclades breaking down into five isoforms naming FtsZ1-1, FtsZ1-2, FtsZ2-1, FtsZ2-2 and FtsZ3 [21,22]. Utilizing reverse genetics, different essential functionality of the two main classes of FtsZ isoforms (FtsZ1 and FtsZ2) were observed [23]. It has been shown that during evolution, functional characteristics of the FtsZ isoforms has diverged [23]. Coassembly experiments provide evidence that FtsZ2 controls filament morphology and FtsZ1 promotes protofilament turnover. It is suggested, that *in vivo*, FtsZ2 forms the chloroplast Z-ring backbone while FtsZ1 facilitates Z-ring remodeling [24]. These differences in functionality of the FtsZ isoforms point towards existence of distinguishing structural traits between these two isoforms. Moreover, as chloroplasts in loss-of-function mutants show distinct shape defects, FtsZ networks might provide scaffolds that ensure the stability and structural integrity of the chloroplast [25]. Additionally, gene knock-out experiments have shown that the FtsZ network is capable of undergoing large deformations upholding its structural integrity [23]. This adaptive stability is presumably linked to the developed structural characteristics of FtsZ network; making the cytoskeletal FtsZ network an ideal first application for introducing and testing a simulation-based method that aims at identifying a link between structural features of a cytoskeletal network and its mechanical functions.

To date, investigating the characteristics of FtsZ isoform network structure has been mostly carried out in form of visual inspection and description [24,26]. We have previously developed a framework for quantitatively describing the structure of cytoskeletal protein networks [27] and utilized it to compare the FtsZ1-2 and FtsZ2-1 network structures [28]. As the next step, automatically finding the distinguishing characteristics of these two isoforms and linking their structural features to their functional characteristics may shed light on the manifestation of functionality in the network structure. State-of-the-art microscopy imaging techniques permit resolving micro-structural details of protein networks. Computational analysis of acquired images facilitates the quantification of components and its assembly to networks [27], and may allow tracking structural changes of the network assembly triggered by internal or external stimuli, i.e., connecting the structure to functionality or distinguishing between network types [28]. Machine learning (ML) algorithms have proven to be remarkably capable of automating such complex image analysis tasks [29] and of correlating image content to biological structural functionality [30–32]. Recently, the concept of ML-based surrogate models has proven to be highly advantageous in accelerating the performance of numerical simulations of complex mechanical environments [33] as well as predicting material properties [34]. A ML-based approach could link structural features to mechanical characteristics and would provide a way to answer questions like "How are FtsZ biopolymers capable of exhibiting adaptive stability?" or "Interplay of which structural changes in the cytoskeleton of a cancerous cell leads to adapting stiffness?"

To overcome the challenge of relating structure to function of cytoskeletal protein networks, we present an ML approach applied to 3D live laser scanning confocal microscopy images. The outcome is an end-to-end tool that links structural features associated with the cytoskeletal network type to its mechanical behaviour and therefore enables a fast evaluation of structure-function relations on the sub-cellular scale. This is carried out by combining an in silico mechanical characterization of protein networks through 3D nano finite element modeling and an automatic mapping of structural features to the mechanical network responses. The introduced method consists of two different models. The first model classifies the respective protein networks based on their structural features by exploiting an image processing and a gradient boosting classification model. The second one creates an in silico surrogate model to predict the sub-cellular mechanical responses of the network. Analyzing the prediction process of the surrogate model based on the structural feature allows us to deduct the presumed structure-function relationship. The method is tested and applied to elucidate isoform-specific structure-function relationships of FtsZ networks.

2. Materials and methods

2.1. Materials

The "Gransden 2004" ecotype of the moss *Physcomitrella patens* ((Hedw.) Bruch & Schimp., IMSC accession number 40001) was cultivated in bioreactors [35].

2.2. Molecular biology and moss transfection

RNA isolation, molecular cloning and moss transfection were described previously in detail [27,28] and are therefore given here only in a shortened version. Total RNA was isolated from wild type *Physcomitrella patens* protonema using TRIzol Reagent (Thermo Fisher Scientific) and used for cDNA synthesis using Superscript III reverse transcriptase (Life Technologies, Carlsbad, CA, USA).

The coding sequences of PpFtsZ1-2 and PpFtsZ2-1 were PCR-amplified from this cDNA and cloned into the reporter plasmid pAct5::Linker:EGFP-MAV4 (modified from [36]) to generate the fusion constructs PpAct5::PpFtsZ1-2::linker::EGFP-MAV4 and PpAct5::PpFtsZ2-1::linker::EGFP-MAV4. Moss protoplasts were isolated and transfected with 50 µg of each of these plasmids, according to the protocol described by Hohe et al. [37]. The transfected protoplasts were incubated for 24 h in the dark, subsequently being returned to normal conditions (25.1 °C; light-dark regime of 16 : 8 h light flux of 55 $\mu mol\, s^{-1}\, m^{-2}$ from fluorescent tubes, Philips TL – 19-65W/25).

2.3. Laser scanning confocal microscopy imaging

In 4 – 7 days after transfection, the protoplasts were concentrated to a volume of 100 µl, and 20 µl of this protoplast suspension was used for imaging. Confocal laser scanning microscopy (CLSM) images were taken with a Leica TCS SP8 microscope (Leica Microsystems, Wetzlar, Germany), using the HCX PL APO 100x/1.40 oil objective and applying the microscopy conditions described previously [27,28]. A selection of images visualising FtsZ networks is depicted in Fig. 1a. To summarise, the zoom factor was 10.6, the voxel sizes were 21 nm in the $X-Y$ dimensions and 240 nm in the Z dimension and the pinhole was adjusted to $0.70AU$ (66.8 µm). For the excitation, a WLL laser was applied at 488 nm with an intensity of 4%. The detection ranges were set to 503–552 nm for the EGFP fluorescence and 664–725 nm for the chlorophyll autofluorescence. All images were deconvolved using Huygens Professional version 17.04 (Scientific Volume Imaging, The Netherlands). The effect of the deconvolution on the images is illustrated in the Supplementary Materials. The imaging and deconvolution protocol resulted in a dataset of $n = 37$ 3D CLSM images (i.e., 21 FtsZ2-1 and 16 FtsZ1-2 isoforms). The chloroplast morphology was not affected by exposure to the laser for the period of the Z-stack acquisition; hence, phototoxicity is negligible (Supplementary Video).

2.4. Image processing to extract structural features

A set of 26 structural features describing the assembly of protein networks from global and local perspectives is extracted from each network. Here, only a short description of the workflow steps and features are given, details as well as a validation have been reported previously [27].

2.4.1. Image pre-processing

Images are segmented using an adaptive local threshold, $T = m + k\sqrt{\frac{1}{NP}\sum_{i=1}^{NP}(p_i - m)^2}$, with $NP = 10 * 10 * 10$ being the local window size, m is the average pixel intensity in the window, p_i denotes the intensity of pixel i and $k = 10$ as a constant value. Next, by calculating the convex hull of the segmented network a solid outer surface representing the volume enclosing the network is determined. To extract the structural features of the network, a transformation to a spatial graph consisting of points, nodes and segments is performed. This transformation consists of following steps: 1. determining the centerline of each filament based on calculating a distance map for each foreground voxel from the edge voxels, and 2. placing points at the centerline of the filaments where either thickness or the direction of the filament changes. The resulting hierarchy of structural elements of the spatial graph reads as: 1. points, 2. elements as the connection between points, 3. nodes, as points that are connected to more than two other points, 4. segments as summation of elements from one node to another (filaments), 5. connections as the meeting points of filaments in a node. This numerical representative in form of a spatial graph allows determining structural features (Fig. 1b).

2.4.2. Features describing 3D network

26 structural features are designed to quantitatively describe the structure of the protein network from two perspectives including overall network morphology and detailed network structure. Seven shape descriptors, which are extracted by analyzing the segmented network and its convex hull, quantitatively describe the overall morphology of the network [27]. Moreover, 19 local structural features of the network are determined employing a spatial graph generation method [27]. These consist of nodal features, segment features and connection feature. The designed features and their definitions are presented in table 1. All morphological and structural features are calculated by a previously validated set of in-house Matlab codes (Matlab 2019a, MathWorks, USA) [27].

2.5. Mechanical nano-FE modeling

To investigate the mechanical response of the protein networks to external load, we designed a generic in silico experiment reflecting a compression against a plate, hence, a scenario that is typically used to experimentally investigate the mechanical behaviour of whole cells [38]. To capture the overall mechanical behavior of each network in a comparative manner, compression tests along the three principal axis of each system were modeled employing a nano-FE approach. All simulations were done with the finite element analysis software Abaqus 6.14 (Dassault Systémes, France). It is important to note that, the goal of this setup is to depict the mechanical behavior of the protein network morphology rather than replicating the real physical condition and dynamics of the biopolymers in their biological roles. Such a task would require consideration of highly complicated interactions of the network with its surrounding, which are not completely understood to date.

2.5.1. 3D protein network model generation

For all samples, protein network surface meshes were defined from the segmented images using a triangular approximation algorithm coupled with a best isotropic vertex placement algorithm to achieve high triangulation quality [39]. The surface area of the resulting surface mesh was calculated and further remeshed using $n_t = \rho_t A_t$ triangles for the remeshed surface, where $\rho_t = 900\,[triangles/\mu m^2]$ is the constant surface mesh density and A_t denotes the surface area. Furthermore, the remeshed surface was smoothed by shifting the vertices towards the average position of its neighbours. The enclosed surfaces were filled with volumetric tetrahedral elements, resulting in an adaptive multi-resolution grid (Fig. 1b) using FEI Amira 6.3.0 (Thermo Fisher Scientific, USA).

The principal directions of a network were determined based on its convex hull and shape matrix. The eigenvectors of the shape matrix ($EV1$, $EV2$ and $EV3$), which are orthogonal to each other, represent the network's principal directions, V_i ($i = 1,2,3$). The mesh is then transformed to the coordinate system spanned by V_i. Afterwards, along each V_i, the pair of nodes exhibiting the largest distance in-between the two points and in the direction of V_i were determined and named N_{i1} and N_{i2} and $i = 1,2,3$, respectively.

For each considered protein network, the geometry of the protein network was imported to Abaqus and three compression simulations (one per direction d) were carried out. For each simulation, the initial set-up was determined by first identifying the initial position of two parallel rigid plates, which are defined

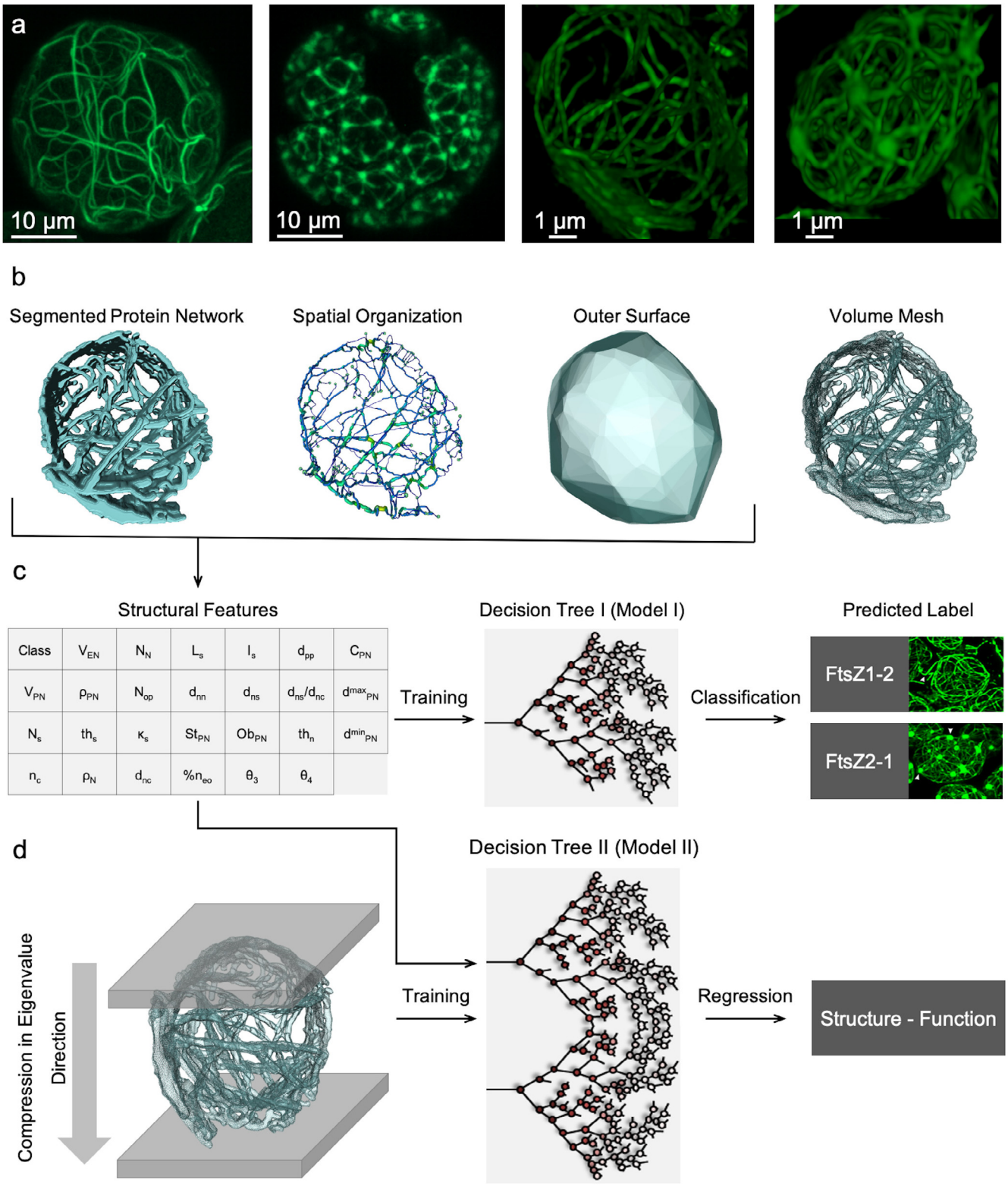

Fig. 1. Correlating structure of FtsZ network to its mechanical functionality utilizing CLSM images. a) Sample 3D CLSM images of FtsZ1-2 and FtsZ2-1 networks in a cell (two left images) and single FtsZ1-2 and 2-1 networks (two right images), respectively. b) Sample of a 3D segmented image and its spatial graph, convex hull and mesh. c) The 26 shape and element descriptors that are extracted and used as input features to train a gradient boosting model for classifying FtsZ1-2 and FtsZ2-1 isoforms. d) A second gradient boosting model (regression) is trained on the structural features to predict the results of the mechanical simulation of compressing the network in its principal directions (3 Eigenvectors determined from segmented images).

for each simulation in direction d by its normal vector $\overrightarrow{N_{d1}N_{d2}}$ (with $d = 1, 2,$ or 3), and the respective nodal points N_{d1} and N_{d2} (cf. Fig. 2a-c).

2.5.2. Governing equations

The simulations were carried out by solving the balance of linear momentum,

Table 1
Definitions of the employed structural features describing the protein network morphology.

Network gross morphology	
Feature	Definition
V_{EN}	volume of the convex hull
d_{PN}^{max}	greatest diameter of the network
V_{PN}	volume of the segmented protein network
d_{PN}^{min}	smallest diameter of the network
ρ_{PN}	network volume density
St_{PN}	stretch of the network
Ob_{PN}	oblateness of the network
Detailed network structure	
Feature	Definition
N_n	number of nodes
th_n	thickness of nodes
ρ_n	node density
d_{nn}	node-to-node distance
d_{ns}	node-to-surface distance
d_{nc}	node-to-centre distance
N_{op}	number of open nodes
C_{PN}	$(d_{nc} - d_{ns})/d_{nc}$
d_{ns}/d_{nc}	node-to-surface to node-to-center ratio
N_s	number of segments
L_s	segment length
κ_s	segment curvature
th_s	mean segment thickness
I_s	segment inhomogeneity
d_{pp}	segment point-to-point distance
n_{oe}	percentage of open nodes
n_c	mean number of connections per node
θ_3 and θ_4	mean angles between segments in a connection with 3and 4 filaments meeting

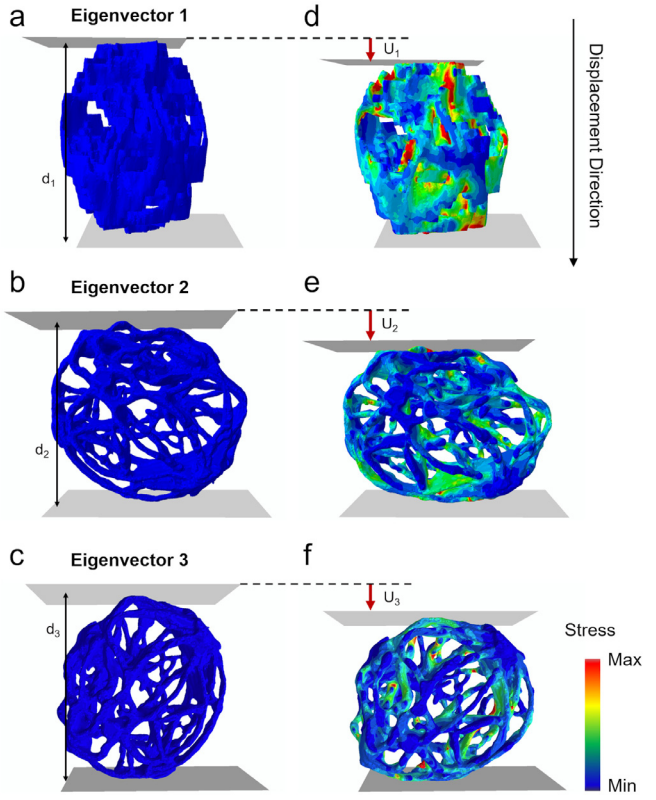

Fig. 2. Simulation setups. a–c) Initial set-up for virtual compression experiments of a sample protein network in primary directions EV1, EV2 and EV3, respectively. d–f) Stress distribution after applying a displacement ($\alpha = 0.02$) to the upper plate in EV1, EV2 and EV3 directions, respectively. The displacements are scaled by a factor of 5.

$$\rho\dot{\mathbf{v}} - \nabla \cdot \boldsymbol{\sigma} - \rho\mathbf{b} = \mathbf{0}, \tag{1}$$

in an explicit manner, where ρ is the mass density, $\mathbf{v}$ denotes the velocity, $\boldsymbol{\sigma}$ describes the Cauchy stress, and $\mathbf{b}$ are the body forces. The contact between the protein network and the rigid plates was chosen as a rough contact meaning that any two points, which come in contact, will stick together (relative penetration tolerance: $1\,e-3$). The governing equation is discretized using the FE method, tetrahedral elements and linear spatial Ansatz functions.

2.5.3. Boundary conditions

The generic boundary conditions for each simulation setup (one simulation for each primary directions: $EV1$, $EV2$ and $EV3$) consist of applying displacement boundary conditions at node N_{d1} to mimic compression experiments. The displacement itself is applied in the $EV_d = \overrightarrow{N_{d1}N_{d2}}$ direction and in fractions (α) of the initial distance, $\|\overrightarrow{N_{d1}N_{d2}}\|$, between the two plates (cf. Fig. 2). Therefore, the amount of displacement along the respective Eigenvector EV_1 is defined by $U_d = \alpha \cdot \|\overrightarrow{N_{d1}N_{d2}}\|$.

To investigate anisotropy in the mechanical response of the network, compression tests along all three primary directions were performed and compared. Furthermore, for analyzing changes in the structural behavior with increasing deformation grade, we increased the displacement of the plate gradually in steps of $\delta\alpha = 0.02$. We applied a total of 10 steps, which is equivalent to $\alpha = 0.20$. Due to no apparent significant differences in the mechanical behaviour of the network between the three directions at $\alpha = 0.02$ and the required computational resources, we chose to focus only on one direction to continue the simulations for $\alpha = 0.02 \rightarrow 0.20$.

2.5.4. Constitutive law and material parameters

Employing the concept of linear elasticity, the stress tensor $\boldsymbol{\sigma}$ is given by

$$\boldsymbol{\sigma} = \lambda \mathrm{tr}(\boldsymbol{\epsilon})\mathbf{I} + 2\mu\boldsymbol{\epsilon}, \tag{2}$$

where $\boldsymbol{\epsilon}$ denotes the strain tensor and $\boldsymbol{I}$ is the second-order identity tensor. Further, λ and μ are the first and the second Lamé coefficients, respectively. The Lamé coefficients are related to Young's modulus E and Poisson's ratio ν by

$$\lambda = \frac{E\nu}{(1+\nu)(1-2\nu)}, \quad \mu = \frac{E}{2(1+\nu)}. \tag{3}$$

In classical continuum mechanics, the material parameters (here λ and μ or E and ν) are obtained by making a constitutive assumption, i.e., selecting a particular phenomenological constitutive law (here the form of $\boldsymbol{\sigma}$) and ensuring that the computed stresses match the experimental ones. The mechanical behavior of filamentous biopolymers is, however, commonly quantified by means of the flexural rigidity, κ [40–42], which is the force couple required for one unit of curvature [43]. The flexural rigidity, κ, is defined as $\kappa = EI$, where I is the second moment of inertia. In the context of protein filament mechanics, the flexural rigidity is calculated as $\kappa = k_B T l_p$, where $k_B = 1.38 \times 10^{-23}$ J/K is the Boltzmann constant, T defines the temperature, and l_p denotes the corresponding thermal persistence length. Recently, persistence length and flexural rigidity of FtsZ filaments reported by Turner et al. as $\kappa = 4.7 \pm 1.0 \times 10^{-27}$ Nm^2 and $l_p = 1.15 \pm 0.25\,\mu$m are commonly employed [44,45]. The average thickness of filamentous elements of the FtsZ network has been reported to be 117 ± 28 nm [28]. Assuming circular cross sections, I equals 1.81×10^{-29} m^4 [46]. Based on these values, we set within our simulations the elasticity modulus to $E = 2.6 \times 10^2$ Pa and the Poisson's ratio to $\nu = 0.5$, i.e., assuming incompressibility [47–49].

2.5.5. Calculated mechanical parameters

We performed a total of 111 simulations (3 simulations per network) on a CPU cluster with 32 cores (4 AMD Opteron Socket G34 Eight-Core 6328, 3.2 GHz, 8C, Abu Dhabi). One simulation took on average 19 ± 7 hours (for the entire 20% compression).

To quantitatively assess the mechanical behavior of the protein network isoforms, we determined four mechanical parameters using NanoFE simulations. These four parameters consists of mean mechanical stress ($\bar{\sigma}$), mean mechanical strain ($\bar{\epsilon}$), buckling failure factor (*FB*) and rupture failure factor (*FR*). Stress represents the internal forces that neighbouring particles of a continuous material exert on each other, strain refers to the deformation of a material in response to a mechanical loading. $\bar{\sigma}$ and $\bar{\epsilon}$ are determined by averaging the L1 norms of the von Mises stresses and principal strains over the elements of the NanoFE model, respectively. *FB* and *FR* are introduced to describe the mechanical failure behaviour of the protein networks. Since a local failure might not lead to a collapse of the whole network structure, we define failure factors based on the assumption that if a certain portion m of all elements ($m \cdot n_{all}^{elem}$) exhibit stresses or strains above the critical stress or strain value, the whole structure will fail by buckling or rupture of an individual or several segments, as shown for other biological materials [50,38]. Since for protein networks, these threshold values have not been experimentally investigated yet, we report only the portion of elements that exceed a particular critical stress or strain value, i.e., the higher the values the higher the failure probability. We define the buckling failure factor as $FB = n_{\sigma_{crit}}^{elem}/n_{all}^{elem}$, and the rupture failure factor as $FR = n_{\epsilon_{crit}}^{elem}/n_{all}^{elem}$, where n_{all}^{elem} is the total number of elements, $n_{\sigma_{crit}}^{elem}$ and $n_{\epsilon_{crit}}^{elem}$ are the number of elements with stress and strain exceeding the critical buckling or rupture values, respectively.

Cytoskeletal structures are reported to fail by buckling or rupture [51]. We therefore, further analyzed the structural stability of the network by calculating a buckling failure factor based on critical stresses σ_{crit} and a rupture failure factor based on critical strains ϵ_{crit}. Buckling of a single filament is assumed to occur if local von Mises stresses exceed a critical value. A filament is assumed to rupture, if strains locally exceed a critical strain value. To our knowledge, σ_{crit} and ϵ_{crit} of FtsZ have not been measured to date. However, despite fundamental structural differences, F-actin and FtsZ show similar mechanical behavior. The rigidity of F-actin is assumed to be $\kappa = 7.5 \cdot 10^{-26}$ Nm^2 [52,40] whereas the rigidity, which we assume for FtsZ filaments, is $\kappa = 4.7 \pm 1.0 \times 10^{-27}$ Nm^2 (l_p of F-actin: 1.77 μm and l_p of FtsZ: 1.15 µm [40,44]). Therefore, we use the values reported for F-actin ($\sigma_{crit} = 3.2$ Pa and $\epsilon_{crit} = 0.2$ [52,48]).

2.6. Data-driven analysis using machine learning

To relate the mechanical functionality of protein networks to their structure, we utilize a ML approach. To do so, we trained two sets of ML models on the dataset containing the 26 calculated structural features of the $n = 37$ protein networks. The aim of these ML models is to 1) perform a classification of the networks as well as an analysis of the structural features dominating the decision process and 2) map the structural features of the network to its mechanical behavior by employing a regression model. This further allows us to identify the most dominant structural features contributing to specific mechanical traits of the network.

2.6.1. Classification of FtsZ isoforms

We designed and trained a gradient boosting model [53], to perform the prediction task based on the extracted features (Fig. 1c):

$$f^*(\vec{x}) = y, \tag{4}$$

where we produce a prediction model (f^*) in the form of an ensemble of weak prediction models to map the set of protein structural features:

$$\vec{x} = \Big[V_{EN}, V_{PN}, \rho_{PN}, d_{PN}^{max}, d_{PN}^{min}, St_{PN}, Ob_{PN}, N_n, th_n, \rho_n, d_{nn}, d_{ns}, d_{nc}, d_{ns}/ d_{nc}, C_{PN}, N_{op}, N_s, L_s, \kappa_s, th_s, I_s, d_{pp}, n_c, n_{oe}, \theta_3, \theta_4\Big],$$

to an isoform (y: FtsZ1-2 or FtsZ2-1). During learning, we consecutively fit new models to provide a more accurate prediction [54]. The new models are constructed to be maximally correlated with the negative gradient of the loss function Ψ, associated with how wrong the prediction is. Given N_n training examples: $\{(\vec{x}_1, y_1), \cdots, (\vec{x}_N, y_N)\}$, where $\vec{x}_i \in \vec{x}$ and $y_i \in y$, the gradient boosting decision tree model estimates the function f of future $\vec{x}$ by the linear combination of individual decision trees

$$f_M(\vec{x}) = \sum_{m=1}^{M} T(\vec{x}; \theta_m), \tag{5}$$

where $T(\vec{x}; \theta_m)$ denotes the m-th decision tree, θ_m is its parameter set, M is the number of decision trees. The final estimation is determined in a forward stage-wise fashion, i.e. based on an initial model $f_0(\vec{x})$ of $\vec{x}$, the model of step m is determined as:

$$f_m(\vec{x}) = f_{m-1}(\vec{x}) + T(\vec{x}; \theta_m), \tag{6}$$

where $f_{m-1}(\vec{x})$ is the model in step $m-1$. θ_m is learned by empirical loss minimization as

$$\theta_m = \arg\min_{\theta_m} \sum_{i=1}^{M} \Psi(y_i, f_{m-1}(\vec{x}) + T(\vec{x}; \theta_m)), \tag{7}$$

with the loss function Ψ. The assumption of linear additivity of the base function, leads to the estimation of θ_m for the best fitting residual $\Psi(y - f_{m-1}(\vec{x}))$. To this end, the negative gradient of the loss function at f_{m-1} is used to approximate the residual R:

$$R_{m,i} = -\left[\partial\Psi\left(\frac{y, f(x_i)}{\partial f(x_i)}\right]_{f(x)=f_{m-1}(x)}\right., \tag{8}$$

with i as the index of the i-th example.

We randomly divided the dataset into 80% training set ($n = 30$) and 20% test set ($n = 7$) with test set being stratified for number of isoforms. The model is trained on the training set employing a 5-fold cross validation [55] and applied on the test set.

2.6.2. Surrogate mechanical model for predicting function from structure

To investigate the structural approach(es) employed by nature to provide networks with specific mechanical functionality, e.g. adaptive stability, a set of surrogate models (SM) in the form of regression boosted gradient models are designed as:

$$\mathrm{SM}_i(\vec{x}) = M. \tag{9}$$

These 8 surrogate models serve as a tool to map the structural features ($\vec{x}$, including the isoform class) of the protein networks to their mechanical behavior (Fig. 1d), with $M \in [\bar{\sigma}, \bar{\epsilon}, FB, FR]$ as the calculated mechanical parameter in *EV*3 direction (cf. 2.5.5) for small and large deformations ($\alpha = 0.02$ and 0.20) separately. The design and training of the regression surrogate models are similar to the described classifier model with the only difference being the definition of the loss function.

Here, the dataset is randomly divided into 90% training set ($n = 33$) and 10% test set ($n = 4$) with test set being stratified for number of isoforms. The model is trained on the training set employing a 5-fold cross validation and applied on the test set. Moreover, to further validate the performance of the gradient

Table 2
Classification and regression model parameters and definition.

Classification output: Network isoform	
y	Definition
FtsZ1-2	An isoform of FtsZ, which presumably promotes protofilament turnover
FtsZ1-2	An isoform of FtsZ, which presumably controls filament morphology
Regression output: Network mechanical behaviour	
M	Definition
$\bar{\sigma}$	Mean L1 norms of the von Mises stresses in simulated model
$\bar{\epsilon}$	Mean L1 norms of the principal strains in simulated model
FB	Ratio of elements failed due to buckling to all elements
FR	Ratio of elements failed due to rupture to all elements

boosting model for both classification and regression tasks, the results are compared to a random forest model as another ensemble method [56]. Definitions of parameters predicted by the two ML model types in equations 4 and 9 are presented in Table 2.

2.6.3. Analyzing feature importance

We determined the importance of each structural feature in both, the classifier and the surrogate mechanical models. To do so, each feature is noised up and the plurality of out-of-bag vote and the reality are determined allowing to measure a wrong prediction rate [57] for each feature.

2.7. Statistical analysis and model performance evaluation

To distinguish the mechanical behavior of FtsZ1-2 isoforms from FtsZ2-1 isoforms, statistical analysis of $\bar{\sigma}, \bar{\epsilon}, FB$ and FR was performed using repeated measures ANOVA and paired or unpaired student's t-tests, as appropriate, followed by Bonferroni corrections for multiple comparisons. All values are presented as mean ± standard deviation and statistical significance was set to $p < 0.05$.

The performance of the both classifer model and the surrogate models is assessed by applying the models on the corresponding test set. The classifier model is evaluated by calculating the $F1$-score and the accuracy of prediction [58]. The performance of each surrogate models is assessed by calculating R^2-values between the model predictions and simulation results. To further measure the differences between the surrogate model predictions and the true values of simulation results, we calculated the mean percentage error (MPE) of prediction as well as the slope of the linear fit for scattered data of model predictions vs. simulation results ($\hat{y}$).

3. Results

3.1. Effect of load direction on mechanical response

$N = 16$ FtsZ1-2 and $n = 21$ FtsZ2-1 isoforms images (see examples in Fig. 3a, e) were processed. For each protein network, the image processing resulted in distinct spatial graphs (Fig. 3b, f), convex hulls (Fig. 3c, g; Supplementary Video), and FE meshes (Fig. 3d, h).

All mechanical parameters ($\bar{\sigma}, \bar{\epsilon}, FR$ and FB) were affected by load direction as well as isoform type (ANOVA, $p < 0.01$). $\bar{\sigma}$, in the FtsZ1-2 isoform, was significantly lower for the EV2 load case than for the other two load cases ($p = 0.01$, Fig. 4a). In FtsZ2-1 bucking failure (FR) was significantly lower for the EV1 load case than the EV3 load case ($p < 0.01$, Fig. 4d). Comparing the mechanical parameters ($x, \bar{\epsilon}, FR, FB$) between the isoforms revealed that all mechanical parameters of the FtsZ2-1 isoform were for the EV2 loading case significantly higher than for the FtsZ1-2 isoform ($p \leq 0.04$; Fig. 4a-d). Additionally, FtsZ2-1 responded to compression in EV3 direction with a significant higher $\bar{\epsilon}$ than FtsZ1-2 ($p = 0.049$; Fig. 4b).

3.2. Effect of displacement on mechanical response

The mechanical response with increasing compression was investigated only in EV3 direction. Increasing the compression of the isoforms from 2% to 20% (Fig. 5; Supplementary Video 2) revealed that, at all displacement steps, no significant difference between the two isoforms in the four calculated mechanical parameters occurred (Fig. 5c–f). For mean stress, mean strain and rupture failure factor (Fig. 5 gradual increase in both network types was detected with increasing displacements. In contrast, with increasing compression, FB converges toward a buckling failure factor of 1% (Fig. 5f).

Comparing the mechanical responses at 20% displacement shows no significant difference between the two FtsZ isoforms (Fig. 5h). At 20% displacement, the buckling failure factor (FtsZ1-2: $1.0 \pm 0.1\%$ and FtsZ2-1: $1.0 \pm 0.1\%$) is significantly higher than the rupture failure factor (FtsZ1-2: $0.3 \pm 0.2\%$ and FtsZ2-1:

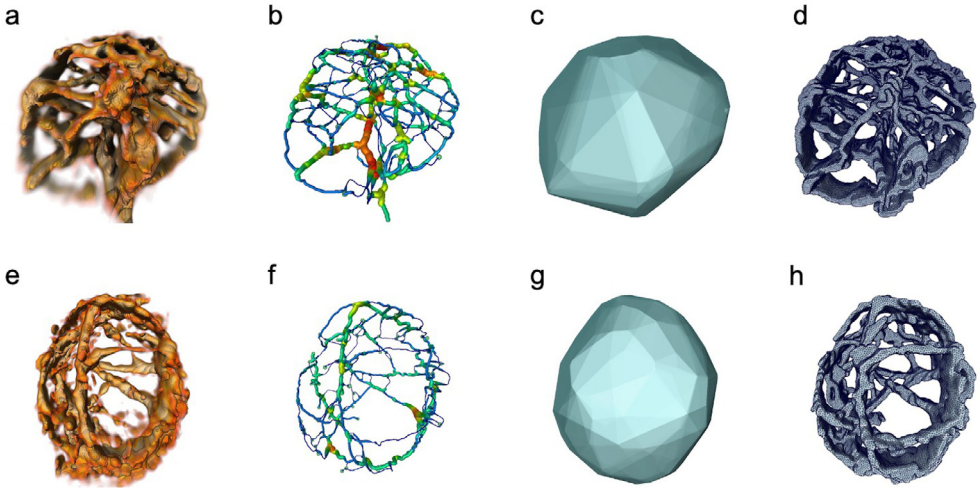

Fig. 3. Image pre-processing of FtsZ1-2 (a–d) and FtsZ2-1 (e–h) isoforms. a) Sample 3D CLSM image of FtsZ1-2 isoform, b) resulting spatial graph, c) resulting convex hull and d) resulting volume mesh. e) Sample 3D CLSM image of FtsZ2-1 isoform, f) resulting spatial graph, g) resulting convex hull and h) resulting volume mesh.

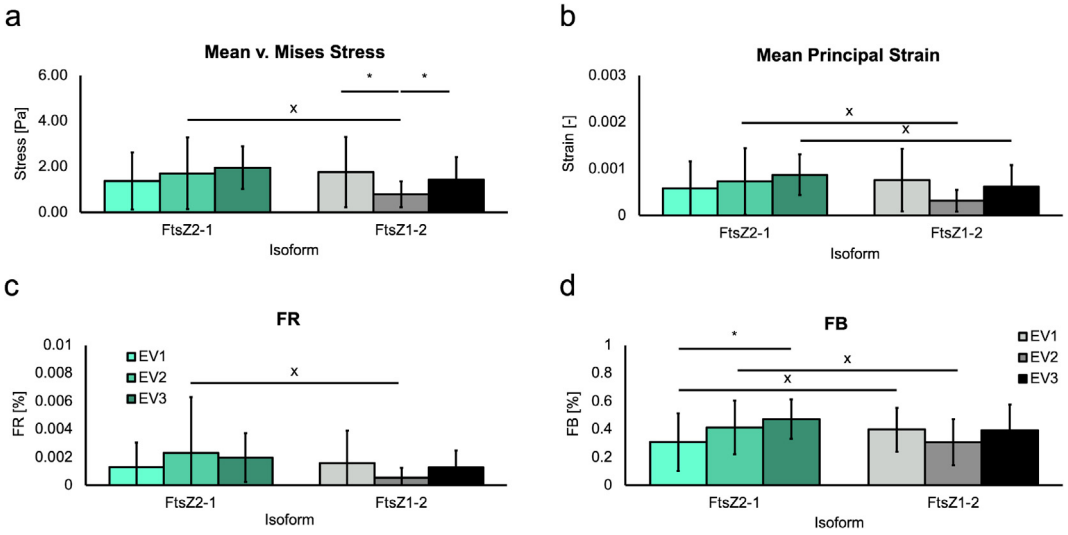

Fig. 4. Mechanical responses to small deformations (2% compression). a) $\bar{\sigma}$. b) $\bar{\epsilon}$. c) *FR* d). *FB*. Data is shown as mean±standard deviation. * denotes a significant difference between load directions (student's t-test, Bonferroni correction), X denotes a significant difference between isoforms. Data is shown as mean±standard deviation.

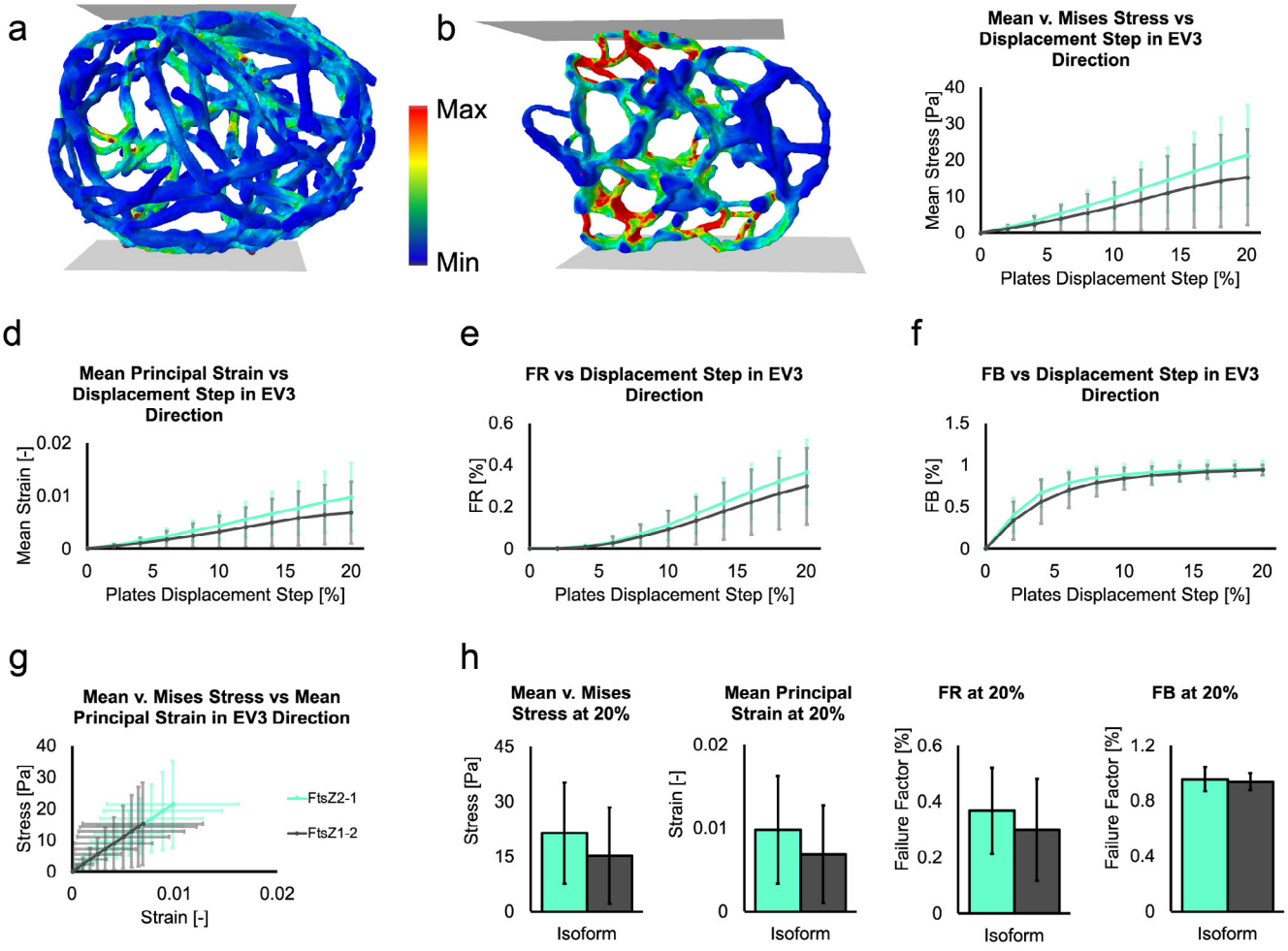

Fig. 5. Changes in mechanical response with increasing compression. a, b) Stress distributions at 20% displacement in sample networks of FtsZ1-2 an FtsZ2-1, respectively. c) $\bar{\sigma}_{EV3}$. d) $\bar{\epsilon}_{EV3}$. e) FR_{EV3}. f) FB_{EV3}. g) Mean stress vs mean strain in EV3 direction. h) Calculated mechanical parameters ($\bar{\sigma}_{EV3}, \bar{\epsilon}_{EV3}, FR_{EV3}$ and FB_{EV3} respectively) at the 20% displacement step. Results are presented as mean±standard deviations. Displacement step size is 2% with minimum displacement of 2% and maximum displacement of 20%.

$0.4 \pm 0.2\%, p \leq 0.01$). However, the first derivative of the failure factors with respect to the displacement (*FR*: FtsZ1-2: 0.04, FtsZ2-1: 0.05 and *FB*: FtsZ1-2: 0.00, FtsZ2-1: 0.00) shows that with increasing displacement, *FR* would presumably become the dominating failure factor.

3.3. Isoform classification based on structural features

The receiver operating characteristics (ROC) curve and the area under the ROC curve (AUC) for each of the folds during training and the test set is shown in Fig. 6a. The model classification metrics on the test set are: $F1-\text{score} = 0.89, AUC = 0.91$ and accuracy of 6 out of 7 correct predictions. A correctly classified FtsZ2-1 isoform and a correctly classified FtsZ1-2 isoform, as well as the wrongly classified isoform (FtsZ1-2) are depicted in Fig. 6b, respectively. Furthermore, the gradient boosting model outperformed the random forest classifier with higher classification metrics on the test set (Table 3).

Analyzing the importance of each of the structural features in the classification model reveals which of the features contribute most and which least in terms of classifying isoform-inherent structural properties. The five most important structural features are the node-surface to node-center ratio (11%), network stretch (10%), largest diameter (9%), node density (8%) and the node-surface distance (7%). These have in total 44% of the overall importance in the classification model (Fig. 6d). Interestingly, three of these five are nodal features of the network and two represent the overall morphology.

3.4. Mechanical behavior prediction based on structural features

The trained surrogate models predicted the mechanical response (simulation results) for small (2%) and large (20%) deformations purely based on structural features. Despite the high correlations between simulated and predicted mechanical parameters in both small and large deformations (Table 3), the performance metrics of the trained surrogate models increase with advancing compression (Fig. 7a-h). This specifically holds for predicting failure factors (*FR* in $\alpha = 0.02 \rightarrow 0.20 : R^2 = 0.96 \rightarrow 0.97, MPE = 10\% \rightarrow 5\%$ and $\hat{y} = 0.95 \rightarrow 1.03$; *FB* in $\alpha = 0.02 \rightarrow 0.20 : R^2 = 0.81 \rightarrow 0.95, MPE = 11\% \rightarrow 1\%$ and $\hat{y} = 0.91 \rightarrow 1.00$). Moreover, comparing the R^2 of surrogate models for all four predicted parameters (*SMs*) in case of small and large deformation shows considerable higher predictive capability of the trained gradient boosting models than the random forest models (Table 4).

The structural features show different importance in predicting the mechanical parameters in case of large deformations. For both surrogate models predicting the mean stresses and strain, most important features (more than 70% of total importance) are net-

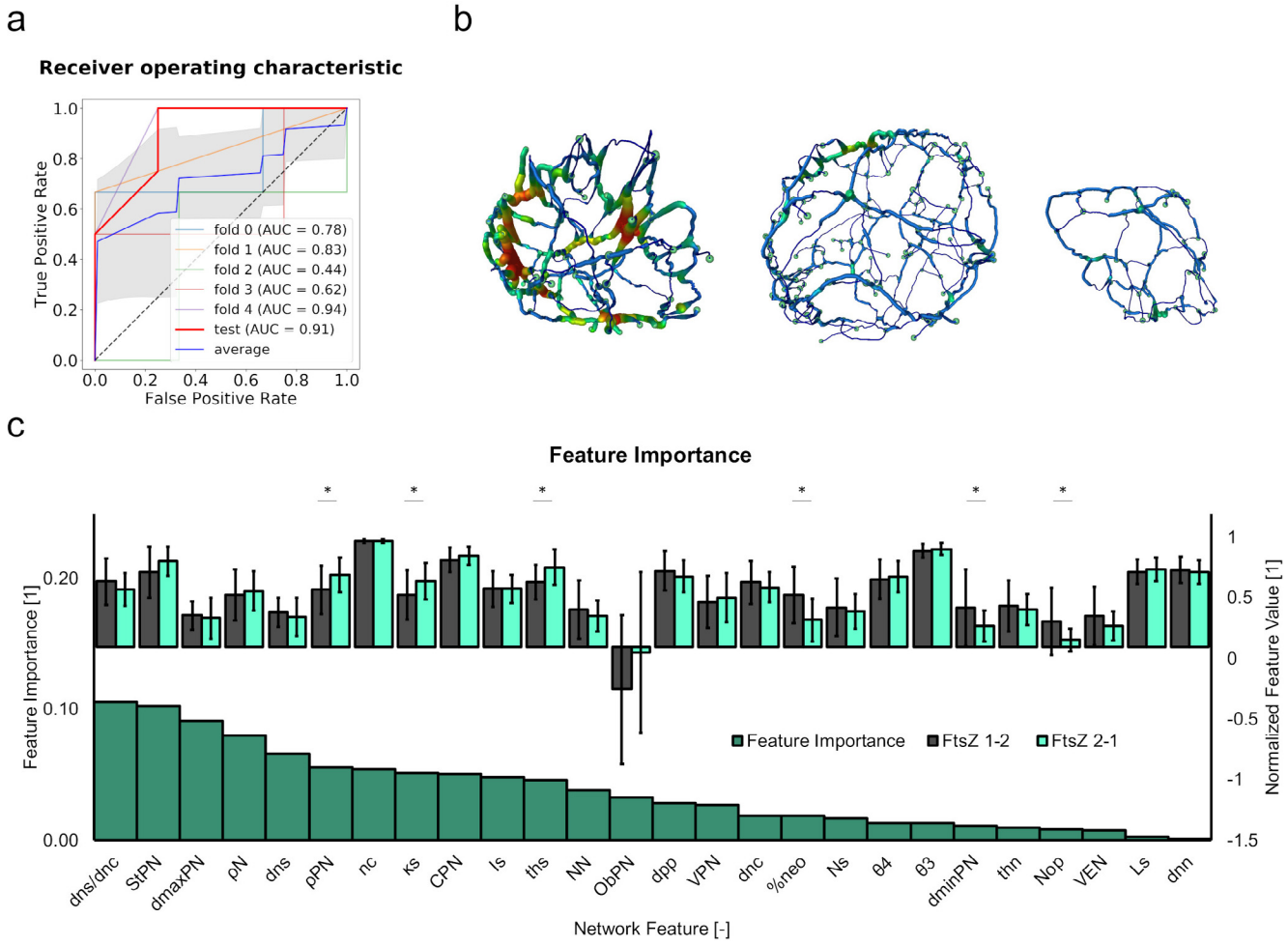

Fig. 6. Classification of isoforms. a) Receiver operating characteristics of the trained classification model for 5-fold cross validation and the test set. The blue line and the gray area mark the average performance of the model on 5 folds and the standard deviation, respectively. b) Sample spatial graphs of a correctly classified FtsZ2-1 (left), a correctly classified FtsZ1-2 (middle) and a wrongly classified FtsZ1-2 (right). c) Feature importance for the classification model as well as normalized feature values (normalized to maximum of each feature). Data shown as mean ± standard deviation. * indicates a significant difference between isoform (un-paired students' t-test).

Table 3
Classification metrics of the 5 folds and the test set for the gradient boosting (GB) compared to the model random forest (RF) models.

	Gradient boosting			Random forest		
Step	Accuracy	F1	AUC	Accuracy	F1	AUC
training fold 1	0.5	0.40	0.67	0.67	0.5	0.83
training fold 2	0.67	0.50	0.83	0.67	0.5	0.44
training fold 3	0.67	0.57	0.44	0.67	0.67	0.67
training fold 4	0.67	0.40	0.75	0.67	0.5	0.75
training fold 5	0.83	0.67	0.94	0.83	0.8	1.0
test	0.88	0.89	0.91	0.75	0.67	0.88

work density ($\rho_{PN}; \bar{\sigma}: 56\%$ and $\bar{\epsilon}: 50\%$) and node density ($\rho_N; \bar{\sigma}: 13\%$ and $\bar{\epsilon}: 22\%$; Fig. 7i, j). The prevailing structural features for predicting rupture failure factor are node density ($\rho_N: 15\%$), segment inhomogeneity ($I_S: 15\%$), network density ($\rho_{PN}: 14\%$) and percentage of open nodes ($\%n_{oe}: 9\%$; Fig. 7k) with cumulative importance of more than 50%. In case of predicting buckling failure factor, segment inhomogeneity ($I_S: 41\%$), node-center distance ($d_{nc}: 12\%$) and node density ($\rho_{PN}: 9\%$) are the most important features with cumulative importance of more than 60% (Fig. 7l).

4. Discussion

Structure and functionality of cytoskeletal protein networks are deeply linked. Although biochemical aspects have been thoroughly studied, little is known about the interplay between the structural characteristics of these networks with their mechanical functionality. Here we proposed a data-driven approach to investigate this structure-function relationship and presented the application to FtsZ protein network.

4.1. Mechanical response of FtsZ isoforms

We were able to show the mechanical response of a protein network to external loads and specifically that the precise structural response of FtsZ networks to compression is independent of the load direction and is different for FtsZ1-2 and FtsZ2-1 isoforms. This is to our knowledge the first detailed in silico investigation of the mechanical behaviour of cytoskeletal protein structures in response to external micro-environmental stimuli. This identified isoform-specific mechanical responses support the assumption of potentially different structural roles of these two main FtsZ isoforms [23]. Further, the isoform-specific mechanical responses are in accordance with the functional- and morphology-related observations of the same isoforms in yeast cells [24,20].

In case of mechanical response of FtsZ in large deformation, we focus only on one direction due to four reasons: 1) the similar mechanical response at 2% displacement for all loading cases, 2) according to calculations in [27], variations of the two significant different parameters for all three principal directions (FtsZ1-2: $\bar{\sigma} = 55\%$ and FtsZ2-1: $FB = 35\%$) can be explained by differences in stretch (FtsZ1-2 $St = 0.76 \pm 0.11$, FtsZ1-2 $St = 0.67 \pm 0.20, p = 0.05$), 3) the need of significant computational resources and 4) EV3 has the overall highest (combining FtsZ2-1 and FtsZ1-2) mean values at 2% plate displacement in all four parameters ($\bar{\sigma} = 1.7 \pm 1.3\ Pa; \bar{\epsilon} = 0.7 \pm 0.6\mathrm{e}-3; FR = 1.6 \pm 3\mathrm{e}-3; FB = 4.2 \pm 1.9\mathrm{e}-1$). The similarity of mechanical responses between the two isoforms suggests that they contribute in response to large chloroplast deformation in a comparable (or combined) fashion to the plastid mechanics. Moreover, FtsZ isoforms show a semi-nonlinear increase in stress and strain with an increase in network deformation. This is similar to previously reported behavior of microtubule [59,60] and actin filaments [61]. This points toward similar load-bearing functionality of FtsZ (as plastoskeleton). Furthermore, for compression up to 20%, buckling remains the prevailing failure factor. The convergence of *FB*, which reaches its limit at about 1%, suggests that the network minimizes the buckling probability, indicating an adaptive stability of FtsZ networks as previously suggested based on experimental observations [23,62]. Although the rupture failure probability steadily increases with increasing compression, it remains significantly below the bulking failure factor, thus rupture failure is a less defining parameter for network failure. One reason for this might be, that FtsZ filaments experiencing high strain values leading to rupture only after buckling and at the location of bucking, which would be similar to the fragmentation of buckled actin filaments [63].

Previous studies employing simplified geometries, such as tensegrity models allowed theoretical studies of cellular mechanism such as cell reorientation [12,16]. More detailed FE models have been developed to investigate the mechanical role of cytoskeletal components [14] and cell mechano-sensitivity [64]. However, the generic and strongly simplified geometrical representation of the cells, e.g ellipsoid [65,66] (even when cytoskeleton filament directions were considered), potentially prevents comprehensive studies of the influence of structural features on the mechanical behavior of cytoskeletal protein networks. Our approach of performing nanoFE simulations on segmented 3D network geometries of life networks allows one to analyse structure-related aspects of protein network mechanics in a sample specific manner. Further, investigating the sub-cellular components separated from their surrounding allows to decouple protein network mechanics from whole cell mechanics. To date, contributions of cytoskeletal structures to whole cellular mechanics can only be indirectly inferred from experimental techniques such as AFM [67] and optical tweezers [68]. However, since the mechanical behavior of a cellular structure is determined by many components, such as of the structure of the cortical, intra-cellular (non-cortical) cytoskeletal, and nuclear networks, as well as their distribution in space, decoupling the individual components remains challenging [14,69,70]. It has been suggested that AFM measurements with sharp tips tend to emphasize biomechanical properties of the cell cortex, whereas AFM measurements with round-tips tend to emphasize stiffness of the intra-cellular network [70]. Combing such measurements with a structural detailed models, as shown here, would possibly further advance the understanding of cellular mechanics.

4.2. Structure-function relationship in FtsZ network

Our mechanical surrogate model is capable of predicting the mechanical behavior of protein networks in response to external loading. This holds for 2% as well as 20% compression ($0.81 \leqslant R^2 \leqslant 0.96$ and $0.93 \leqslant R^2 \leqslant 0.97$, respectively). However, the relatively higher accuracy for large network compression points toward higher correspondence of the extracted structural features to the mechanical behavior of the network in response to large deformations. This specifically coheres to the hypothesis that these networks are able to undergo large deformations with-

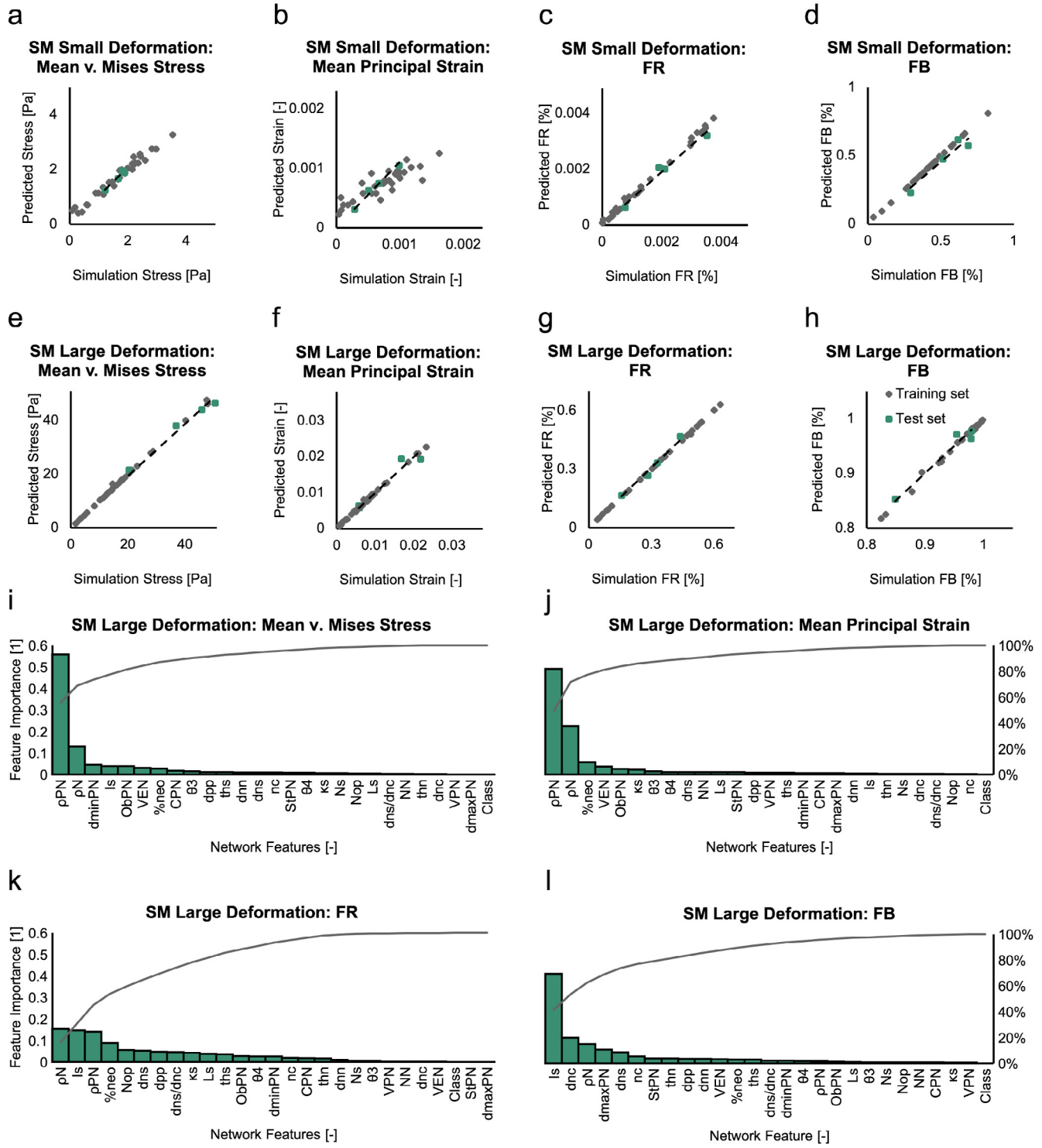

Fig. 7. Surrogate mechanical models. a–d) Simulation results vs. surrogate model prediction for the test and training sets networks for small deformation in *EV*3 primary direction. a) $\bar{\sigma}$. b) $\bar{\epsilon}$. c) *FR*. d) *FB*. e–h) Simulation results vs. Surrogate model prediction for the test and training set networks for large deformation in *EV*3 primary direction. e) $\bar{\sigma}$. f) $\bar{\epsilon}$. g) *FR*. h) *FB*. Training and test datasets are shown in gray and green, respectively. Dashed line represents a liner fit to the data points. i–l) Importance of structural features for the set of surrogate model predicting each mechanical parameter, i) $\bar{\sigma}$. j) $\bar{\epsilon}$. k) *FR*. l) *FB*. The gray line represents cumulative importance. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

out losing their structural integrity, as previously postulated [23]; hence, possessing structural features conforming to response in case of large deformations. Moreover, the high accuracy in mapping the structural features to the mechanical behavior of the networks further demonstrates the potential load bearing functionality of FtsZ protein in chloroplasts. Furthermore, this shows that the capability of the network to keep its stability by undergoing deformations relies not only on material properties of the biopolymer, but probably more prominently, on the structural features of the network. This is in accordance with the effects of the network architecture on the overall mechanical behavior reported in actin protein network [71]. In summary, to our knowl-

Table 4
Surrogate models performance in predicting calculated mechanical parameters in case of small ($\alpha = 0.02$) and large deformations ($\alpha = 0.20$) for the gradient boosting (GB) compared to the model random forest (RF) models.

	GB						RF	
	$\alpha = 0.02$			$\alpha = 0.2$			$\alpha = 0.02$	$\alpha = 0.2$
SM	Pred. R^2	*MPE*	Lin. fit $\hat{y}$	Pred. R^2	*MPE*	Lin. fit $\hat{y}$	Pred. R^2	Pred. R^2
$\bar{\sigma}$	0.90	4%	1.04	0.96	5%	0.97	0.39	−4.5
$\bar{\epsilon}$	0.89	14%	1.11	0.93	9%	0.99	0.14	−5.16
FR	0.96	10%	0.95	0.97	5%	1.03	−0.14	0.20
FB	0.81	11%	0.91	0.95	1%	1.00	0.66	−0.35

edge, this is the first detailed investigation of these sample-specific structure-based mechanical analysis of performance correlations. This enables us to not only have a image-based virtual mechanical testing method, but also a method to investigate the manifestation of the mechanical characteristics of structural network features. In addition, comparing the prediction performance metrics from the gradient boosting models to random forest models confirmed the superior effectiveness of the chosen classification and regression models.

By analyzing the importance of the features of the surrogate models in predicting stresses and strains of the network, we could show that the network density and the node density are the structural features mostly contributing to load bearing characteristic of the network. However, the mechanical failure behaviour of the networks is mostly corresponding to more local structural characteristics. Specifically, in the case of buckling failure behaviour, the surrogate model shows that local changes in the filament (segment inhomogeneity) and the distance of the nodes to the center dictate the observed mechanical behaviour. This can be interpreted as the network being capable of stopping the increase in failure possibility due to buckling of its filaments by possessing an arrangement of the nodes and filaments with the specific architecture that includes: local changes of direction and thickness of filaments in FtsZ1-2 and FtsZ2-1: $I_S = 18.8$ and the distance between the nodes and the center of the network: $d_{nc} = 1.85$ µm. This could potentially be used to design adaptively stable structures capable of undergoing large deformations [72] or mechanically optimized and synthetically engineered biomaterials [73,74].

Our feature-based classification and regression models achieved on par accuracy with deep learning based protein network classification methods [75,76], while adding the ability of extracting specific structural features enabling the predictions. This is specifically beneficial in the context of mechanical functionality, where our designed extracted features correspond to mechanical traits in their nature. Whereas, extracted features utilizing convolutional networks tend to be abstract and at best very challenging to interpret [77,30]. Moreover, unlike deep learning models that require extremely large dataset for meaningful training, our models could reach high accuracy on $n = 37$, 3D CLSM images.

4.3. Outlook

Here, we could show, that besides the biochemical properties and dynamic behaviour of the two FtsZ isoforms being different [78], also the mechanical function encoded in the network structure seems to be different between the two isoforms. Since the two FtsZs represent gene duplication during chloroplast evolution from cyanobacteria to algae and further [78–80], combined structural and functional analysis focusing on mechanical aspects of FtsZ, as introduced here, might open new doors to gain deeper insight into chloroplast evolution and structural consequences of the gene duplication of FtsZ in eukaryotic forms. Unlike FtsZ found in eukaryotic plastids, FtsZ in bacteria does not have these two phylogenetically distinct families of FtsZ, which co-localize at the division site to form the Z-ring. As the function of FtsZ in bacteria is to date also not completely revealed, application of the method presented here may allow to identify similarities and differences between eukaryotic and prokaryotic FtsZ and may allow studying cytoskeletal evolution from prokaryotes to eukaryotes and to elucidate common structural mechanisms involved in division. This may also allow insights into FtsZ filament dynamics and force generation in bacterial cells [81].

FtsZ being an ancestor of tubulin and the homology between these two [82,83], might further allow comprehending structural and functional aspects of microtubule by understanding mechanism of FtsZ dynamics. The structural evidence obtained from studies pointed out that ring polymers are homologous across tubulin family [84]. As functions of these rings are expected to be found in disassembly and force generation due to shifting from one conformation to the other, a transfer of knowledge gained from the simpler FstZ or an application of the method presented here to microtubule will might allow to further elucidate these mechanical functions. In line with this, the introduced automatic analysis of subcellular structure-function relationships can be potentially utilized to perform in silico investigations of the effects of structural alterations on subcellular mechanical functions. For instance, treatment of various cancers e.g. breast and prostate cancers are carried out using agents influencing microtubule stability and promoting microtubule assembly [85–87]. The similarity between FtsZ and microtubule as well as the capability of our method to analyze structural stability of the protein networks can be potentially utilized to investigate or monitor treatment-induced changes in stability of microtubule proteins. Another possible application is analyzing the reported structural changes of cytoskeletal proteins correlated to EMT. EMT-induced alterations in cytoskeletal protein networks such as irregular and flattened morphology of cancer-associated-fibroblasts [88] and increase in compactness of fibroblast membrane leading to higher cell membrane fluidity [89,88] can potentially be detected by the designed structural and morphological features i.e. stretch and oblateness of the cell cytoskeleton as well as compactness of the cytoskeletal protein in fibroblasts.

4.4. Limitations

Our study has limitations. First, the imaging resolution might affect the simulation results as well as the mapping of the surrogate models. However, we have previously shown that our quantitative imaging method is capable of resolving the micro-structure of FtsZ networks [28]. Second, the commonly used linear elastic material model in FE simulations of cytoskeletons [14,15,64,90] might not completely capture the mechanical behavior of the network. However, to our knowledge, to date no constitutive law has been developed for describing mechanical behavior of FtsZ. Although more complicated constitutive laws for the mechanical

behavior of single actin filaments have been proposed [91–93], linear elasticity is the prevailing choice for cytoskeletal networks in whole-cell models [14,15,64,90]. Moreover, our focus was not on exactly matching the mechanical behaviour, but on the influences of structural features. Future studies could focus on combining our approach of precisely modeling the micro-structure with experimental techniques, such as atomic force microscopy, to further investigate material properties of the FtsZ-based plastoskeleton. Third, the loading conditions of our simulations are not an exact duplication of reality, where a combination of active dynamic forces [44] as well as osmotic pressure [94] drive the morphological changes of the network. Moreover, the FtsZ isoform is surrounded by other proteins as well as other materials, such as inter-organelle fluids. Our designed simulation setup provides a generic platform to investigate the structure-function relationships in FtsZ protein network rather than a one-to-one simulation of dynamics of plastids. The failure criteria used in this study are experimentally derived from actin filaments [95,48], since no failure criteria has been experimentally derived for FtsZ to date. However, due to the assumed similarity in structural functionality between the FtsZ network and actin networks and the similarity of rigidity in FtsZ and actin filaments, actin failure criteria might represent FtsZ behavior to a certain extent. Finally, the dataset size might restrict the generality of conclusions made by means of the introduced data-driven approach. Although the in silico experimental setup increases the dataset size, more images would further improve the performance of the designed methodology. Future studies could benefit from synthetically generated network geometries to augment the training data, which has the potential to further increase the robustness of the developed models.

5. Conclusions

In this work, we showed that combing confocal microscopy imaging with nanoFE analysis employing a machine learning framework allows for an image-based surrogate model capable of predicting cellular mechanical responses to external stimuli. Additionally, by providing a way to identify structural features determining the mechanical response with respect to a given stimulus, we were, for the first time, able to directly investigate the structure-function relationship of individual protein networks in a sample-specific manner. Our ML surrogate model trained on in silico data generates highly accurate and fast predictions of isoform classification and the mechanical behavior on the sub-cellular level. Therefore, the method provides a framework to further investigate structural functionality of protein networks in plants as well as in humans, as it would allow to monitor the structure-function relationships of cytoskeletal components during morphological and, hence, time-dependent, changes, e.g., actin-driven cell shaping. This may also lead to an improved understanding of the mechanical aspects of cell-biomaterial interaction, and would provide insights into designing micro/nanoengineered functional biomaterials for future research on regenerative medicine, cell biology, development and diseases, as well as drug development.

Conflict of interest statement

All authors declare no conflicting interests.

Data and code availability

Codes for the pre-processing and training the models as well as the trained classification and mechanical surrogate models are available at: https://github.com/pazadeh/FEML.

CRediT authorship contribution statement

Pouyan Asgharzadeh: Conceptualization, Writing - original draft, Writing - review & editing. **Annette I. Birkhold:** Conceptualization, Writing - original draft, Writing - review & editing. **Zubin Trivedi:** Writing - review & editing. **Bugra Özdemir:** Writing - review & editing. **Ralf Reski:** Conceptualization, Writing - review & editing, Funding acquisition. **Oliver Röhrle:** Conceptualization, Writing - review & editing, Funding acquisition.

Acknowledgements

This study was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2075–390740016 (SimTech) and EXC 2189 (CIBSS), and as a part of Transregio/SFB TRR141, project A09.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, athttps://doi.org/10.1016/j.csbj.2020.09.024.

References

[1] Jordan MA, Wilson L. Microtubules as a target for anticancer drugs. Nat Rev Cancer 2004;4:253.
[2] Parent CA. Making all the right moves: chemotaxis in neutrophils and dictyostelium. Curr Opin Cell Biol 2004;16:4–13.
[3] Flitney EW, Kuczmarski ER, Adam SA, Goldman RD. Insights into the mechanical properties of epithelial cells: the effects of shear stress on the assembly and remodeling of keratin intermediate filaments. FASEB J 2009;23:2110–9.
[4] Betz O, Birkhold A, Caliaro M, Eggs B, Mader A, Knippers J, Röhrle O, Speck O. Adaptive stiffness and joint-free kinematics: actively actuated rod-shaped structures in plants and animals and their biomimetic potential in architecture and engineering. Springer International Publishing: Cham. pp. 135–167. .
[5] Guck J, Schinkinger S, Lincoln B, Wottawah F, Ebert S, Romeyke M, Lenz D, Erickson HM, Ananthakrishnan R, Mitchell D, et al. Optical deformability as an inherent cell marker for testing malignant transformation and metastatic competence. Biophys J 2005;88:3689–98.
[6] Suresh S, Spatz J, Mills J, Micoulet A, Dao M, Lim C, Beil M, Seufferlein T. Connections between single-cell biomechanics and human disease states: gastrointestinal cancer and malaria. Acta Biomater 2005;1:15–30.
[7] Mendez MG, Kojima S-I, Goldman RD. Vimentin induces changes in cell shape, motility, and adhesion during the epithelial to mesenchymal transition. FASEB J 2010;24:1838–51.
[8] Liu C-Y, Lin H-H, Tang M-J, Wang Y-K. Vimentin contributes to epithelial-mesenchymal transition cancer cell mechanics by mediating cytoskeletal organization and focal adhesion maturation. Oncotarget 2015;6:15966.
[9] Hendrick J, Franz-Wachtel M, Moeller Y, Schmid S, Macek B, Olayioye MA. The polarity protein scribble positions dlc3 at adherens junctions to regulate rho signaling. J Cell Sci 2016;129:3583–96.
[10] Ketene AN, Schmelz EM, Roberts PC, Agah M. The effects of cancer progression on the viscoelasticity of ovarian cell cytoskeleton structures. Nanomed Nanotechnol Biol Med 2012;8:93–102.
[11] Fletcher DA, Mullins RD. Cell mechanics and the cytoskeleton. Nature 2010;463:485.
[12] Ingber DE. Tensegrity: the architectural basis of cellular mechanotransduction. Annu Rev Physiol 1997;59:575–99.
[13] Milner JS, Grol MW, Beaucage KL, Dixon SJ, Holdsworth DW. Finite-element modeling of viscoelastic cells during high-frequency cyclic strain. J Funct Biomater 2012;3:209–24.
[14] Barreto S, Clausen CH, Perrault CM, Fletcher DA, Lacroix D. A multi-structural single cell model of force-induced interactions of cytoskeletal components. Biomaterials 2013;34:6119–26.
[15] Katti DR, Katti KS. Cancer cell mechanics with altered cytoskeletal behavior and substrate effects: a 3d finite element modeling study. J Mech Behav Biomed Mater 2017;76:125–34.
[16] Xu G-K, Li B, Feng X-Q, Gao H. A tensegrity model of cell reorientation on cyclically stretched substrates. Biophys J 2016;111:1478–86.
[17] Reski R. Rings and networks: the amazing complexity of ftsz in chloroplasts. Trends Plant Sci 2002;7:103–5.
[18] Adams DW, Errington J. Bacterial cell division: assembly, maintenance and disassembly of the z ring. Nat Rev Microbiol 2009;7:642.
[19] Osteryoung KW, Pyke KA. Division and dynamic morphology of plastids. Annu Rev Plant Biol 2014;65:443–72.

[20] TerBush AD, MacCready JS, Chen C, Ducat DC, Osteryoung KW. Conserved dynamics of chloroplast cytoskeletal ftsz proteins across photosynthetic lineages. Plant Physiol 2018;176:295–306.
[21] Zimmer AD, Lang D, Buchta K, Rombauts S, Nishiyama T, Hasebe M, Van de Peer Y, Rensing SA, Reski R. Reannotation and extended community resources for the genome of the non-seed plant physcomitrella patens provide insights into the evolution of plant gene structures and functions. BMC Genomics 2013;14:498.
[22] Lang D, Van Gessel N, Ullrich K, Reski R. The genome of the model moss physcomitrella patens. In: Advances in botanical research. vol. 78. Elsevier; 2016. pp. 97–140. .
[23] Martin A, Lang D, Hanke ST, Mueller SJ, Sarnighausen E, Vervliet-Scheebaum M, Reski R. Targeted gene knockouts reveal overlapping functions of the five physcomitrella patens ftsz isoforms in chloroplast division, chloroplast shaping, cell patterning, plant development, and gravity sensing. Mol Plant 2009;2:1359–72.
[24] TerBush AD, Osteryoung KW. Distinct functions of chloroplast ftsz1 and ftsz2 in z-ring structure and remodeling. J Cell Biol 2012;199:623–37.
[25] Osteryoung KW. Conserved cell and organelle division. Nature 1995;376:473–4.
[26] TerBush AD, Porzondek CA, Osteryoung KW. Functional analysis of the chloroplast division complex using schizosaccharomyces pombe as a heterologous expression system. Microsc Microanal 2016;22:275–89.
[27] Asgharzadeh P, Özdemir B, Reski R, Röhrle O, Birkhold AI. Computational 3d imaging to quantify structural components and assembly of protein networks. Acta Biomater 2018;69:206–17.
[28] Özdemir B, Asgharzadeh P, Birkhold AI, Mueller SJ, Röhrle O, Reski R. Cytological analysis and structural quantification of ftsz1-2 and ftsz2-1 network characteristics in physcomitrella patens. Sci Rep 2018;8:11165.
[29] Kan A. Machine learning applications in cell image analysis. Immunol Cell Biol 2017;95:525–30.
[30] Asgharzadeh P, Röhrle O, Willie BM, Birkhold AI. Decoding rejuvenating effects of mechanical loading on skeletal aging using in vivo μ ct imaging and deep learning. Acta Biomater 2020.
[31] Masso M, Vaisman II. Accurate prediction of stability changes in protein mutants by combining machine learning with structure based computational mutagenesis. Bioinformatics 2008;24:2002–9.
[32] Baldi P, Pollastri G. A machine learning strategy for protein analysis. IEEE Intell Syst 2002;17:28–35.
[33] Fritzen F, Fernández M, Larsson F. On-the-fly adaptivity for nonlinear twoscale simulations using artificial neural networks and reduced order modeling. Front Mater 2019;6:75.
[34] Evans JD, Coudert F-X. Predicting the mechanical properties of zeolite frameworks by machine learning. Chem Mater 2017;29:7833–9.
[35] Hohe A, Reski R. Optimisation of a bioreactor culture of the moss physcomitrella patens for mass production of protoplasts. Plant Sci 2002;163:69–74.
[36] Kircher S, Wellmer F, Nick P, Rügner A, Schäfer E, Harter K. Nuclear import of the parsley bzip transcription factor cprf2 is regulated by phytochrome photoreceptors. J Cell Biol 1999;144:201–11.
[37] Hohe A, Egener T, Lucht JM, Holtorf H, Reinhard C, Schween G, Reski R. An improved and highly standardised transformation procedure allows efficient production of single and multiple targeted gene-knockouts in a moss, physcomitrella patens. Curr Genet 2004;44:339–47.
[38] Ofek G, Wiltz DC, Athanasiou KA. Contribution of the cytoskeleton to the compressive properties and recovery behavior of single cells. Biophys J 2009;97:1873–82.
[39] Zilske M, Lamecker H, Zachow S. Adaptive remeshing of non-manifold surfaces. Proc Eurogr 2008;27.
[40] Gittes F, Mickey B, Nettleton J, Howard J. Flexural rigidity of microtubules and actin filaments measured from thermal fluctuations in shape. J Cell Biol 1993;120:923–34.
[41] Felgner H, Frank R, Schliwa M. Flexural rigidity of microtubules measured with the use of optical tweezers. J Cell Sci 1996;109:509–16.
[42] Felgner H, Frank R, Biernat J, Mandelkow E-M, Mandelkow E, Ludin B, Matus A, Schliwa M. Domains of neuronal microtubule-associated proteins and flexural rigidity of microtubules. J Cell Biol 1997;138:1067–75.
[43] Landau LD, Lifshitz EM, Berestetskii V, Pitaevskii L. Course of theoretical physics. Theory Elasticity 1995.
[44] Turner DJ, Portman I, Dafforn TR, Rodger A, Roper DI, Smith CJ, Turner MS. The mechanics of ftsz fibers. Biophys J 2012;102:731–8.
[45] Arlot S, Celisse A, et al. A survey of cross-validation procedures for model selection. Stat Surveys 2010;4:40–79.
[46] Erickson HP, Taylor DW, Taylor KA, Bramhill D. Bacterial cell division protein ftsz assembles into protofilament sheets and minirings, structural homologs of tubulin polymers. Proc Natl Acad Sci 1996;93:519–23.
[47] Kojima H, Ishijima A, Yanagida T. Direct measurement of stiffness of single actin filaments with and without tropomyosin by in vitro nanomanipulation. Proc Natl Acad Sci 1994;91:12962–6.
[48] Rajagopal V, Holmes WR, Lee PVS. Computational modeling of single-cell mechanics and cytoskeletal mechanobiology. Wiley Interdiscip Rev Syst Biol Med 2018;10: e1407.
[49] Erickson HP, Osawa M. Ftsz constriction force–curved protofilaments bending membranes. In Prokaryotic cytoskeletons. Springer; 2017. pp. 139–160. .
[50] Pistoia W, Van Rietbergen B, Lochmüller E-M, Lill C, Eckstein F, Rüegsegger P. Estimation of distal radius failure load with micro-finite element analysis models based on three-dimensional peripheral quantitative computed tomography images. Bone 2002;30:842–8.
[51] Kubitschke H, Schnauss J, Nnetu KD, Warmt E, Stange R, Kaes J. Actin and microtubule networks contribute differently to cell response for small and large strains. New J Phys 2017;19: 093003.
[52] Janmey PA, Euteneuer U, Traub P, Schliwa M. Viscoelastic properties of vimentin compared with other filamentous biopolymer networks. J Cell Biol 1991;113:155–60.
[53] Friedman JH. Greedy function approximation: a gradient boosting machine. Ann Stat 2001:1189–232.
[54] Freund Y. Boosting a weak learning algorithm by majority. Inf Comput 1995;121:256–85.
[55] Picard RR, Cook RD. Cross-validation of regression models. J Am Stat Assoc 1984;79:575–83.
[56] Breiman L. Random forests. Mach Learn 2001;45:5–32.
[57] Friedman JH. Stochastic gradient boosting. Comput Stat Data Anal 2002;38:367–78.
[58] James G, Witten D, Hastie T, Tibshirani R. An introduction to statistical learning, vol. 112. Springer; 2013.
[59] Lin Y-C, Koenderink GH, MacKintosh FC, Weitz DA. Viscoelastic properties of microtubule networks. Macromolecules 2007;40:7714–20.
[60] Yang Y, Bai M, Klug WS, Levine AJ, Valentine MT. Microrheology of highly crosslinked microtubule networks is dominated by force-induced crosslinker unbinding. Soft Matter 2013;9:383–93.
[61] Gardel ML, Kasza KE, Brangwynne CP, Liu J, Weitz DA. Mechanical response of cytoskeletal networks. Methods Cell Biol 2008;89:487–519.
[62] Asgharzadeh P, Özdemir B, Müller SJ, Röhrle O, Reski R. Analysis of physcomitrella chloroplasts to reveal adaptation principles leading to structural stability at the nano-scale. In Biomimetic research for architecture and building construction. Springer; 2016. pp. 261–275. .
[63] Enrique M, Martiel J-L, Blanchoin L. Mechanical heterogeneity favors fragmentation of strained actin filaments. Biophys J 2015;108:2270–81.
[64] Yang X, Sun L-W, Du C-F, Wu X-T, Fan Y-B. Finite element analysis of osteocytes mechanosensitivity under simulated microgravity. Microgravity Sci Technol 2018;1–13.
[65] Gladilin E, Gonzalez P, Eils R. Dissecting the contribution of actin and vimentin intermediate filaments to mechanical phenotype of suspended cells using high-throughput deformability measurements and computational modeling. J Biomech 2014;47:2598–605.
[66] Pivkin IV, Peng Z, Karniadakis GE, Buffet PA, Dao M, Suresh S. Biomechanics of red blood cells in human spleen and consequences for physiology and disease. Proc Natl Acad Sci 2016;113:7804–9.
[67] Li M, Xi N, Wang Y, Liu L. Advances in atomic force microscopy for single-cell analysis. Nano Res 2019;12:703–18.
[68] Dao M, Lim CT, Suresh S. Mechanics of the human red blood cell deformed by optical tweezers. J Mech Phys Solids 2003;51:2259–80.
[69] Curry N, Ghézali G, Kaminski Schierle GS, Rouach N, Kaminski CF. Correlative sted and atomic force microscopy on live astrocytes reveals plasticity of cytoskeletal structure and membrane physical properties during polarized migration. Front Cell Neurosci 2017;11:104.
[70] Vahabikashi A, Park CY, Perkumas K, Zhang Z, Deurloo EK, Wu H, Weitz DA, Stamer WD, Goldman RD, Fredberg JJ, et al. Probe sensitivity to cortical versus intracellular cytoskeletal network stiffness. Biophys J 2019;116:518–29.
[71] Enrique M, Gardel ML. Actin mechanics and fragmentation. J Biol Chem 2015;290:17137–44.
[72] Reski R, Özdemir B, Asgharzadeh P, Birkhold A, Röhrle O. The plastid skeleton: a source of ideas in the nano range, in: Biomimetics for Architecture. Learning from Nature, Birkhäuser, 2019. pp. 163–166. .
[73] Li X, Cui R, Sun L, Aifantis KE, Fan Y, Feng Q, Cui F, Watari F. 3d-printed biopolymers for tissue eng. application. Int Natl J (Wash) Polym Sci 2014.
[74] Liu X, Yuk H, Lin S, Parada GA, Tang T-C, Tham E, de la Fuente-Nunez C, Lu TK, Zhao X. 3d printing of living responsive materials and devices. Adv Mater 2018;30:1704821.
[75] Kraus OZ, Grys BT, Ba J, Chong Y, Frey BJ, Boone C, Andrews BJ. Automated analysis of high-content microscopy data with deep learning. Mol Syst Biol 2017;13:924.
[76] Pärnamaa T, Parts L. Accurate classification of protein subcellular localization from high-throughput microscopy images using deep learning. Genes Genomes Genet 2017;7:1385–92.
[77] Yosinski J, Clune J, Nguyen A, Fuchs T, Lipson H. Understanding neural networks through deep visualization, arXiv preprint arXiv:1506.06579 (2015). .
[78] Osteryoung KW, Stokes KD, Rutherford SM, Percival AL, Lee WY. Chloroplast division in higher plants requires members of two functionally divergent gene families with homology to bacterial ftsz. Plant Cell 1998;10:1991–2004.
[79] Osteryoung KW, McAndrew RS. The plastid division machine. Annu Rev Plant Biol 2001;52:315–33.
[80] Miyagishima S-Y, Nozaki H, Nishida K, Nishida K, Matsuzaki M, Kuroiwa T. Two types of ftsz proteins in mitochondria and red-lineage chloroplasts: the duplication of ftsz is implicated in endosymbiosis. J Mol Evol 2004;58:291–303.
[81] Mateos-Gil P, Tarazona P, Vélez M. Bacterial cell division: modeling ftsz assembly and force generation from single filament experimental data. FEMS Microbiol Rev 2019;43:73–87.
[82] Erickson HP. The ftsz protofilament and attachment of zipa–structural constraints on the ftsz power stroke. Curr Opin Cell Biol 2001;13:55–60.

[83] van den Ent F, Amos L, Löwe J. Bacterial ancestry of actin and tubulin. Curr Opin Microbiol 2001;4:634–8.
[84] Erickson HP, Stoffler D. Protofilaments and rings, two conformations of the tubulin family conserved from bacterial ftsz to alpha/beta and gamma tubulin. J Cell Biol 1996;135:5–8.
[85] Battaje RR, Panda D. Lessons from bacterial homolog of tubulin, ftsz for microtubule dynamics. Endocr Relat Cancer 2017;24:T1–T21.
[86] Singh P, Rathinasamy K, Mohan R, Panda D. Microtubule assembly dynamics: an attractive target for anticancer drugs. IUBMB life 2008;60:368–75.
[87] Field JJ, Kanakkanthara A, Miller JH. Microtubule-targeting agents are clinically successful due to both mitotic and interphase impairment of microtubule function. Bioorg Med Chem 2014;22:5050–9.
[88] Angelucci C, Maulucci G, Lama G, Proietti G, Colabianchi A, Papi M, Maiorana A, De Spirito M, Micera A, Balzamino OB, et al. Epithelial-stromal interactions in human breast cancer: effects on adhesion, plasma membrane fluidity and migration speed and directness. PLoS One 2012;7: e50804.
[89] Sok M, Šentjurc M, Schara M, Stare J, Rott T. Cell membrane fluidity and prognosis of lung cancer. Ann Thorac Surg 2002;73:1567–71.
[90] Deguchi S, Ohashi T, Sato M. Tensile properties of single stress fibers isolated from cultured vascular smooth muscle cells. J Biomech 2006;39:2603–10.
[91] Holzapfel GA, Ogden RW. On the bending and stretching elasticity of biopolymer filaments. J Elast 2011;104:319–42.
[92] Blundell J, Terentjev E. Stretching semiflexible filaments and their networks. Macromolecules 2009;42:5388–94.
[93] Unterberger MJ, Schmoller KM, Bausch AR, Holzapfel GA. A new approach to model cross-linked actin networks: multi-scale continuum formulation and computational analysis. J Mech Behav Biomed Mater 2013;22:95–114.
[94] Glynn JM, Miyagishima S-Y, Yoder DW, Osteryoung KW, Vitha S. Chloroplast division. Traffic 2007;8:451–61.
[95] Wang Y, Qian J. Buckling of filamentous actin bundles in filopodial protrusions. Acta Mech Sin 2019;1–11.