Physiochemical Properties of *Caulobacter crescentus* Holdfast: a Localized Bacterial Adhesive

Cécile Berne [†], *Xiang Ma* [‡], *Nicholas A. Licata* [§], *Bernardo R.A. Neves* [/], *Sima Setayeshgar* [//], *Yves V. Brun** [†], *Bogdan Dragnea** [‡]

[†] Department of Biology, Indiana University, Bloomington, IN 47405, USA

[‡] Department of Chemistry, Indiana University, Bloomington, IN 47405, USA

[§] Department of Natural Sciences, University of Michigan-Dearborn, Dearborn, MI 48128, USA

[/] Universidade Federal de Minas Gerais, Belo Horizonte, MG 30123-970, Brazil.

[//] Department of Physics, Indiana University, Bloomington, IN 47405, USA

SUPPLEMENTARY INFORMATION.

1. Quantitative Analysis of the Dependence of the Rupture Force on Dwell Time: Reaction-diffusion model of adhesin-surface association

1.1 Multistep surface kinetic scheme

To model the time dependence of the adhesion force on dwelling time, a reaction-diffusion surface attachment mechanism is proposed. The model was born out from the requirement of a minimal set of parameters reproducing the most salient features of the experimental results. Thus, we assume that adhesins, denoted *P*, can diffuse within the holdfast mass and bind irreversibly to the surface, denoted S. An intermediate step in the binding process is the reversible association of P with the surface. The coupling between bulk and surface concentrations of P is achieved by

including reversible adsorption and desorption kinetics from the surface, where we denote the concentration of surface associated P by $[P_a]$. The strength of holdfast attachment is proportional to the concentration of the irreversibly bound surface bound complex, $[P-S]$, and more weakly to the concentration of the reversible surface associated form $[P_a]$. We note that $[P]$ is a volume concentration, while $[S]$, $[P_a]$, and $[P-S]$ are concentrations of the substrate, adhesin, and bound complex on the surface, respectively. The resulting reaction-diffusion equations are:

$$\frac{\partial [P]}{\partial t} = D\nabla^2[P], \qquad (1)$$

$$\frac{d\,[P_a]}{dt} = -k\,[S]\,[P_a] + \,k_a[P]|_{z=0} - k_d[P_a], \quad (2)$$

$$\frac{d\,[P-S]}{dt} = k\,[S]\,[P_a] = k\,[P_a]\,([S_0] - \,[P-S]). \quad (3)$$

where D is the diffusion constant of adhesin within the holdfast mass, k_a, and k_d are reaction rates describing association and dissociation of P with the surface, and k describes the rate at which the irreversibly bound complex, $[P-S]$, is formed from the surface associated adhesin, P_a. Approximating the holdfast by a cubic volume of size $h \approx 60\,nm$ with the reactive surface at $z=0$, the boundary conditions for P are no-flux at all non-reactive boundaries ($\hat{n}\cdot D\vec{\nabla}[P]=0$, where $\hat{n}$ is the unit normal to a domain boundary). At $z=0$, the boundary condition is

$$D\frac{\partial P}{\partial z}(x,y,z=0,t) = \,-k_a[P(x,y,z=0,t)] + k_d[P_a(x,y,t)]. \quad (4)$$

It is useful to introduce dimensionless variables (denoted with hat) by rescaling all lengths by h, i.e. $\hat{x}=x/h$, and rescaling time as $\hat{t}=t/\tau_D$, where $\tau_D = h^2/D$ is a diffusive time scale. Concentrations are rescaled by the appropriate powers of h, which defines $\hat{P}=[P]h^3$, $\widehat{P-S}=[P-S]h^2$, $\hat{P_a}=[P_a]h^2, \hat{S}=[S]h^2$, and reaction rates are rescaled as $\widehat{k_a} = \,k_a\frac{h}{D}, \widehat{k_d} = \,k_d\frac{h^2}{D}$ and $\hat{k}=k/D$. Numerical solutions are obtained for different choices of total concentrations $\widehat{S_0}$ and $\widehat{P_0}$, and parameters $(\widehat{k_a}, \widehat{k_d}, \hat{k})$. A representative example of the numerical solution to the model is shown in Figure S1, which plots the dimensionless concentrations as a function of time.

Given that the time to rupture is much shorter than the experimental dwell times, we assume that the strength of the adhesion force is proportional to the concentrations of the surface associated adhesin, $[P_a]$ and $[P-S]$, that have accumulated for a given dwell time, t. Specifically,

$$F(\hat{t}) = \delta f_1 \iint \widehat{P_a}(\hat{x},\hat{y},\hat{t})d\hat{x}\,d\hat{y} + \delta f_2 \iint \widehat{P-S}(\hat{x},\hat{y},\hat{t})d\hat{x}\,d\hat{y}\,, \qquad (5)$$

where we have assumed that each $P-S$ bond contributes a constant force δf_2 to the total rupture force, and that each surface associated adhesin P_a contributes a constant force δf_1. We expect $\delta f_1 < \delta f_2$, as discussed below, based on the experimental observation that for the much longer dwell times characterizing earlier micropipette experiments of Tsang *et al.*, the measured rupture forces are greater than those obtained on the shorter time scale of dwell times accessible to AFM measurements. As we discuss later, one possible mechanism for the onset of larger forces on longer time scales is the accumulation of the $P-S$ complex, characterized by a long reaction time scale relative to the time scale of the AFM experiments (small reaction rate).

The solutions to this reaction-diffusion model provide insight into the observed experimental trends on the adhesion force. To make direct comparison between the model and data, we convert the dimensionless dwell time $\hat{t}$ into the dimensioned dwell time t according to the time scale, $\tau_D = h^2/D$. We note that although the value of $h \approx 60\,nm$ is known from the experiments, the diffusion coefficient of the adhesin D within the bulk of the holdfast is unknown. Hence, we treat τ_D as a model parameter. The final parameters to be specified are δf_1 and δf_2. The parameters δf_1 and δf_2 are determined by minimizing the cost function $\chi^2 = \sum_{i=1}^{N} \frac{1}{\sigma_i^2}(F(t_i) - F_{exp}(t_i))^2$ subject to the constraint that $\delta f_2 \geq n\delta f_1$, where n is some factor, chosen to be $n = 2$ in the fits reported here. Here σ_i is the standard deviation of the experimentally measured rupture force $F_{exp}(t_i)$ for dwell time t_i. This constraint ensures that a $P-S$ bond contributes a larger force than a surface associated contact, P_a. To produce model fits, as in Figure 12 and Figure S2, we discretize parameter space for a broad range of values for the parameters $(\hat{P}_0, \hat{S}_0, \widehat{k_a}, \widehat{k_d}, \hat{k}, \tau_D, \delta f_1, \delta f_2)$. We set initial values of δf_1 and δf_2 consistent with the above mentioned constraint and compute the cost function. We then perform a random walk in the parameter space of $(\delta f_1, \delta f_2)$, with a step size of $(0.01 \times \delta f_1, 0.01 \times \delta f_2)$. Steps which decrease the cost are accepted with probability 1, and those which increase the cost are accepted with probability 0.05. The random walk is continued until the cost function converges to its minimum value. Discretizing parameter space in this manner and obtaining a numerical solution at each point to construct the cost function, χ^2, does not reveal an explicit minimum. Thus, many good fits are characterized by similar values of χ^2, disallowing tight constraints on parameter values. Further characterization of the

adhesin and its biochemical properties is required to provide constraints on some model parameters, allowing others to be better determined than currently possible through a similar fitting procedure. Hence, we proceed instead to highlight trends present in the solutions to the model and their connection key experimental features.

In this model, the lag time observed in the dependence of adhesion on dwell time is related to the rate at which the surface associated adhesin, P_a, and the bound complex, $P-S$, accumulate at the surface. At fixed total substrate concentration, $[S_0]$, the rate of $P-S$ accumulation depends on the rate of irreversible adhesin-substrate complex formation, k, and the concentration of surface associated adhesin, $[P_a]$. For a fixed amount of total adhesin, P_0, within the holdfast mass, the concentration of P_a at a given dwell time depends on: (i) the diffusive time scale $\tau = h^2/D$ which sets an upper bound on the rate of accumulation of P at the surface, and (ii) the rates (k_a, k_d) governing the adsorption of P to and desorption of P_a from the surface, respectively. While (i) is likely independent of the nature of the substrate, (ii) may depend on the surface composition. We note that the parameters (k_a, k_d) governing the surface localization of the adhesin from its bulk concentration could strongly depend on the hydrophobicity/hydrophillicity of the surface. Furthermore, depending on the nature of the adhesin-substrate bond, surfaces with different chemical composition could present varying adsorption site concentrations $[S_0]$, also affecting the onset of the adhesion force.

The strength of adhesion depends on the force required to rupture a single adhesin-substrate $(P-S)$ bond, δf_2, and the force required to rupture a single surface associated (P_a) contact, δf_1. In performing fits to the data, this parameter is required to be less than the maximum forces of adhesion, which are measured to be approximately 1 nN, and additionally constrained by $\delta f_2 \geq n\delta f_1$. (Here, the factor $n = 2$ is chosen to ensure that a $P-S$ bond contributes a larger force than a P_a contact. The fit values are not sensitive to this factor.) For comparison, the strength of a single covalent bond has been measured in AFM experiments where single polysaccharide molecules covalently anchored between a surface and an AFM tip were stretched until they became detached [1]. It was found that the silicon-carbon bond ruptured at 2 ± 0.3 nN, and the sulfur-gold anchor ruptured at 1.4 ± 0.3 nN, exceeding the total rupture forces measured in this work. We conclude that for the dwell times probed, the nature of the chemical bonds between the adhesin and substrate is likely noncovalent (ionic, van der Waals, hydrogen). Within the framework of the current model, we demonstrate that noncovalent

bonds can still lead to the much larger rupture forces measured by the micropipette assay of Tsang *et al.* [2], as discussed below.

1.2. Comparison of strength of adhesion on short and long time scales

In the Tsang *et al.* [2] measurements, cells were allowed to attach to a thin flexible pipette whose force constant has been calibrated by AFM. A suction pipette is then used to grab an attached cell and pull in a direction perpendicular to the flexible pipette. The force of adhesion is then calculated by the amount of bending in the flexible pipette prior to rupture of the cell-pipette contact. To make direct comparison with these results, the force measurements from Tsang *et al.* [2] have been rescaled by the ratio of areas $h^2/h_s^2 \approx 1/4$, where h^2 is the contact area between the shed holdfast and the surface in the current AFM experiments, and h_s^2 is the cross sectional area of the stalk. This rescaling is necessary because in the micropipette experiments, the holdfast is in contact with the surface over an area which is at least as large as the cross sectional area of the stalk. The qualitative picture that emerges within the framework of the current model is as follows: For short dwell times, the rupture force ($F \lesssim 1$ nN) is due mostly to the surface associated contacts, P_a, which have not yet undergone the irreversible transition to the surface bound form $P-S$. If the rate k for this transition is small, larger rupture forces will only be measured for longer dwell times. Fig 12A shows a fit of the numerical solution to the model to the maximum adhesion force from the current AFM measurements for dwell times, $t \leq 100$ s. This figure also shows that for the same fit parameters, on the longer time scale, $t \sim 2 \times 10^5$ s, of the micropipette pulling measurements of Tsang *et al.* [2], significantly larger forces consistent with those measurements are obtained.

In the present model, the multiple time scales arise from the two-step surface reaction scheme: the small reaction rate (and long time scale) characterizing the formation of $P-S$ bonds leads to the stronger forces measured on the longer time scales of micropipette experiments, while the faster time scales associated with diffusion of adhesin within the bulk of the holdfast and its surface adsorption/desorption kinetics lead to smaller forces on the shorter time scale of the AFM experiments. However, other mechanisms leading to multiple time scales are also possible, and we discuss two plausible mechanisms below.

1.3. Alternative Mechanisms

Slow diffusion limit: Within the framework of the above model, if the second, slow surface reaction step is absent, then the adhesion force would arise entirely from the surface associated adhesin concentration, $[P_a]$. In this case, if the diffusion constant associated with movement of the adhesin within the holdfast mass is very slow, then the larger forces on longer time scales could arise from the diffusion-limited rate of arrival of adhesin to the surface. Thus, the smaller forces on shorter time scales result from the presence of adhesin molecules already in contact with the surface, with the dependence of the adhesion force on dwell time characterized by the rates of adsorption and desorption, $k_{a,d}$.

To justify a small diffusion constant, a common biophysical mechanism is the reduction of the diffusion constant of a chemical species due to its rescaling in the presence of traps [11]. If the adhesin binds to the holdfast matrix, denoted G, with on /off rates, k_+ and k_-, respectively, then its dynamics within the holdfast mass is given by

$$\frac{\partial P}{\partial t} = D\nabla^2[P] - k_+\,[G]\,[P] + \,k_-[G-P], \quad (6)$$

$$\frac{d\,[G-P]}{dt} = \,k_+[G][P] - k_-[G-P]. \quad (7)$$

In the limit that the dynamics of binding/unbinding of adhesin to/from the matrix are fast compared to the time scale of its diffusion, then $\frac{d\,[P-G]}{dt} \approx 0$, i.e., the adhesin-matrix complex, $[G-P]$, is assumed to be in pseudo-equilibrium. Furthermore if $[G-P] \ll [G_0]$, an approximation that holds true if the concentration of total adhesin is much smaller than that of available matrix binding sites, $[G_0]$, then $[G-P] \approx k_+[G_0][P]/k_-$. Adding Eqs. (6) and (7), we find that the dynamics of the adhesin can be approximated by the diffusion equation, with a diffusion constant rescaled according to $D \to D/\alpha$, where $\alpha = \frac{k_+[G_0]}{k_-}$. If the kinetics of adhesin-matrix binding is characterized by a very small unbinding rate, a condition required for overall strength of the holdfast, then α can be very large, leading to a small, effective diffusion constant.

We carried out numerical simulations with D small, eliminating the slow surface reaction step ($k = 0$). In this case, the force of adhesion is due entirely to the formation of the surface associated form P_a. For concreteness, the diffusion constant is rescaled according to $\alpha = 1000$. We have also investigated $\alpha = 100$, with no qualitative change in conclusions. As demonstrated in Figure S2, there are two plateaus in the concentration of surface-adsorbed adhesin and rupture force. The first plateau occurs as the initial

adhesin molecules within the holdfast layer proximal to the surface reach equilibrium in their binding to the surface. The second plateau occurs as the remaining adhesin molecules within the bulk of the holdfast diffuse to the surface. In Figure S2, the second plateau occurs for a force of approximately 10 nN, which is an order of magnitude larger than the forces as measured by AFM for dwell times less than 100 seconds. However, this plateau is less than characteristic forces on the order of 100 nN measured for much longer dwell times reported by Tsang *et al.* [2]. Due to the dimensionality and dynamic range of the parameter space, it is difficult to rule out the possibility that the slow diffusion scenario with single-step reaction kinetics can better capture both the short and long time scale regimes. Future experimental studies that constrain the experimentally accessible parameter space will allow a more direct examination of the detailed assumptions of the reaction-diffusion model of surface adhesion elaborated above.

Multiple species of adhesin or N-acetylglucosamine modification: A second possible mechanism leading to the difference in the strength of adhesion on short and long time scales is the existence of multiple adhesive components, with surface binding kinetics acting on different time scales. One species of adhesin may be responsible for weaker binding with faster kinetics, while a second species may adhere with slower rates. Finally, modification of the *N*-acetylglucosamine residues could impact their adhesive properties. One of the genes of the holdfast synthesis gene cluster, *hfsH*, encodes a predicted polysaccharide deacetylase. Deacetylation of the holdfast *N*-acetylglucosamine residues impact its charge and may provide sites for modification, including crosslinking.

Cross-linking of the holdfast polymer matrix: AFM measurements of the adhesive properties of shed holdfast reported in this work likely involve multiple bonds with the surface, as assumed above. This distinguishes the present study from the large body of single molecule AFM studies of specific receptor-ligand binding [3] and protein unfolding [4], both theoretically and experimentally. The use of shed holdfast also distinguishes this work from whole cell adhesion AFM measurements, allowing the adhesive holdfast patch to be studied in isolation without interference from other cell surface interactions. Following previous works [5], a simple physical framework for modeling holdfast attachment is that of N_0 bonds acting in parallel with the load distributed (approximately) uniformly among them. Each bond is described by a fixed surface binding site and a

flexible polymer connected to the holdfast, assumed to be a Hookean spring with spring constant, K_p. As a bond ruptures, the load on the remaining bonds increases, thereby increasing their probability of rupture according to the load-dependent dissociation constant for a single bond given previously [6,7] ,

$$k_0(t) = k_0 e^{F_b(t)x_b/k_BT} , \qquad (8)$$

where k_0 is the dissociation rate in the absence of a force; x_b, is the distance between the binding potential and the barrier; $F_b(t)$, is the force on a single bond, and $\square_B T$ is the product of Boltzmann's constant and temperature. Additionally, we treat the holdfast as an elastic element with spring constant K_h, which is in turn coupled to the AFM cantilever arm with spring constant K_c, pulled at constant speed v, as shown in Fig 12B. As long as a bond is intact, it is stretched by x_p assumed to be the same for all bonds. For simplicity, rest lengths of all springs are taken to be zero. With the extension of the transducer given by $x_c(t) = vt$, force balance gives the force on the single bond, $F_b(t)$, as

$$F_b(t) = \frac{K_p}{1+N(t)K_p\left(\frac{1}{K_p}+\frac{1}{K_t}\right)} vt \qquad (9)$$

and the force on the transducer is given by $F_c(t) = N(t)F_b(t)$. Note that in our study the initial number of bonds, $N_0 = N(t=0)$, in each measurement of the rupture force is a function of the dwell time, $N_0 = N_0(\tau)$.

Previous work [5 8] has considered the dependence of the rupture force on the initial number of bonds, N_0, for both irreversible bonds (where a bond once ruptured does not rebind over the course of the pulling experiment), and reversible bonds (where the rebinding rate is not zero). The loading parameter, μ, given by

$$\mu = \frac{x_p v}{k_0\, k_B T}\frac{K_t K_h}{K_t+K_h} \qquad (10)$$

has been modified to additionally include the elasticity of the holdfast in our experiments. The scaling of the rupture force with N_0 has been determined in fast and slow loading regimes in the limits of soft and stiff transducers.

For the experiments reported in this work, we estimate the loading parameter μ as follows: The cantilever retraction rate in the experiments reported here is $v \approx 2\ \mu m/s$, $k_BT = 4\ pN \cdot nm$ at room temperature, and we estimate $x_p \sim 1\ nm$ [9]. From numerical fits of simulations to the AFM rupture force data reported above, we take the rate $k_0 \sim k_d \sim 10^{-4}\ s^{-1}$ (see Figure 12A and Figure S2). Hence, we can write the loading parameter as

$$\mu \sim (5 \times 10^6\ nm \cdot pN^{-1})\ \frac{K_t K_h}{K_t + K_h} \quad (11)$$

In the limit that $K_t \ll K_h$ or $K_t \sim K_h$, we can approximate $\frac{K_t K_h}{K_t + K_h} \approx K_t = 25\ pN/nm$, giving $\mu \sim 10^8$. For $K_t \gg K_h$, we have $\mu \sim (5 \times 10^6\ nm \cdot pN^{-1})\ K_h$. From numerical simulations reported in Figure 12A and Figure S2, the number of bound adhesins does not exceed 10^3. Hence, for plausible values of K_h, we expect $\mu > N_0$, putting our experiments in the fast loading regime.

In the fast loading regime, it has been shown that the relevant scaling of the rupture force with N_0 is linear (with logarithmic correction) [5] [8]. This scaling is the justification for taking the rupture force to be linearly proportional to the surface concentration of adhesin in Section 1. Furthermore, the rupture force, F^*, is found to increase (logarithmically) with the loading parameter, μ

$$F^* = (k_B T / x_b)\ N_o \log(\mu / N_0) \quad \text{for } N_0 \ll \mu \quad (12)$$

We note from Eq. (9) that the loading parameter μ increases with increasing holdfast spring constant, K_h, leading to an increase in the rupture force for the same number of initial surface bonds. A possible mechanism for a time-dependent increase in the holdfast stiffness, $K_h(t)$, is cross-linking of the holdfast polymer network, akin to actin-filamin gels [10]. In this scenario, in addition to binding the holdfast to the surface, the adhesin also cross-links the holdfast matrix over time. The stiffening of the holdfast matrix ensures that an applied load is uniformly distributed over the multiple parallel bonds with the surface. If this were not the case, then the force on some bonds for a given applied load would be larger, leading to a greater probability of rupture according to Eq. (8), resulting in a cascade of ruptures with shorter rupture time and therefore smaller rupture force. Indeed, in the large load regime, the rupture time is determined to be inversely proportional to μ [5].

The cross-linking of the holdfast matrix can be achieved by the putative cross-linker/adhesin or alternatively through chemical modification of *N*-acetylglucosamine as described above. Regardless of how cross-linking is achieved, it provides an alternative mechanism for generating the separation of time scales leading to smaller adhesion forces on short time scales and large forces on longer time scales. In this scenario, the strength of adhesion is additionally derived from stiffening of the holdfast, thereby uniformly distributing an applied load among multiple parallel bonds with the surface.

2. Analysis of Rupture Event Force Distributions

2.1 Identifying Rupture Events from AFM Retraction Curves

The majority (~60%) of the DFS retraction curves are characterized by a single contact-rupture event as in Figure S3A, the others containing several local minima corresponding to partial rupture events, as demonstrated in Figure S3B. To perform a quantitative statistical analysis of the rupture events, we identified rupture events in two ways: i) by visual inspection (carried out independently by two experimenters), and ii) based on an automated procedure of identifying local maxima/minima in the AFM retraction curves. In the latter approach, the first step is to parse the curve into a rupture portion (when the holdfast adhered to the AFM tip is in contact with the surface), shown in red in Figure S3B, and a detached portion (after the holdfast on the AFM tip is completely detached from the surface), shown in blue. The rupture portion of the retraction curve may contain many local minima, as seen in Figure S3B. The algorithm identifies all of the local minima m_i, circled in black, and their corresponding local maxima M_i, circled in green. The portion of the AFM curve connecting m_i to M_i is classified as a candidate rupture event with a rupture force of magnitude $\Delta F_i = F(M_i) - F(m_i)$. The true rupture events are selected from the set of candidate events by setting a force threshold, which is a consequence of instrumental noise.

In this analysis, we apply two different methods of setting a force threshold, a constant force threshold and a variable force threshold. For the constant force threshold, the true rupture events are selected from the candidate events provided they satisfy the criterion $\Delta F_i > \Delta F_{crit}$, where ΔF_{crit} is a constant force threshold, fixed for all AFM retraction curves. Typical choices of $\Delta F_{crit} = 5 - 25$ pN are consistent with threshold values used in previous works [9]. This thresholding truncates the lower half of force histograms up to the threshold values. We find that the fit results to the resulting force distributions are little affected by the choice of ΔF_{crit} within the range given above. For the variable force threshold, we first determine the variability in each individual experiment by calculating the standard deviation of the force σ_i for the detached portion of each individual DFS retraction curve. Across the 157 different curves, this value σ_i varies between 3 and 10 pN, as demonstrated in Figure S3C. True rupture events are selected from the candidate events provided they satisfy the criterion $\Delta F_i > m\sigma_i$, where m is a numerical parameter

which is held constant across all curves, and σ_i is the standard deviation for curve *i*. In words, only candidate rupture events whose magnitude is greater than m standard deviations are classified as true ruptures. We find that for $m = 2 - 4$, yielding force thresholds that are consistent in value to the constant thresholds used in the former method, the fit results are little affected by the choice of *m*.

2.2 Analysis of Rupture Force Distributions

Distributions of rupture forces obtained using the methods described above were analyzed in the context of two semi-empirical hypotheses. The first hypothesis considers a given rupture event to correspond to the breaking of an integer number of adhesin bonds with the substrate, each with a common, characteristic adhesion force. Assuming this characteristic force is larger than the AFM resolution, rupture force histograms would feature a series of regularly spaced peaks. Visual inspection of the data suggested the possibility of such structure with a peak spacing of approximately 30 pN. One example, obtained from the algorithmic rupture event identification approach, is plotted in Figure S4. We have carried out a series of fits to assess the statistical compatibility of the data with this intriguing hypothesis. To interpret the resulting goodness-of-fit metrics, we compared with results obtained considering a second, “null” hypothesis. The functional form taken for this hypothesis, the sum of two falling exponential distributions, was motivated by inspection of semi-logarithmic plots of the rupture force histograms, for example as shown in the right panel of Figure S4. Two roughly linear regions (from 25 to 120 pN, and from 120 to 400 pN) can be identified, spanning larger ranges of rupture forces than the separation between the putative peaks: the question is whether the peaks (and intervening troughs) are consistent with statistical fluctuations relative to some smooth underlying dependence. A sum of two exponential distributions would appear to provide a viable description of the main trends in the data if one ignores the peak-like indications.

The main results reported here are based on analysis of rupture force histograms in 5 pN-wide bins over a range from 25 to 250 pN. Analyses employing other bin widths and fit ranges gave consistent results, as shown below. Accounting for the limited number of rupture events per bin, we have employed a maximum likelihood approach.

This approach naturally incorporates the Poisson errors intrinsically present in binned data under the assumption that rupture events are mutually independent. We report as the goodness-of-fit metric the value of the −2ln $\mathcal{L}$ quantity being minimized, where $\mathcal{L}$ is the product of the per-bin likelihoods that represent the Poisson probability that the fit function value for a given bin has fluctuated to yield the number of rupture events actually observed,

$$\mathcal{L} = \Pi_i \frac{\mu_i^{N_i} e^{-\mu_i}}{N_i!} \quad (13)$$

where μ_i is the value of the fit function (rupture force) at the *i*th bin, and N_i is the number of rupture events in this bin. For the case of the multiple-peak hypothesis, we assume:

1. Rupture events occur in integer multiples of a characteristic adhesion force.

2. Noise in the AFM data introduces systematic smearing of the measured rupture forces, assumed to be Gaussian distributed and with a constant width over the range of measured forces (i.e., independent of the number of ruptured bonds).

3. Additionally, there could be some background level of incorrectly inferred rupture forces, assumed to be uniformly distributed.

The nominal fit function consists of six Gaussians, each with a mean corresponding to a particular integer (n = 1–6) multiple of the characteristic rupture force, and with identical widths, plus a flat "background". With this, there are 9 free parameters: 7 scale factors A_n for the six Gaussians and background level, one parameter $\bar{x}$ representing the mean characteristic rupture force, and one parameter σ for the common root-mean-square width of the Gaussians. Thus, each Gaussian takes the form:

$$G_n(x_i;\, A_n, \bar{x}, \sigma) = \frac{A_n}{\sqrt{2\pi}\sigma}\, exp\left(\frac{(x_i - n\bar{x})^2}{2\sigma^2}\right) \quad (14)$$

where x_i represents the force (in pN) of the *i*th bin of the histogram, and n = 1, ..., 6.

We have varied the number of Gaussians as well as the form of the background function, obtaining consistent results for the common parameters. Despite the large number of free parameters in the nominal six-Gaussian plus flat background function, fits

are well-behaved, with roughly parabolic $-2 \ln \mathcal{L}$ surfaces in parameter space out to (at least) several standard deviations relative to best-fit parameter values, as is illustrated below. The null-hypothesis function has four free parameters. These are the scale factor and slope for each of the two exponential functions.

2.3 Results from fits to multiple-peak hypothesis

The fit of the multiple-peak function to the algorithmically derived rupture force histogram obtained with a constant 20 pN threshold is presented in the main text in Figure 9A. For comparison, in Figure 9B we display the corresponding fit to the rupture force histogram obtained from visual identification of rupture event candidates. Results obtained from a number of fits for key parameter and goodness-of-fit values are presented in Table S1, with the two cases plotted in Figure 9 entered in the first two rows.

In general, good agreement is found among the different analyses, with a best fit value for the characteristic rupture force of approximately 30 pN, and common widths of the Gaussian peaks of approximately 9 pN. For the purposes of reporting a single result in the main text, we have taken the analysis shown in the first row of the table as the nominal one. Considering the variations in values across the different analyses, we estimate a systematic uncertainty of 0.7 pN on the characteristic rupture force, which, when combining with the statistical uncertainty gives an overall uncertainty of 1.0 pN.

2.4 Robustness of fits to multiple-peak hypothesis

In assessing further the robustness of the multiple-peak analysis, several points can be made. Despite the large number of fit parameters, it is notable that the constraint embedded within the fit function that the peaks are equally spaced with common widths is a strong one. Furthermore, the peaks are well separated, limiting the degree of correlation among fit parameters. Consequently, best-fit parameter values are found to be insensitive to the number of peaks included and the range over which the fit is performed.

To further illustrate the robust nature of the fit procedure and results, in Figure S5 we plot the value of $-2\ ln\ \mathcal{L}$ versus x from a scan over this parameter in which the fit was redone at each point with all other parameters floating. The sharp minimum around 30 pN, and the parabolic shape in the vicinity of the minimum are as expected given the statistical uncertainties reported by the fit.

2.5 Results from fits to null hypothesis

For all fits performed with the multiple-peak hypothesis, values for goodness-of-fit metrics are within acceptable ranges, as indicated by the correspondence between the $-2\ ln\ \mathcal{L}$ values and the number of degrees of freedom. However, by itself this finding does not exclude the possibility that alternate hypotheses, without peaks, could explain the DFS data. In this section we address to what extent the multiple-peak structure suggested by the data is present at a statistically significant level by comparing with fits to the double-exponential function described earlier.

In Figure S6, we present results that correspond to those shown in Figure 9, but with the double-exponential fit function. Inspection of the distributions of residuals indicates that the double-exponential function describes well the rupture force histogram obtained from the algorithmic analysis, but not so well that obtained from the visual-identification analysis. This observation is quantified in Table S2, which compares goodness-of-fit metrics.

Two sets of *p*-values are given in the table. First, given the best-fit double-exponential function for a given rupture force histogram, one can ask what fraction of experiments would give a $-2\ ln\ \mathcal{L}$ value larger than what was actually observed. This was addressed by simulating 10,000 toy experiments for each of the two histograms (visual and algorithmic rupture-event identification), and histogramming the $-2\ ln\ \mathcal{L}$ values from these experiments, as shown in Figure S7. As suggested by the residual plots in Figure S6, the null hypothesis cannot be excluded for the algorithmic analysis, with a more than healthy p-value of 88.7%, nor for the visual-identification analysis, with a less likely, but still reasonable p-value of 4.7%.

However, the multiple-peak hypothesis is favored in both analyses, and the

second set of p-values in the table is obtained from application of the well-known likelihood-ratio test for discrimination between hypotheses. Applying this test to the two hypotheses, one would expect the difference in $-2\,ln\,\mathcal{L}$ values to be distributed according to the χ^2 probability distribution for 5 degrees of freedom (the difference in number of fit parameters) if both hypotheses are equally likely. The *p*-values for this test, 7.25% for algorithmic analysis and 0.01% for the visual-identification analysis show some discrimination. While the former value is not statistically significant, the latter indicates that the visual-identification histogram strongly disfavors the null hypothesis relative to the multiple-peak hypothesis. (It is important to note however that this application of the likelihood ratio test is approximate in this case, since the parameters of the double-exponential function do not constitute a subset of those of the multiple-peak function.)

From these studies, we conclude that the visual-identification analysis yields a strong indication for the multiple-peak hypothesis as providing the best description of the DFS data. While we prefer the algorithmic approach as being most reproducible and amenable to testing, we are unable to imagine any bias in the visual-identification procedure that would artificially give rise to a multiply-peaked structure in the rupture force distribution. The facts that goodness-of-fit metrics for the algorithmic analysis also favor the multiple-peak hypothesis and that it gives results for fit parameters that are consistent with those from the visual-identification analysis provide additional support for this hypothesis.

S**UPPLEMENTARY TABLES**

Table S1: Results from fits of rupture force histograms to multiple-peak hypothesis. The histograms fitted are obtained from the DFS data based on algorithms employing different thresholds or on the visual identification of rupture events, or constructed with different bin widths or ranges. The notation d.o.f. refers to the number of degrees of freedom in the fit (number of bins minus number of fit parameters). The results presented include best- fit values for the characteristic rupture force $\bar{x}$ and common Gaussian width σ with their one-standard deviation uncertainties, as well as the best-fit value of $-2\ ln\ \mathcal{L}$. The notation (a) refers to a histogram generated with a version of the constant threshold algorithm, that combines rupture event candidates occurring close in time.

Bin size	Threshold	Range (pN)	d.of.	$\bar{x}$ (pN)	σ	$-2\ ln\ \mathcal{L}$
5 pN	20 pN	25-250	36	29.7 ± 0.6	9.1 ± 0.8	28.6
5 pN	visual	25-250	36	30.5 ± 0.4	8.4 ± 0.6	41.7
5 pN	20 pN	25-400	66	29.8 ± 0.5	9.1 ± 0.8	59.2
5 pN	visual	25-400	66	30.5 ± 0.4	8.5 ± 0.6	67.9
5 pN	15 pN	25-250	36	29.7 ± 0.6	9.1 ± 0.8	28.6
5 pN	25 pN	25-250	36	29.9 ± 0.6	9.7 ± 1.0	31.9
5 pN	20 pN (a)	25-250	36	29.8 ± 0.8	10.3 ± 1.4	40.8
3 pN	20 pN	24-252	67	28.6 ± 0.8	9.0 ± 0.9	62.9
4 pN	20 pN	24-252	49	28.6 ± 0.7	8.9 ± 0.9	55.1

Table S2: Goodness-of-fit metrics for fits to the different hypotheses shown in Figure 9 and Figure S6. The $-2\ ln\ \mathcal{L}$ values correspond to identical fit ranges and bin-widths, leading to 36 degrees of freedom for the multiple-peak hypothesis and 41 degrees of freedom for the null hypothesis. Entries under the heading "*p*-value for Null Hypothesis" are determined from simulations of 10,000 experiments drawn from a parent distribution governed by the corresponding best-fit double-exponential function. Entries under the heading "Likelihood Ratio *p*-value" are determined assuming the difference in $-2\ ln\ \mathcal{L}$ values for the two hypotheses would be governed by a χ^2 probability distribution corresponding to 5 degrees of freedom (equal to the difference between 41 and 36).

Rupture Event Identification Method	Best-fit ($-2\ ln\ \mathcal{L}$) Multiple-Peak	value Null	*p*-value for Null Hypothesis	Likelihood Ratio *p*-value
Algorithmic	28.6	38.5	88.7%	7.25%
Visual	41.7	66.9	4.7%	0.01%

SUPPLEMENTARY FIGURE LEGEND.

Figure S1: The dimensionless concentrations as a function of time. The parameter values of the fit are: h = 60 nm, k = 7.2 × 10^{-3} nm^2 s^{-1}, k_a = 1.5 × 10^{-1} nm s^{-1}, k_d = 0.2 s^{-1}, S_0 = 0.56 nm^{-2}, P_0 = 1.5 × 10^{-2} M , τ = 5 s.

Figure S2: The rupture force as a function of dwell time, as measured by DFS (circles), and in the model (solid line). The parameter values of the fit are: h = 60 nm, k = 0 nm^2 s^{-1}, k_a = 240 nm s^{-1}, k_d = 2.0 × 10^{-4} s^{-1}, S_0 = 1.0 nm^{-2}, P_0 = 4.6 × 10^{-3} nm^{-3}, τ =13.8 h, δf_1 = 12 pN. Note that in this slow diffusion scenario, since k = 0 no $P - S$ is ever formed, and the term proportional to δf_2 does not contribute to the adhesion force. The mean force as measured by Tang *et al.* reduced by a factor of 4 is shown as a triangle, which corresponds to $h_S/h = 2$. The vertical line indicates the range of measured values.

Figure S3: a. The force (N) as a function of the height of the AFM tip z (m) for a sample retraction curve. This curve is characterized by a single contact-rupture event. b. The force (N) as a function of the height of the AFM tip z (m) for a sample retraction curve. The curve is parsed into a rupture portion, when the holdfast adhered to the AFM tip is in contact with the surface (shown in red), and a detached portion, when the holdfast adhered to the AFM tip is no longer in contact with the surface (shown in blue). This curve is characterized by multiple rupture events. Each individual rupture event corresponds to the portion of the trace between the local minima (black circles) and the local maxima (green circles). c. The standard deviation of the force (N) for different experimental trials. The standard deviation of the force is calculated using the deflection of the AFM tip for the detached portion of the trace, shown in blue in (b).

Figure S4: Histogram of rupture forces inferred from the algorithmic analysis of the DFS data with a fixed 20 pN threshold. a. linear scale. b. logarithmic scale.

Figure S5: Scan of characteristic rupture force values. Plot of $-2\,ln\,\mathcal{L}$ values (filled circles) from fits with fixed values of $\bar{x}$, scanning over the region near the best fit value of 29.8 pN for the case of algorithmic identification of rupture events with 20 pN threshold, 5 pN bin width, and fit range of 25–400 pN. The curve overlaid is a fit of the interior six points to a parabola.

Figure S6: Plots showing fits of rupture event histograms to the null (double-exponential) hypothesis. The lower panel of each image displays the histogram of rupture-event forces, corresponding to the algorithmic (20 pN threshold, a) and visual (b) identification of rupture events in the DFS force-displacement data. The corresponding best-fit function comprising a sum of two exponential functions is overlaid. The fits were performed over the range 25 to 250 pN. The upper panel of each image displays the normalized residuals, computed as (data−fit)/fit).

Figure S7: Distributions of $-2\ ln\ \mathcal{L}$ values for simulated null-hypothesis□actual values from the corresponding fits to data are 38.5 for the algorithmic analysis and 66.9 for the visual-identification analysis.

SUPPLEMENTARY REFERENCES

(1) Grandbois, M.; Beyer, M.; Rief, M.; Clausen-Schaumann, H.; Gaub, H. E. How Strong Is a Covalent Bond? *Science* **1999**, *283*, 1727-1730.

(2) Tsang, P. H.; Li, G.; Brun, Y. V.; Freund, L. B.; Tang, J. X. Adhesion of Single Bacterial Cells in the Micronewton Range. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 5764-5768.

(3) Marszalek, P. E.; Dufrene, Y. F. Stretching Single Polysaccharides and Proteins Using Atomic Force Microscopy. *Chem. Soc. Rev.* **2012**, *41*, 3523-3534.

(4) Hoffmann, T.; Dougan, L. Single Molecule Force Spectroscopy Using Polyproteins. *Chem. Soc. Rev.* **2012**, *41*, 4781-4796.

(5) Seifert, U. Rupture of Multiple Parallel Molecular Bonds under Dynamic Loading *Phys. Rev. Lett.* **2000**, *84*, 2750-2753.

(6) Bell, G. I. Models for the Specific Adhesion of Cells to Cells. *Science* **1978**, *200*, 618-627.

(7) Evans, E.; Ritchie, K. Dynamic Strength of Molecular Adhesion Bonds. *Biophys. J.* **1997**, *72*, 1541-1555.

(8) Seifert, U. Dynamic Strength of Adhesion Molecules: Role of Rebinding and Self-Consistent Rates. *Europhys. Lett.* **2002**, *58*, 792.

(9) Getfert, S.; Reimann, P. Hidden Multiple Bond Effects in Dynamic Force Spectroscopy. *Biophys. J.* **2012**, *102*, 1184-1193.

(10) Kang, H.; Wen, Q.; Janmey, P. A.; Tang, J. X.; Conti, E.; MacKintosh, F. C. Nonlinear Elasticity of Stiff Filament Networks: Strain Stiffening, Negative Normal Stress, and Filament Alignment in Fibrin Gels. *J. Phys. Chem. B* **2009**, *113*, 3799-3805.