Biophysical Journal, Volume *115*

Supplemental Information

Thermodynamic Modeling of the Statistics of Cell Spreading on Ligand-Coated Elastic Substrates

Eoin McEvoy, Siamak S. Shishvan, Vikram S. Deshpande, and J. Patrick McGarry

SUPPLEMENTARY INFORMATION

1. The homeostatic statistical mechanics framework

Here, we provide a brief overview of the homeostatic mechanics framework of Shishvan *et al.* (1) with the aim to provide the reader the key aspects of the framework required for fully appreciating the computational results presented in the main text. Readers are referred to Shishvan *et al.* (1) for a more complete treatment including the derivations of the relevant equations.

Making the ansatz that *living cells are entropic*, Shishvan *et al*. (1) introduced the concept of the *homeostatic ensemble* with cellular homeostasis providing the additional constraints and mechanisms for entropy maximisation. This defined the notion of a (dynamic) *homeostatic equilibrium* state that intervenes to allow living cells to elude thermodynamic equilibrium. They thus developed a statistical mechanics framework for living cells using the notions of statistical inference (2) applicable over a timescale from a few hours to a few days as long as the cell remains as a single undivided entity (i.e. the interphase period of the cell cycle). The key ideas behind the framework can be summarised as follows. A system comprising the cell and the extracellular matrix (ECM) is an open system with the cell exchanging nutrients with the surrounding bath. These nutrients fuel a large number of coupled biochemical reactions that include actin polymerisation, treadmilling and dendritic nucleation that effect changes to the cell morphology. These biochemical reactions change the morphology of the cell but are not precisely controlled, and this manifests via the observed morphological fluctuations of the cell. Shishvan *et al*. (1) made the *ansatz* that these biochemical reactions provide the mechanisms to maximise the morphological entropy of the cell, but constrained by the fact that the cell maintains a homeostatic state over the interphase period. Cellular homeostasis is the ability of cells to actively regulate their internal state, and maintain the concentration of all internal species[1] at specific average values over their morphological fluctuations independent of the environment.

[1] Chemical species here are defined in a manner analogous to the Gibbs definition for a grand canonical ensemble, viz. chemical species are an ensemble of chemically identical molecular entities that can explore the same set of molecular energy levels on the timescale of a morphological microstate.

1.1. Morphological microstates, entropy, fluctuations, and the homeostatic temperature

Controlling only macro variables (i.e. macrostate) such as the temperature, pressure, and nutrient concentrations in the nutrient bath results in inherent uncertainty (referred to here as missing information) in micro variables (i.e. microstates) of the system. This includes a level of unpredictability in homeostatic process variables, such as the spatio-temporal distribution of chemical species, that is linked to Brownian motion and the complex feedback loops in the homeostatic processes. Thus, this system not only includes the usual lack of precise information on the positions and velocities of individual molecules associated with the thermodynamic temperature, but also an uncertainty in cell shape resulting from imprecise regulation of the homeostatic processes. The consequent entropy production forms the basis of this new statistical mechanics framework motivated by the following two levels of microstates:

(i) *Molecular microstates*. Each molecular microstate has a specific configuration (position and momentum) of all molecules within the system.

(ii) *Morphological microstates*. Each morphological microstate is specified by the mapping (connection) of material points on the cell membrane to material points on the collagen coated substrate. In broad terms, a morphological microstate specifies the shape and size of the cell.

Shishvan *et al*. (1) identified the (dynamic) homeostatic or equilibrium state of the system by entropy maximisation. Subsequently, we shall simply refer to this state as an equilibrium state to emphasise that it is a stationary macrostate of the system inferred via entropy maximisation as in conventional equilibrium analysis. The total entropy of the system is written in terms of the conditional probability $P^{(i|j)}$ of the molecular microstate (i) given the morphological microstate (j) and the probability $P^{(j)}$ of morphological microstate (j) as

$$I_{\mathrm{T}} = \sum_{j} P^{(j)} I_{\mathrm{M}}^{(j)} + I_{\Gamma}. \tag{S1}$$

In eq. (S1), $I_{\mathrm{M}}^{(j)} \equiv -\sum_{i\in j} P^{(i|j)} \ln P^{(i|j)}$ and $I_{\Gamma} \equiv -\sum_{j} P^{(j)} \ln P^{(j)}$ are the entropies of molecular microstates in morphological microstate (j) and the morphological microstates, respectively. Equilibrium then corresponds to molecular and morphological macrostates that maximise I_{T} subject to appropriate constraints. The molecular macrostate evolves on the order of seconds, limited by processes such as the diffusion of unbound actin. By contrast, transformation of the morphological macrostate involves cell shape changes and therefore,

the morphological macrostate evolves on the order of minutes, limited by co-operative cytoskeletal processes within the cell such as meshwork actin polymerisation and dendritic nucleation. The evolutions of the molecular and morphological macrostates are therefore temporally decoupled, and Shishvan *et al*. (1) showed that eq. (S1) can be maximised by independently maximising $I_{\mathrm{M}}^{(j)}$ at the smaller timescale to determine the equilibrium distribution of molecular microstates for a given morphological microstate, and then maximising I_{Γ} at the larger timescale to determine the equilibrium distribution of the morphological microstates.

Over the (short) timescale on the order of seconds, the only known constraint on the system is that it is maintained at a constant temperature, pressure and strain distribution. The equilibrium of a given morphological microstate (j) obtained by maximising $I_{\mathrm{M}}^{(j)}$ (denoted by $S_{\mathrm{M}}^{(j)}$) corresponds to molecular arrangements that minimise the Gibbs free-energy with $G^{(j)}$. Since the connection between the cell and the collagen coated substrate is fixed for a given morphological microstate, the determination of $G^{(j)}$ is a standard boundary value problem as described in Section 2.2. Over the (long) timescale on the order of several minutes to hours, the equilibrium distribution $P_{\mathrm{eq}}^{(j)}$ is determined by maximising I_{Γ}, but now with the additional constraint that the cell is maintained in its homeostatic state. For the case of a cell on an ECM in a constant temperature and pressure nutrient bath, the homeostatic constraint translates to the fact that the average Gibbs free-energy of the system over all the morphological microstates it assumes, is equal to the equilibrium Gibbs free-energy G_{S} of an isolated cell in suspension (free-standing cell), i.e. the homeostatic processes maintain the average biochemical state of the system equal to that of a cell in suspension. In deriving this result, Shishvan *et al*. (1) did not consider every individual homeostatic process, but rather used just the coarse-grained outcome of the homeostatic processes. The application of this coarse-grained constraint is the key element of the *homeostatic mechanics* framework, with the morphological entropy I_{Γ} parameterising the information lost by not modelling all variables associated with the homeostatic processes.

The maximisation of I_{Γ} while enforcing $\sum_j P^{(j)} G^{(j)} = G_{\mathrm{S}}$ gives the *homeostatic equilibrium* state such that

$$P_{\mathrm{eq}}^{(j)} = \frac{1}{Z} \exp\left(-\zeta G^{(j)}\right), \tag{S2}$$

where $Z \equiv \sum_j \exp\left(-\zeta G^{(j)}\right)$ is the partition function of the morphological microstates, and the distribution parameter ζ follows from the homeostatic constraint

$$\frac{1}{Z}\sum_j G^{(j)} \exp\left(-\zeta G^{(j)}\right) = G_{\mathrm{S}}\,. \tag{S3}$$

The collection of all possible morphological microstates that the system assumes while maintaining its homeostatic equilibrium state is referred to as the *homeostatic ensemble*. The homeostatic ensemble can therefore be viewed as a large collection of copies of the system, each in one of the equilibrium morphological microstates. The copies (j) are distributed in the ensemble such that the free-energies $G^{(j)}$ follow an exponential distribution $P_{\mathrm{eq}}^{(j)}$ with the distribution parameter ζ.

The equilibrium morphological entropy $S_\Gamma = -\sum_j P_{\mathrm{eq}}^{(j)} \ln P_{\mathrm{eq}}^{(j)}$ (i.e. the maximum value of I_Γ) follows from (S2) as

$$S_\Gamma = \zeta G_{\mathrm{S}} + \ln Z, \tag{S4}$$

where $P_{\mathrm{eq}}^{(j)}$ is substituted from eq. (S2). Thus, S_Γ is related to ζ via the conjugate relation $\partial S_\Gamma/\partial G_{\mathrm{S}} = \zeta$. Thus, analogous to $1/T$ that quantifies the increase in uncertainty of the molecular microstates (i.e. molecular entropy $S_{\mathrm{M}}^{(j)}$) with average enthalpy, ζ specifies the increase in uncertainty of the morphological microstates (i.e. morphological entropy S_Γ) with the average Gibbs free-energy. We therefore refer to $1/\zeta$ as the *homeostatic temperature* with the understanding that it quantifies the fluctuations on a timescale much slower than that characterised by T.

2. The equilibrium Gibbs free-energy of a morphological microstate

Similar to conventional statistical mechanics calculations that require a model for the energy of the system, the homeostatic statistical mechanics framework requires a model for the Gibbs free-energy $G^{(j)}$ of morphological microstate (j). Here, we calculate $G^{(j)}$ using the free-energy model of Vigliotti *et al*. (3) (as modified by Shishvan *et al*. (1)) that includes contributions from cell elasticity and the actin/myosin stress-fibre cytoskeleton, with the cell modelled as a two-dimensional (2D) body in the $x_1 - x_2$ plane adhered to a deformable collagen coated substrate, such that the out-of-plane Cauchy stress $\Sigma_{33} = 0$. The key differences are that (i) we include a nucleus that was neglected in Shishvan *et al*. (1) (ii) and also add in explicit contribution from the focal adhesions as described in the main body of the paper. The state of the system changes as the cell moves, spreads, and changes shape on the substrate. Here, we shall give a prescription to calculate the Gibbs free-energy of the system when the cell is in a specific morphological microstate (j), where the connections of material points on the cell membrane to the surface of the collagen coated substrate are specified.

With the system comprising the cell and substrate within a constant temperature and pressure nutrient bath, the equilibrium value of the Gibbs free-energy $G^{(j)}$ of the system in morphological microstate (j) is given by $G^{(j)} = F_{\text{cell}}^{(j)} + F_{\text{sub}}^{(j)} + F_{\text{adh}}^{(j)}$. Here $F_{\text{adh}}^{(j)}$ is the free-energy associated with the formation of focal adhesion as described in the main paper body (see Section 2.3 therein). Moreover, we assume the substrate to be linear elastic so $F_{\text{sub}}^{(j)}$ is calculated directly from knowing the tractions the cells exert on a linear elastic half-space. Thus, we focus our description on the calculation of the free-energy of the cell $F_{\text{cell}}^{(j)}$. In the following, for the sake of notational brevity, we shall drop the superscript (j) that denotes the morphological microstate, as the entire discussion refers to a single morphological microstate.

The Vigliotti *et al*. (3) model assumes only two elements within the cell: (i) a passive elastic contribution from elements such as the cell membrane, intermediate filaments and microtubules, and (ii) an active contribution from contractile acto-myosin stress-fibres that are modelled explicitly. This model was modified in Shishvan *et al.* (1) to incorporate a non-dilute concentration of stress-fibres and here we further modify this model by including the nucleus in the analysis as a passive elastic body, in addition to the cytoplasm comprising the two above mentioned components. We shall first describe the modelling of the active acto-

myosin stress-fibres in the cytoplasm and then discuss the elastic model of both the nucleus and the cytoplasm.

Consider a two-dimensional (2D) cell of thickness b_0 and volume V_0 in its elastic resting state comprising a nucleus of volume V_{N} and cytoplasm of volume V_{C} such that $V_0 = V_{\mathrm{N}} + V_{\mathrm{C}}$. The representative volume element (RVE) of the stress-fibres within the cytoplasm in this resting configuration is assumed to be a cylinder of volume $V_{\mathrm{R}} = \pi b_0 (n^{\mathrm{R}} \ell_0/2)^2$, where ℓ_0 is the length of a stress-fibre functional unit in its ground-state, and n^{R} is the number of these ground-state functional units within this reference RVE. The total number of functional unit packets within the cell is N_0^{T}, and we introduce $N_0 = N_0^{\mathrm{T}} V_{\mathrm{R}}/V_{\mathrm{C}}$ as the average number of functional unit packets available per RVE; N_0 shall serve as a useful normalisation parameter. The state of stress-fibres at location x_i within the cell is described by their angular concentration $\eta(x_i, \varphi)$, and there are $n(x_i, \varphi)$ functional units in series along the length of each stress-fibre in the RVE. Here, φ is the angle of the stress-fibre bundle in the undeformed configuration with respect to the $x_2 -$ direction. Vigliotti *et al.* (3) showed that, at steady-state, the number n^{ss} of functional units within the stress-fibres is given by

$$\hat{n}^{\mathrm{ss}} \equiv \frac{n^{\mathrm{ss}}}{n^{\mathrm{R}}} = \frac{[1 + \varepsilon_{\mathrm{nom}}(x_i, \varphi)]}{1 + \tilde{\varepsilon}_{\mathrm{nom}}^{\mathrm{ss}}}, \tag{S5}$$

where $\tilde{\varepsilon}_{\mathrm{nom}}^{\mathrm{ss}}$ is the strain at steady-state within a functional unit of the stress-fibres, and $\varepsilon_{\mathrm{nom}}(x_i, \varphi)$ is the nominal strain in direction φ. The chemical potential of the functional units within the stress-fibres in terms of the Boltzmann constant k_{B} is given by

$$\chi_{\mathrm{b}} = \frac{\mu_{\mathrm{b}}}{n^{\mathrm{R}}} + k_{\mathrm{B}} T \ln \left[\left(\frac{\pi \hat{\eta}\, \hat{n}^{\mathrm{ss}}}{\widehat{N}_{\mathrm{u}} \left(1 - \dfrac{\hat{\eta}}{\hat{\eta}_{\mathrm{max}}} \right)} \right)^{\frac{1}{n^{\mathrm{ss}}}} \left(\frac{\widehat{N}_{\mathrm{u}}}{\pi \widehat{N}_{\mathrm{L}}} \right) \right] \tag{S6}$$

where the normalized concentration of the unbound stress-fibre proteins is $\widehat{N}_{\mathrm{u}} \equiv N_{\mathrm{u}}/N_0$ with $\hat{\eta} \equiv \eta n^{\mathrm{R}}/N_0$, while $\hat{\eta}_{\mathrm{max}}$ is the maximum normalised value of $\hat{\eta}$ corresponding to full occupancy of all available sites for stress-fibres (in a specific direction). Here, $\widehat{N}_{\mathrm{L}}$ is the number of lattice sites available to unbound proteins. The enthalpy μ_{b} of n^{R} bound functional units at steady-state is given in terms of the isometric stress-fibre stress σ_{max} and the internal energy μ_{b0} as

$$\mu_{\mathrm{b}} = \mu_{\mathrm{b0}} - \sigma_{\mathrm{max}} \Omega (1 + \tilde{\varepsilon}_{\mathrm{nom}}^{\mathrm{ss}}), \tag{S7}$$

where Ω is the volume of n^{R} functional units. By contrast, the chemical potential of the unbound proteins is independent of stress and given in terms of the internal energy μ_{u} as

$$\chi_{\mathrm{u}} = \frac{\mu_{\mathrm{u}}}{n^{\mathrm{R}}} + k_B T \ln\left(\frac{\widehat{N}_{\mathrm{u}}}{\pi\, \widehat{N}_{\mathrm{L}}}\right). \tag{S8}$$

For a fixed configuration of the 2D cell (i.e. a fixed strain distribution $\varepsilon_{\mathrm{nom}}(x_i, \varphi)$), the contribution to the specific Helmholtz free-energy of the cell f from the stress-fibre cytoskeleton follows as

$$f_{\mathrm{cyto}} = \rho_0 \left(\widehat{N}_{\mathrm{u}} \chi_{\mathrm{u}} + \int_{-\pi/2}^{\pi/2} \hat{\eta}\, \hat{n}^{\mathrm{ss}} \chi_{\mathrm{b}} d\varphi\right), \tag{S9}$$

where $\rho_0 \equiv N_0/V_{\mathrm{R}}$ is the number of protein packets per unit reference volume available to form functional units in the cell. However, we cannot yet evaluate f_{cyto} as $\widehat{N}_{\mathrm{u}}(x_i)$ and $\hat{\eta}(x_i, \varphi)$ are unknown. These will follow from the chemical equilibrium of the cell as will be discussed in Section 2.1.

The total stress Σ_{ij} within the cell includes contributions from the passive elasticity provided mainly by the intermediate filaments of the cytoskeleton attached to the nuclear and plasma membranes and the microtubules, as well as the active contractile stresses of the stress-fibres. The total Cauchy stress is written in an additive decomposition as

$$\Sigma_{ij} = \sigma_{ij} + \sigma_{ij}^{\mathrm{p}}, \tag{S10}$$

where σ_{ij} and σ_{ij}^{p} are the active and passive Cauchy stresses, respectively. In the 2D setting with the cell lying in the $x_1 - x_2$ plane, the active stress is given in terms of the volume fraction v_0 of the stress-fibre proteins as

$$\begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix} = \frac{v_0 \sigma_{\max}}{2} \int_{-\pi/2}^{\pi/2} \hat{\eta}[1 + \varepsilon_{\mathrm{nom}}(\varphi)] \begin{bmatrix} 2\cos^2\varphi^* & \cos 2\varphi^* \\ \cos 2\varphi^* & 2\sin^2\varphi^* \end{bmatrix} d\varphi, \tag{S11}$$

where φ^* is the angle of the stress-fibre measured with respect to x_2, and is related to its orientation φ in the undeformed configuration by the rotation with respect to the undeformed configuration. The passive elasticity in the 2D setting is given by a 2D specialization of the Ogden (4) hyperelastic model as derived in Shishvan *et al.* (1). The strain energy density function of this 2D Ogden model is

$$\Phi_{\mathrm{C}} \equiv \frac{2\mu_{\mathrm{C}}}{m_{\mathrm{C}}^2}\left[\left(\frac{\lambda_{\mathrm{I}}}{\lambda_{\mathrm{II}}}\right)^{\frac{m_{\mathrm{C}}}{2}} + \left(\frac{\lambda_{\mathrm{II}}}{\lambda_{\mathrm{I}}}\right)^{\frac{m_{\mathrm{C}}}{2}} - 2\right] + \frac{\kappa_{\mathrm{C}}}{2}(\lambda_{\mathrm{I}}\lambda_{\mathrm{II}} - 1)^2, \tag{S12}$$

for the cytoplasm and

$$\Phi_{\mathrm{N}} \equiv \frac{2\mu_{\mathrm{N}}}{m_{\mathrm{N}}^2}\left[\left(\frac{\lambda_{\mathrm{I}}}{\lambda_{\mathrm{II}}}\right)^{\frac{m_{\mathrm{N}}}{2}} + \left(\frac{\lambda_{\mathrm{II}}}{\lambda_{\mathrm{I}}}\right)^{\frac{m_{\mathrm{N}}}{2}} - 2\right] + \frac{\kappa_{\mathrm{N}}}{2}(\lambda_{\mathrm{I}}\lambda_{\mathrm{II}} - 1)^2, \tag{S13}$$

for the nucleus where λ_{I} and λ_{II} are the principal stretches, μ_{C} (μ_{N}) and κ_{C} (κ_{N}) the shear modulus and in-plane bulk modulus of cytoplasm (nucleus), respectively, while m_{C} (m_{N}) is a material constant governing the non-linearity of the deviatoric elastic response of cytoplasm (nucleus). The cell is assumed to be incompressible, and thus throughout the cell, we set the principal stretch in the x_3 −direction $\lambda_{\mathrm{III}} = 1/(\lambda_{\mathrm{I}}\lambda_{\mathrm{II}})$. The (passive) Cauchy stress then follows as $\sigma_{ij}^{\mathrm{p}} p_j^{(k)} = \sigma_k^{\mathrm{p}} p_i^{(k)}$ in terms of the principal (passive) Cauchy stresses σ_k^{p} ($\equiv \lambda_k \partial\Phi_{\mathrm{C}}/\partial\lambda_k$ for the cytoplasm and $\equiv \lambda_k \partial\Phi_{\mathrm{N}}/\partial\lambda_k$ for the nucleus) and the unit vectors $p_j^{(k)}$ ($k = \mathrm{I}, \mathrm{II}$) denoting the principal directions. The total specific Helmholtz free-energy of the cytoplasm is then $f = f_{\mathrm{cyto}} + \Phi_{\mathrm{C}}$ while that of the nucleus is $f = \Phi_{\mathrm{N}}$.

2.1. Equilibrium of the morphological microstate

Shishvan *et al*. (1) have shown that equilibrium of a morphological microstate reduces to two conditions: (i) mechanical equilibrium with $\Sigma_{ij,j} = 0$ throughout the system, and (ii) chemical equilibrium such that $\chi_{\mathrm{u}}(x_i) = \chi_{\mathrm{b}}(x_i, \varphi) = \text{constant}$, i.e. the chemical potentials of bound and unbound stress-fibre proteins are equal throughout the cell. The condition $\chi_{\mathrm{u}} = \chi_{\mathrm{b}}$ implies that $\hat{\eta}(x_i, \varphi)$ is given in terms of $\widehat{N}_{\mathrm{u}}$ by

$$\hat{\eta}(x_i, \varphi) = \frac{\widehat{N}_{\mathrm{u}}\, \hat{\eta}_{\max} \exp\left[\frac{\hat{n}^{\mathrm{ss}}(\mu_{\mathrm{u}} - \mu_{\mathrm{b}})}{k_B T}\right]}{\pi \hat{n}^{\mathrm{ss}} \hat{\eta}_{\max} + \widehat{N}_{\mathrm{u}} \exp\left[\frac{\hat{n}^{\mathrm{ss}}(\mu_{\mathrm{u}} - \mu_{\mathrm{b}})}{k_B T}\right]}, \tag{S14}$$

and $\widehat{N}_{\mathrm{u}}$ follows from the conservation of stress-fibre proteins throughout the cytoplasm, viz.

$$\widehat{N}_{\mathrm{u}} + \frac{1}{V_{\mathrm{C}}}\int_{V_{\mathrm{C}}} \int_{-\pi/2}^{\pi/2} \hat{\eta}\, \hat{n}^{\mathrm{ss}} d\varphi \; dV = 1. \tag{S15}$$

Knowing $\widehat{N}_{\mathrm{u}}$ and $\hat{\eta}(x_i, \varphi)$, the stress Σ_{ij} can now be evaluated and these stresses within the system (i.e. cell and substrate) need to satisfy mechanical equilibrium, i.e. $\Sigma_{ij,j} = 0$. In this case, the mechanical equilibrium condition is readily satisfied as the stress field Σ_{ij} within the

cell is equilibrated by a traction field T_i exerted by the substrate on the cell such that $b\Sigma_{ij,j} = -\mathrm{T}_i$, where $b(x_i)$ is the thickness of the cell in the current configuration. The substrate energy F_{sub} is calculated by applying these tractions to a linear elastic half-space as described in Shishvan *et al*. (1).

The equilibrium free-energy is then given as

$$G \equiv \rho_0 V_{\mathrm{C}} \chi_{\mathrm{u}} + \int_{V_{\mathrm{C}}} \Phi_{\mathrm{C}}\, dV + \int_{V_{\mathrm{N}}} \Phi_{\mathrm{N}}\, dV + F_{\mathrm{sub}} + F_{\mathrm{adh}}. \tag{S16}$$

Here, χ_{u} is given by eq. (S8) with the equilibrium value of $\widehat{N}_{\mathrm{u}}$ obtained from eq. (S15). For the purposes of further discussion, we label the equilibrium value $F_{\mathrm{cyto}} \equiv \rho_0 V_{\mathrm{C}} \chi_{\mathrm{u}}$ as the cytoskeletal free-energy of the cell, and $F_{\mathrm{passive}} \equiv \int_{V_{\mathrm{C}}} \Phi_{\mathrm{C}}\, dV + \int_{V_{\mathrm{N}}} \Phi_{\mathrm{N}}\, dV$ as the passive elastic energy of the cell. Moreover, $F_{\mathrm{cell}} \equiv F_{\mathrm{cyto}} + F_{\mathrm{passive}}$.

2.2. Numerical methods

We employ Markov Chain Monte Carlo (MCMC) to construct a Markov chain that is representative of the homeostatic ensemble. This involves three steps: (i) a discretization scheme to represent morphological microstate (j), (ii) calculation of $G^{(j)}$ for a given morphological microstate (j), and (iii) construction of a Markov chain comprising these morphological microstates. Here, we briefly describe the procedure which was implemented in MATLAB with readers referred to (1) for further details. Typical Markov chains comprised in excess of 2.5 million samples.

In the general setting of a three-dimensional (3D) cell, a morphological microstate is defined by the connection of material points on the cell membrane to the surface of the collagen coated substrate. In the 2D context of cells on collagen coated substrates, this reduces to specifying the connection of all material points of the cell to locations within the collagen coated substrate, i.e. a displacement field $u_i^{(j)}(X_i)$ is imposed on the cell with X_i denoting the location of material points on the cell in the undeformed configuration, and these are then displaced to $x_i^{(j)} = X_i + u_i^{(j)}$ in morphological microstate (j). These material points located at $x_i^{(j)}$ are then connected to material points on the collagen coated substrate at the same location $x_i^{(j)}$, completing the definition of the morphological microstate in the 2D setting.

The cell is modelled as a continuum and thus $u_i^{(j)}$ is a continuous field. To calculate the density of the morphological microstates, we define $u_i^{(j)}$ via Non-Uniform Rational B-splines

(NURBS) such that the morphological microstate is now defined by M pairs of weights $\left[U_L^{(j)}, V_L^{(j)}\right]$ $(L = 1, \dots, M)$. In all the numerical results presented here, we employ $M = 16$ with 4×4 weights $U_L^{(j)}$ and $V_L^{(j)}$ governing the displacements in the x_1 and x_2 directions, respectively. The NURBS employ fourth order base functions for both the x_1 and x_2 directions, and the knots vector included two nodes each with multiplicity four, located at the extrema of the interval. We emphasise here that this choice of representing the morphological microstates imposes restrictions on the morphological microstates that will be considered. Therefore, the choice of the discretisation used to represent $u_i^{(j)}$ needs to be chosen so as to be able to represent the microstates we wish to sample, e.g. the choice can be based on the minimum width of a filopodium one expects for the given cell type. Given $u_i^{(j)}$, we can calculate $G^{(j)}$ using the model described in Section 2.1 with the cell discretised using constant strain triangles of size $e \approx R_0/10$, where R_0 is the radius of the cell in its undeformed configuration.

We construct the Markov chain using the Metropolis (5) algorithm that gives a sequence of random samples from the exponential equilibrium distribution (eq. (S5)). We employ the Metropolis algorithm in an iterative manner so as to enforce the homeostatic constraint (eq. (S3)). The scheme is summarised as follows:

(i) Assume a value of ζ and use the undeformed cell configuration as the starting configuration and label it as morphological microstate $j = 0$ with equilibrium free-energy $G^{(0)}$ calculated as described in Section 2.1.

(ii) Randomly pick one pair of the M weights $U_L^{(j)}, V_L^{(j)}$ and perturb them by two independent random numbers picked from a uniform distribution over the interval $[-\Delta, \Delta]$.

(iii) Compute the new free-energy of this perturbed state and thereby the change in free-energy $\Delta G = G^{(j)} - G^{(j-1)}$.

(iv) Use the Metropolis criterion to accept this perturbed state or not, i.e.
   a. if $\Delta G \leq 0$, accept the perturbed state;
   b. if $\Delta G > 0$, compute $P^{acc} = \exp(-\zeta \Delta G)$ and accept the perturbed state if $P^{acc} > \mathcal{R}$, where $\mathcal{R}$ is a random number drawn from a uniform distribution over $[0, 1]$.

(v) If the perturbed state is accepted, add it to the list of samples as a new morphological microstate, else repeat the configuration prior to step (ii) in the sample list and return to step (ii).

(vi) Keep repeating this procedure until a converged distribution is obtained. Here, we typically use the criterion that the average of $G^{(j)}$ within the generated sample list (labelled $\langle G^{(j)} \rangle$ changes by less than 1% over 100,000 steps of the Markov chain.

(vii) If $\langle G^{(j)} \rangle$ is within $\pm 2\%$ of the homeostatic value of G_s, we accept this distribution, else we modify ζ and repeat from step (i).

2.3. Material parameters

All simulations are reported at a reference thermodynamic temperature $T = T_0$, where $T_0 = 310$ K. Most of the parameters of the model are related to the properties of the proteins that constitute stress-fibres. These parameters are thus expected to be independent of cell type. Notable exceptions to this are: (i) the stress-fibre protein volume fraction v_0; and (ii) the passive elastic properties. Here, we use parameters calibrated for SMCs that give good correspondence with the wide range of measurements reported here. The passive elastic parameters of the cytoplasm are taken to be $\mu_{\mathrm{C}} = 1.67$ kPa, $\kappa_{\mathrm{C}} = 35$ kPa and $m_{\mathrm{C}} = 5$, while the corresponding values for the nucleus are $\mu_{\mathrm{N}} = 3.33$ kPa, $\kappa_{\mathrm{N}} = 35$ kPa and $m_{\mathrm{N}} = 20$ (6–8). The maximum contractile stress $\sigma_{\max} = 240$ kPa is consistent with a wide range of measurements on muscle fibres (9), and the density of stress-fibre proteins was taken as $\rho_0 = 3 \times 10^6\ \mu\mathrm{m}^{-3}$ with the volume fraction of stress-fibre proteins $v_0 = 0.032$. Following Vigliotti *et al*. (3), we assume that the steady-state functional unit strain $\tilde{\varepsilon}^{\mathrm{ss}}_{\mathrm{nom}} = 0.35$ with $\mu_{\mathrm{b0}} - \mu_{\mathrm{u}} = 2.3\, kT_0$ and $\Omega = 10^{-7.1}\ \mu\mathrm{m}^3$. The maximum angular concentration of stress-fibre proteins is set to $\hat{\eta}_{\max} = 1$. The cell in its undeformed state is a circle of radius R_0 and thickness $b_0 = 0.05R_0$, with a circular nucleus of radius $R_{\mathrm{N}} = 10b_0$ whose centre coincides with that of the cell. Results are presented in terms of normalised cell area $\hat{A}^{(j)} \equiv A^{(j)}/A_0$, where $A^{(j)}$ is the area of a morphological microstate (j) while $A_0 = \pi R_0^2$ is the area of the undeformed cell. Thus, we do not explicitly need to specify R_0.

Results are presented for adhesion to incompressible linear elastic substrates with Young's modulus $E_{\mathrm{sub}} = 8\,\mathrm{kPa}, 32\,\mathrm{kPa}$ and a rigid substrate with $E_{\mathrm{sub}} \to \infty$ coated with surface densities of collagen ranging between 6 ng cm^{-2} to 665 ng cm^{-2}. Using a molecular weight of collagen M_{col} of 200 kDa (10) (i.e. 200 kg mol^{-1}) with 1 ligand per molecule, a surface collagen density ρ_{col} of 1 g cm^{-2} corresponds to $3.011\, x10^9$ ligands/cm^2. A range of

ligand densities from $N_{\mathrm{H}} = 175\ \mu\mathrm{m}^{-2}$ to $N_{\mathrm{H}} = 20 \times 10^3\ \mu\mathrm{m}^{-2}$ are analyzed, corresponding to the range of ρ_{col} investigated in the experiments of Engler *et al*. (11). The other parameters of the focal adhesion model are based on commonly accepted ranges; see for example Deshpande *et al*. (12). Specifically, we assumed a uniform surface density of $C_0 = 5 \times 10^3\ \mu\mathrm{m}^{-2}$ of integrin molecules with surface density of integrins sites $C_{\mathrm{r}} = 20 \times 10^3\ \mu\mathrm{m}^{-2}$. The complex stiffness $\kappa_{\mathrm{s}} = \kappa_{\mathrm{p}} = 0.3\ \mathrm{nN}\ \mu\mathrm{m}^{-1}$ while the maximum complex force $\mathrm{F}_{\max} = 0.01\ \mathrm{nN}$ (13).

2.4. Definitions of normalised quantities and observables

Following Shishvan *et al.* (1), the free-energy $G^{(j)}$ can be decomposed as $G^{(j)} = \Upsilon^{(j)} + \Upsilon_0$, where $\Upsilon_0 = \rho_0 V_0 \left[\mu_{\mathrm{u}}/n^{\mathrm{R}} - kT \ln\left(\pi \widehat{N}_{\mathrm{L}}\right)\right]$ is independent of the morphological microstate. It is thus natural to subtract out Υ_0 and define a normalised free-energy as

$$\hat{G}^{(j)} \equiv \frac{\Upsilon^{(j)}}{|G_{\mathrm{S}} - \Upsilon_0|} = \frac{G^{(j)} - \Upsilon_0}{|G_{\mathrm{S}} - \Upsilon_0|}\ , \tag{S17}$$

where G_{S} is the equilibrium free-energy of a free-standing cell (i.e. a cell in suspension with traction-free surfaces). Then, the distribution given by eq. (S6) can be re-written as

$$P_{\mathrm{eq}}^{(j)} = \frac{1}{\hat{Z}} \exp\left[-\hat{\zeta}\hat{G}^{(j)}\right] \tag{S18}$$

with $\hat{Z} \equiv \sum_j \exp\left[-\hat{\zeta}\hat{G}^{(j)}\right]$ and $\hat{\zeta} \equiv \zeta |G_S - \Upsilon_0|$. It then immediately follows that the distributions of states are not influenced by the values of n^{R}, $\widehat{N}_{\mathrm{L}}$ and V_0 and these parameters need not be specified so long as energies are quoted in terms of the normalised energies $\hat{G}^{(j)}$.

Analogously, we define the normalised elastic, cytoskeletal, adhesion, and substrate free-energies of the spread microstate (j) as

$$\hat{F}_{\mathrm{passive}}^{(j)} \equiv \frac{F_{\mathrm{passive}}^{(j)}}{|G_{\mathrm{S}} - \Upsilon_0|}\ , \qquad \hat{F}_{\mathrm{cyto}}^{(j)} \equiv \frac{F_{\mathrm{cyto}}^{(j)} - \Upsilon_0}{|G_{\mathrm{S}} - \Upsilon_0|} \tag{S19}$$

and

$$\hat{F}_{\mathrm{sub}}^{(j)} \equiv \frac{F_{\mathrm{sub}}^{(j)}}{|G_{\mathrm{S}}|}, \qquad \hat{F}_{\mathrm{adh}}^{(j)} \equiv \frac{F_{\mathrm{adh}}^{(j)}}{|G_{\mathrm{S}} - \Upsilon_0|}, \tag{S20}$$

respectively.

For the SMCs with volume $V_0 = \pi R_0^2 b_0$ modelled here, the equilibrium free-standing microstate is a spatially uniform circular cell with $(G_{\mathrm{S}} - \Upsilon_0)/(V_0 k T_0) \approx -4.53\ x\ 10^6\ \mu\mathrm{m}^{-3}$. The free-standing cell is a standard boundary value problem with traction-free boundary

conditions. An iterative FE scheme as described in Shishvan *et al.* (1) is used to solve this boundary value problem and calculate G_S. This solution predicts that the free-standing is circular cell with radius $0.92\ R_0$.

Bibliography

1. Shishvan, S.S., A. Vigliotti, and V.S. Deshpande. 2018. The homeostatic ensemble for cells. Biomech. Model. Mechanobiol. https://doi.org/10.1007/s10237-018-1048-1.
2. Jaynes, E.T. 1957. Information Theory and Statistical Mechanics. Phys. Rev. 106: 620–630.
3. Vigliotti, A., W. Ronan, F.P.T. Baaijens, and V.S. Deshpande. 2016. A thermodynamically motivated model for stress-fiber reorganization. Biomech. Model. Mechanobiol. 15: 761–789.
4. Ogden, R.W. 1972. Large Deformation Isotropic Elasticity - On the Correlation of Theory and Experiment for Incompressible Rubberlike Solids. Proc. R. Soc. A Math. Phys. Eng. Sci. 326: 565–584.
5. Metropolis, N., A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller. 1953. Equation of State Calculations by Fast Computing Machines. J. Chem. Phys. 21: 1087–1092.
6. Ronan, W., V.S. Deshpande, R.M. McMeeking, and J.P. McGarry. 2012. Numerical investigation of the active role of the actin cytoskeleton in the compression resistance of cells. J. Mech. Behav. Biomed. Mater. 14: 143–157.
7. Dowling, E.P., W. Ronan, G. Ofek, V.S. Deshpande, R.M. McMeeking, K. a Athanasiou, and J.P. McGarry. 2012. The effect of remodelling and contractility of the actin cytoskeleton on the shear resistance of single cells: a computational and experimental investigation. J. R. Soc. Interface. 9: 3469–79.
8. Dowling, E.P., W. Ronan, and J.P. McGarry. 2013. Computational investigation of in situ chondrocyte deformation and actin cytoskeleton remodelling under physiological loading. Acta Biomater. 9: 5943–5955.
9. Lucas, S.M., R.L. Ruff, and M.D. Binder. 1987. Specific tension measurements in single soleus and medial gastrocnemius muscle fibers of the cat. Exp. Neurol. 95: 142–54.
10. Raj, C. V, I.L. Freeman, R.L. Church, and S.I. Brown. 1979. Biochemical characterization of procollagen-collagen synthesized by rabbit corneal endothelial cells in culture. Invest. Ophthalmol. Vis. Sci. 18: 75–84.
11. Engler, A., M. Sheehan, H.L. Sweeney, and D.E. Discher. 2003. Substrate compliance vs ligand density in cell on gel responses. Eur. Cells Mater. 6: 7–8.
12. Deshpande, V.S., M. Mrksich, R.M. McMeeking, and A.G. Evans. 2008. A bio-mechanical model for coupling cell contractility with focal adhesion formation. J. Mech. Phys. Solids. 56: 1484–1510.
13. Kong, F., A.J. García, A.P. Mould, M.J. Humphries, and C. Zhu. 2009. Demonstration of catch bonds between an integrin and its ligand. J. Cell Biol. 185: 1275–84.

ADDITIONAL FIGURES

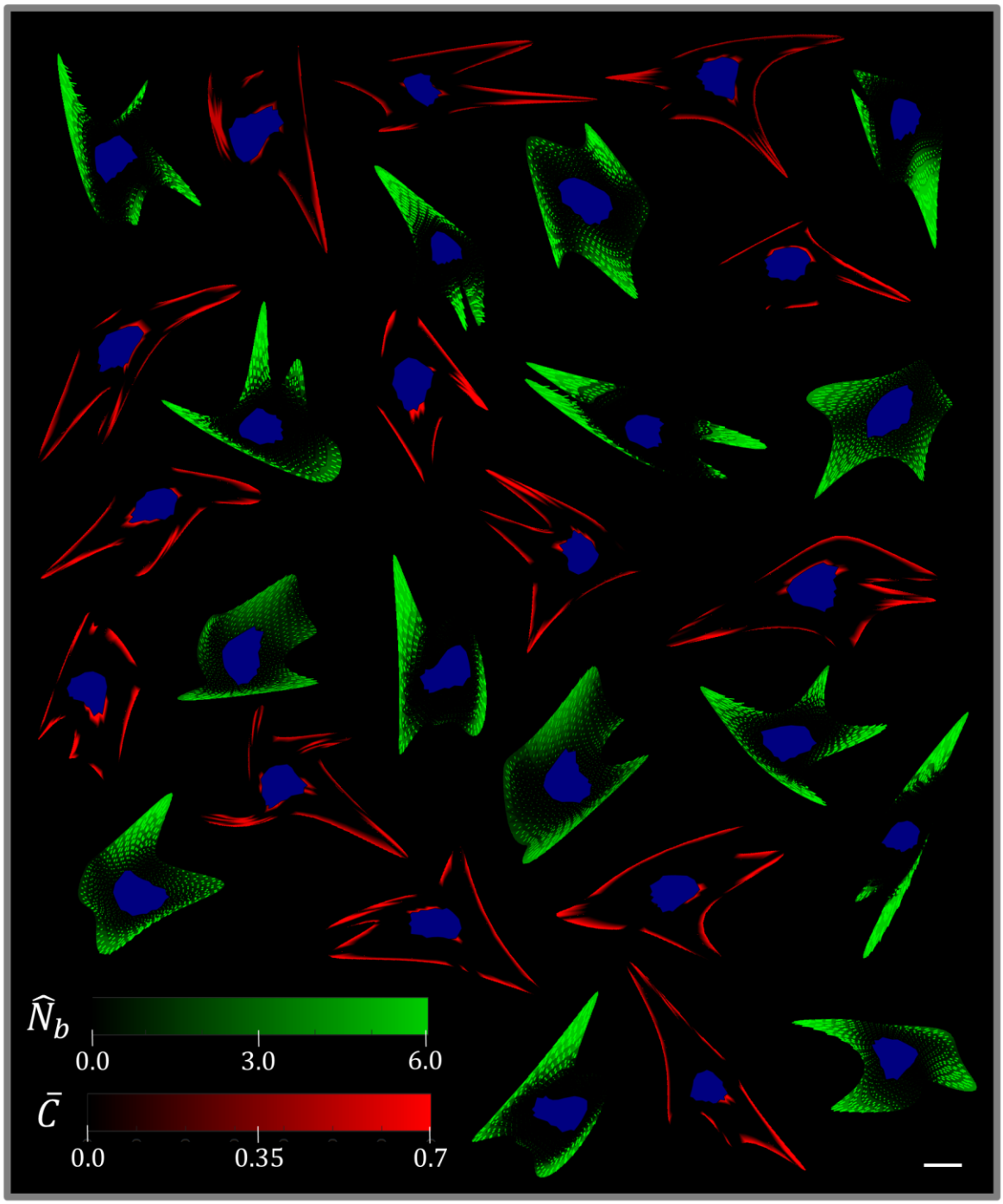

Figure SI-1: Additional configurations of bound stress-fibre protein concentrations $\widehat{N}_b$ (green) with dominant alignment, and focal adhesion distributions $\widehat{C}$ (red). The substrate is rigid, and nucleus is highlighted in blue. The configurations are all from the median of the free-energy $\widehat{G}$ distribution for cells spread on a rigid substrate with a surface collagen density ρ_{col} of 250 ng/cm^2. Scale bar indicates undeformed cell radius R_0.

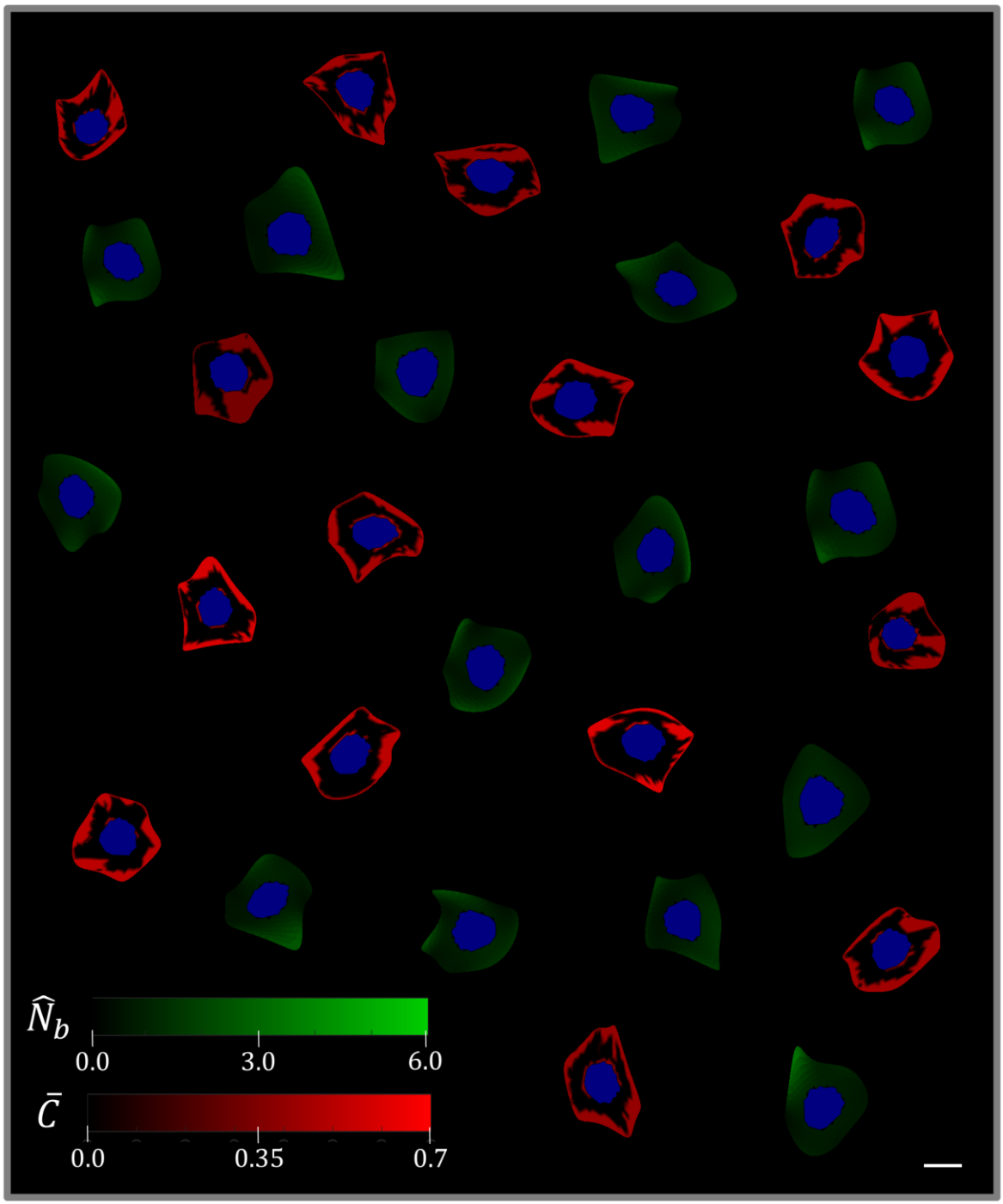

Figure SI-2: Additional configurations of bound stress-fibre protein concentrations $\widehat{N}_\mathrm{b}$ (green), and focal adhesion distributions $\widehat{C}$ (red). The substrate is rigid, and nucleus is highlighted in blue. The configurations are all from the median of the free-energy $\widehat{G}$ distribution for cells spread on a rigid substrate with a surface collagen density ρ_{col} of 6 $\mathrm{ng/cm^2}$. Scale bar indicates undeformed cell radius R_0.

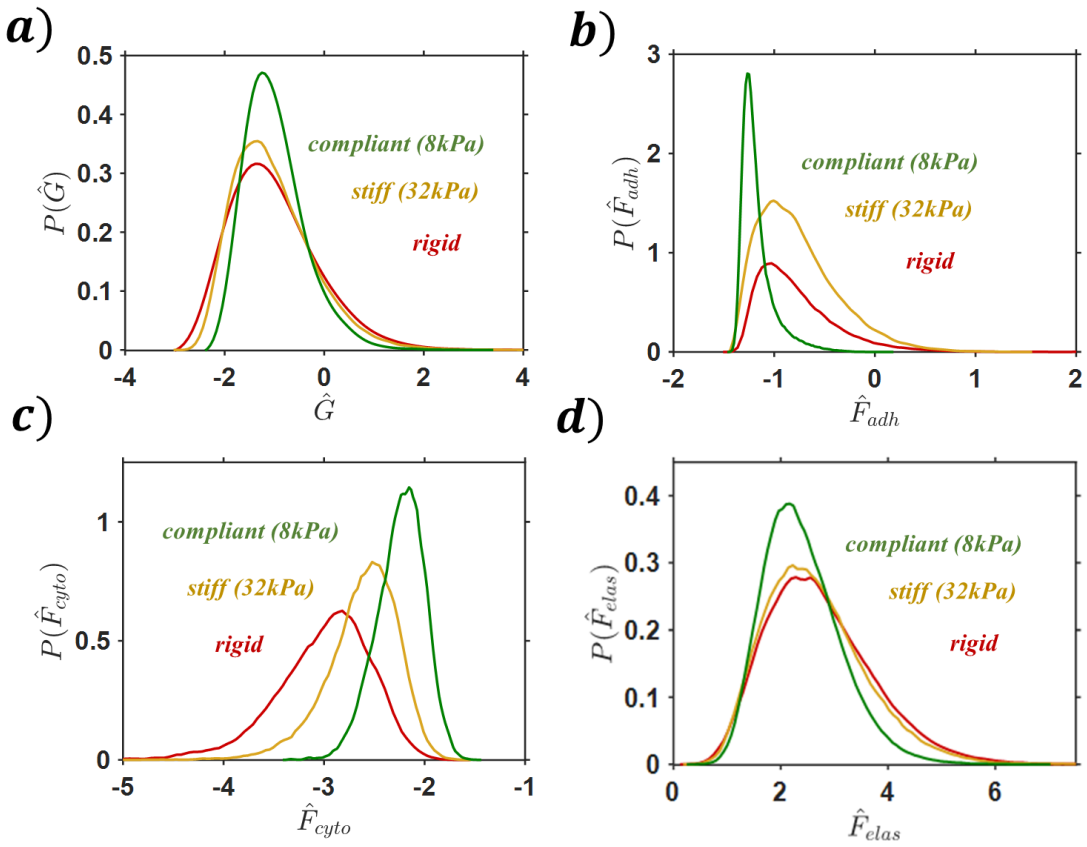

Figure SI-3: Probability density functions for cells spread on substrates of different stiffness at a collagen density of 33 $ng\ cm^{-2}$, of (a) Gibbs free-energy, (b) adhesion free-energy, (c) cytoskeletal free-energy, and (d) elastic free-energy ($\widehat{\boldsymbol{F}}_{\boldsymbol{elas}} = \widehat{\boldsymbol{F}}_{\boldsymbol{passive}} + \widehat{\boldsymbol{F}}_{\boldsymbol{sub}}$).

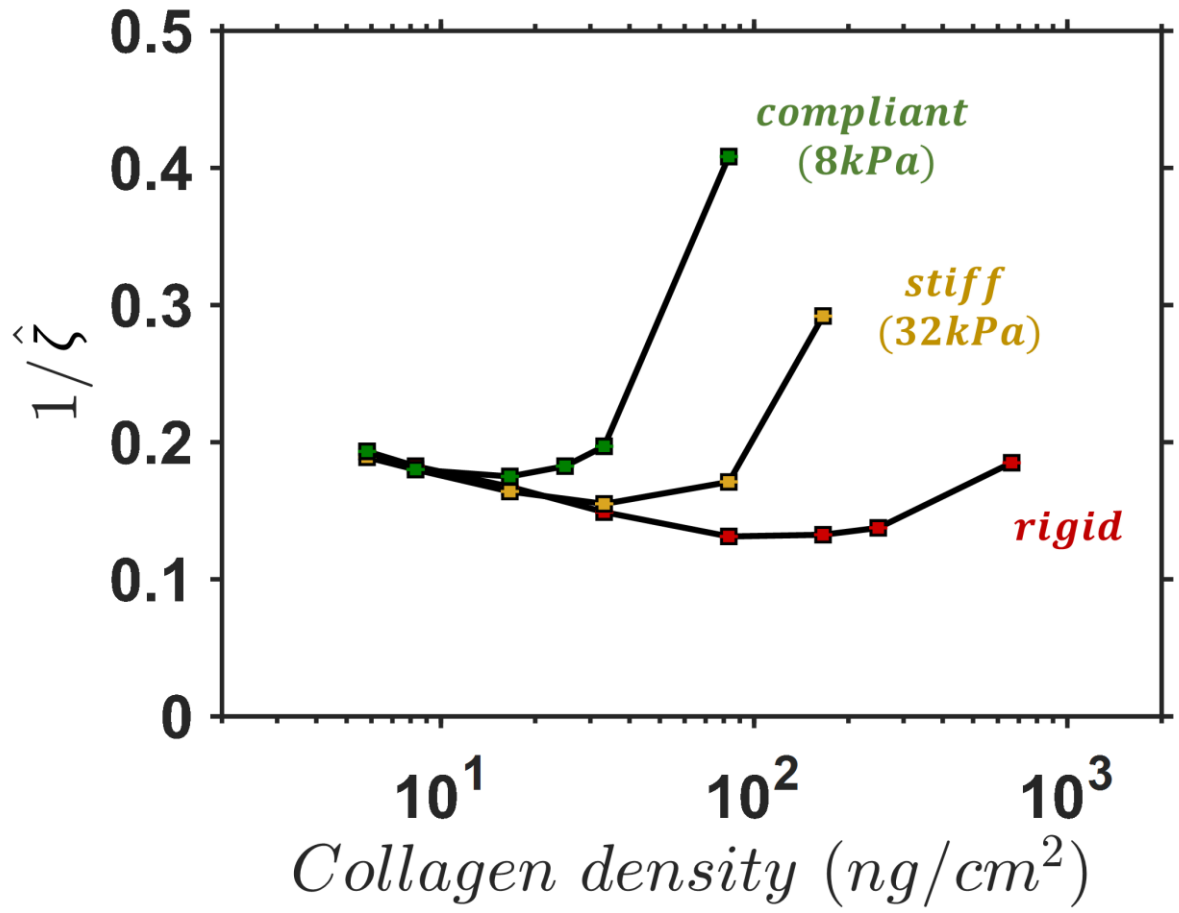

Figure SI-4: Predictions of the normalised homeostatic temperature $1/\hat{\zeta}$.