

Single-molecule force spectroscopy studies of integrin-mediated cell signaling

Jan Opfer



Munich, July 18, 2012

Single-molecule force spectroscopy studies of integrin-mediated cell signaling

Dissertation

Faculty of Physics

Ludwig Maximilian University of Munich

submitted by

Jan Opfer

Munich, July 18, 2012

1st referee: Prof. Dr. Kay-Eberhard Gottschalk
2nd referee: Prof. Dr. Joachim Rädler
Day of the defense: September 12, 2012

Contents

Zusammenfassung	9
Abstract	11
1. Introduction.....	13
1.1. Atomic force spectroscopy.....	13
1.2. Pushing the force resolution limit.....	15
1.3. Integrins.....	16
2. Noise reduction.....	21
2.1. Materials and methods	24
2.1.1. Stationary wavelet transform	24
2.1.2. The ReNoiR algorithm	24
2.1.3. Automatic parameter optimization	29
2.1.4. Evaluation of the signal recovery performance.....	30
2.1.5. Estimate of the noise reduction performance for fundamental signal forms.....	31
2.1.6. Analysis of the tether onset force.....	34
2.1.7. Analysis of the SERS spectra	34
2.1.8. Analysis of the optical tweezers measurements	34
2.2. Results	35
2.2.1. Synthetic data	35
2.2.2. Experimental data	43
2.2.3. Ideal filter parameters	46
2.3. Discussion.....	49
3. Step detection.....	51
3.1. Materials and methods	52
3.1.1. Detection method of Kerssemakers et al.	52
3.1.2. MSF algorithm.....	52
3.1.3. Generation of the test signals.....	53
3.1.4. Evaluation of the step detection algorithms	56
3.1.5. Calculation of the step heights	57
3.1.6. Recording of the AFM spectra	57
3.2. Results and discussion.....	57
3.2.1. Step detection performance.....	57
3.2.2. Distribution of false-positives and -negatives	63
3.2.3. Height resolution	65
3.2.4. Reproduction of continuous height distributions.....	69
3.2.5. Computational cost.....	71
3.2.6. Analysis of AFM force spectra.....	71
3.2.7. Analysis of kinesin motor experiments.....	73
3.2.8. Parameter optimization	74
4. Effect of SDF-1α on the mechanics of the integrin $\alpha_4\beta_1$ environment	77
4.1. Materials and methods	77
4.1.1. Reagents.....	77

4.1.2.	Lymphocytes	77
4.1.3.	Substrate preparation	78
4.1.4.	AFM measurements	78
4.1.5.	Data analysis	79
4.1.6.	FACS analysis	80
4.2.	Results	80
4.2.1.	VLA-4 expression	80
4.2.2.	Adhesion rates	81
4.2.3.	Forces	84
4.2.4.	Step positions	85
4.2.5.	Lymphocyte stiffness	86
4.3.	Discussion	88
5.	Role of the trans-membrane domains in integrin-mediated adhesion and signaling	93
5.1.	Materials and methods	94
5.1.1.	In vitro site-directed mutagenesis	94
5.1.2.	Cell culture	94
5.1.3.	Atomic force spectroscopy	94
5.2.	Results	94
5.3.	Discussion	95
6.	Conclusions	103
7.	Appendix	105
7.1.	Curvalyser reference manual	105
7.1.1.	Basic configuration	105
7.1.2.	Program execution	106
7.1.3.	Baseline correction	107
7.1.4.	Noise reduction	108
7.1.5.	Contact point calibration	108
7.1.6.	Step detection	109
7.1.7.	Fitting procedures	110
7.1.8.	Output files	110
7.1.9.	Overview of important configuration parameters	111
7.2.	Paramalyser reference manual	115
7.2.1.	Configuration	115
7.2.2.	Program execution	117
7.2.3.	Tasks	118
7.2.4.	Parameters	119
References		121
Publications		131
Acknowledgment		133

Zusammenfassung

Integrine stellen eine wichtige Klasse von Zelladhäsionsrezeptoren dar, denn sie übermitteln Informationen bidirektional zwischen dem Zytoplasma biologischer Zellen und der sie umgebenden extrazellulären Matrix. Mittels Rasterkraftmikroskopie wurden spektroskopische Messungen der spezifischen Wechselwirkungen zwischen Integrinen und ihren zugehörigen Liganden durchgeführt. Prinzipiell ermöglichen diese Experimente tiefe Einblicke in die zelluläre Signaltransduktion, aber trotz aufwändiger Systeme zur Isolierung gegen Vibrationen weisen die aufgenommenen Daten sehr geringe Signal-Rausch-Verhältnisse auf, die eine exakte Auswertung beeinträchtigen. Diesem Nachteil wurde mit einem neuartigen Nachbearbeitungsverfahren begegnet, das das Rauschen signifikant reduziert und so das Signal-Rausch-Verhältnis verbessert. Damit können zuvor unsichtbare Merkmale der Signale erkennbar gemacht werden.

Ein weiterer wichtiger Arbeitsschritt bei der Auswertung dieser Experimente besteht in der Identifizierung stufenförmiger Übergänge, die Abrissen der Rezeptor-Ligand-Bindungen entsprechen. Zu diesem Zweck wurde eine Methode entwickelt, die daran angepasst werden kann, sehr niedrige oder schmale Stufen zu erkennen, auch wenn sie weich verlaufen und vom Rauschen verborgen sind.

Durch Anwendung des Rauschunterdrückungs-Algorithmus auf Kraftspektroskopie-Daten, die mit lebenden T-Lymphozyten aufgenommen wurden, konnte erstmalig die initiale Kraft beobachtet werden, die für die Extraktion eines Membranschlauchs erforderlich ist. Mit der Stufenerkennungsmethode wurde ein deutlicher Hinweis auf sub-10-pN-Stufen gefunden. Es wurde zudem gezeigt, dass das Chemokin SDF-1 α zu einer Verstärkung der einzelnen Bindungen zwischen VLA-4 – einem Integrin-Typ, der in erster Linie bei der Anfangsphase der Chemokin-induzierten Adhäsion von Lymphozyten beteiligt ist – und dessen Liganden VCAM-1 führt. Die Verstärkung der Adhäsion wird von einer Versteifung der Umgebung der Integrine begleitet. Sie ist unabhängig von einer intrazellulären Bindungsstelle von VLA-4 für Talin, dem für die Affinitätsregulierung von Integrinen wichtigsten intrazellulären Faktor.

Darüber hinaus wurde die Funktion der Transmembran-Domäne von Integrinen bei Rezeptor-Ligand-Wechselwirkungen durch Analyse der Auswirkungen von zwei Mutationen des Integrins $\alpha_v\beta_3$ auf die zelluläre Adhäsion untersucht: eine Chimäre, die die stark dimerisierende Transmembran-Domäne von Glycophorin A enthält, und eine Punktmutation, die bekanntermaßen

zur Dissoziation der Transmembran-Domäne führt. Die Ergebnisse zeigen, dass beide Konstrukte starke Zelladhäsion hervorrufen. Sie stimmen gut mit einem Drei-Zustands-Modell der Integrin-Aktivierung überein. Ein Ruhezustand wird durch intrazelluläre Liganden zu einem Zwischenzustand aktiviert ohne dass die Transmembran-Domäne separiert wird. Die dimerisierende Chimäre imitiert den Zwischenzustand, der die zelluläre Adhäsion verstärkt.

Abstract

Integrins constitute an important class of cell adhesion receptors, as they bidirectionally transduce information between the cytoplasm of biological cells and the surrounding extracellular matrix. By means of atomic force microscopy, spectroscopic measurements of the specific interactions of integrins with their corresponding ligands were performed. Basically, these experiments allow deep insights into cellular signal transduction, but despite sophisticated vibration isolation systems the acquired data exhibit very low signal-to-noise ratios that impair an accurate analysis. This drawback was overcome by a novel post-processing algorithm, which significantly reduces the noise and thus improves the signal-to-noise ratio. Thereby, previously invisible signal features can be revealed.

Another important task when evaluating this kind of experiments is the identification of steplike transitions corresponding to unbinding events between the receptor-ligand bonds. To this end, a technique has been developed that can be adjusted to detect very low or narrow steps even if they are smooth and hidden by noise.

By applying the noise reduction algorithm to force spectroscopy data obtained with living T lymphocytes, the onset force required for the extraction of a membrane tether could be observed for the first time. Using the step detection method, strong evidence of sub-10-pN steps was found. Moreover, it was shown that the chemokine SDF-1 α leads to a strengthening of individual bonds between VLA-4, one type of integrins primarily involved in the early stages of chemokine-induced lymphocyte adhesion, and its ligand VCAM-1. The adhesion strengthening is accompanied by a stiffening of the integrins' environment. It is independent of an intracellular binding site of VLA-4 to talin, the major intracellular factor involved in integrin affinity modulation.

Further, the functional role of the integrin trans-membrane domains in receptor-ligand interactions was explored by analyzing the effects of two mutations of the integrin $\alpha_v\beta_3$ on cellular adhesion: a chimera encompassing the strongly dimerizing trans-membrane domain of glycophorin A and a point mutation known to induce trans-membrane domain dissociation. The results show that both constructs provoke strong cell adhesion. They correspond well to a three-state model of integrin activation. A resting state is activated by intracellular ligands to an intermediate state without trans-membrane domain separation. The dimerizing chimera mimics the intermediate state, which strengthens cellular adhesion.

1. Introduction

1.1. Atomic force spectroscopy

To perform single-molecule force spectroscopy measurements with living cells, state-of-the-art biophysical techniques like atomic force microscopy¹ (AFM) or optical/magnetic tweezers²⁻⁴ are required. They are highly sensitive tools revealing insights into biological processes on the molecular level. The former was deployed in this work, as it allows for more direct control on force and position when probing the interactions between cell receptors and their interaction partners (ligands). Further, AFM is ideally suited to measure interactions between cells and functionalized surfaces presenting a well-defined molecular environment.

Although AFM was originally used for imaging only, the technique can alike be applied to force spectroscopy experiments. To study the receptor molecules on the surface of a living cell, it is attached to the AFM force sensor (cantilever) and brought into contact with the receptor-specific ligands, which are immobilized on a substrate at a well-defined surface density (fig. 1).

The kinetics of the receptor-ligand interactions can be characterized by their life times and unbinding (rupture) forces (fig. 2). To obtain expressive statistics, a large number of force curves must be acquired and analyzed. Therefore, an automated method is necessary to extract these parameters and guarantee constant, objective ratings. The example in fig. 2 obviously contains only one easily identifiable rupture event (step), but low steps hidden by noise cannot be detected manually. However, computer-aided techniques allow for a much higher force resolution, which is limited by the signal-to-noise ratio (SNR) of the recorded signals.

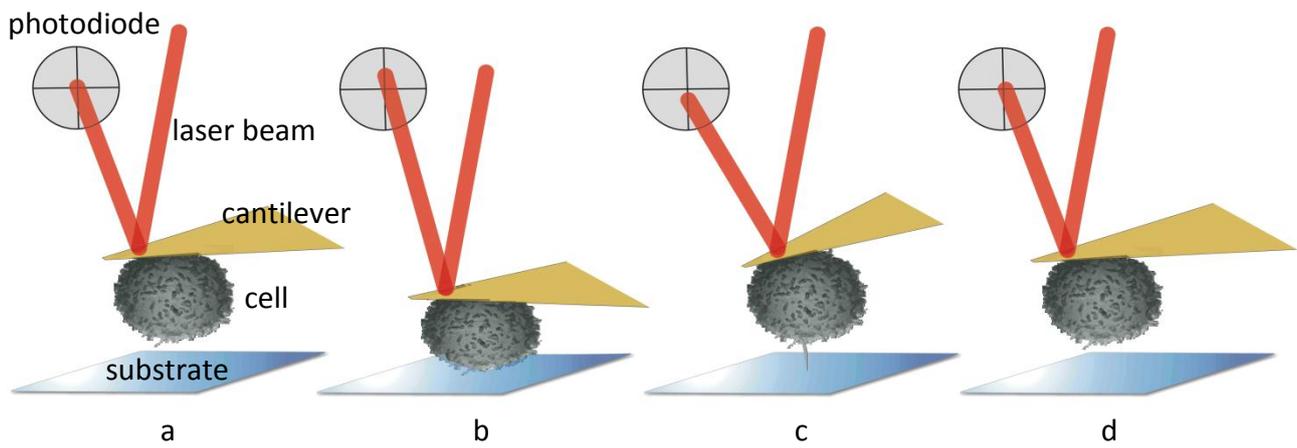


Fig. 1: Force spectroscopy of living cells using an atomic force microscope (modified from Schmitz⁵). Interactions between cell and substrate cause a deflection of the force sensor (cantilever), which results in a change in angle of reflection of a laser beam focused on the cantilever's upper side. The angular change is converted into a voltage signal proportional to the force by a segmented photodiode and recorded electronically. a: The cantilever with a living cell attached to it is approached with constant velocity to a functionalized surface presenting a specific ligand. b: The cell is brought into contact with the substrate at constant indentation force for a given dwell time, allowing the receptors exposed on the cell surface to establish intermolecular bonds with the corresponding ligands on the substrate. c: While the cantilever is retracted, a tether is pulled out of the cell membrane, which exerts a constant force on the cantilever. d: The receptor-ligand interactions are broken and the cantilever is further retracted.

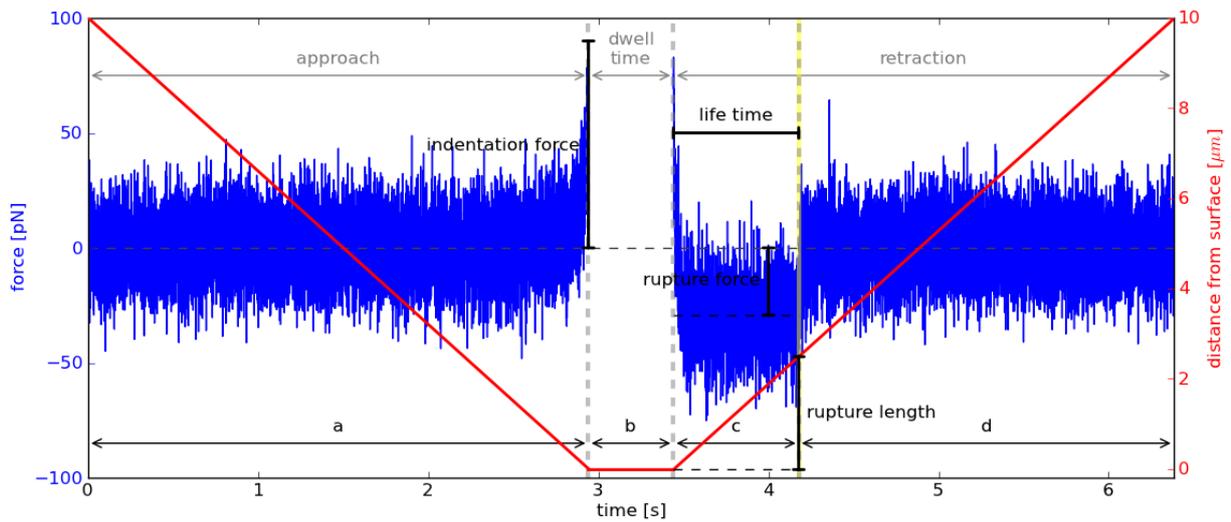


Fig. 2: Typical force curve (blue) resulting from the approach-retraction cycle (red) illustrated in fig. 1. The cycle is repeated many times. For each unbinding event (yellow line), the life time and rupture force are determined automatically and recorded for statistical analysis.

1.2. Pushing the force resolution limit

Basically, two strategies are possible to enhance the force resolution: On the one hand, the noise can be tackled at its source. Most sources of noise in the measurements described above (such as electronic noise or vibrations) may be significantly reduced, e.g. by acoustic damping or vibration isolation systems. In contrast, Brownian noise cannot be cut down at all, as this would mean to lower the temperature and thereby affect the kinetics of the observed organism. On the other hand, as these disturbances cannot be eliminated completely, only post-processing techniques come into question to further push the resolution limit. More particularly, two different approaches were put to the touch: First, the noise was reduced to increase the SNR and make low steps visible. Second, the steps were detected directly within the original signal.

A comparison of existing noise reduction methods has shown that stationary wavelet transform^{6, 7} (SWT) is the most efficient algorithm in the sense of minimum deviation between noisy and clean signal (quantified by the root mean square error). Unfortunately, the de-noised signals contain spike-shaped artifacts (known as pseudo-Gibbs phenomena), which arise from filtering out small wavelet coefficients required to represent the high-frequency characteristics of discontinuities⁸.

Filtering only larger coefficients increases the root mean square error, i.e. the artifacts constitute an inevitable side effect. Although they are significantly less pronounced than those resulting from conventional noise reduction based on discrete wavelet transform (DWT), they still heavily impair the detection of steps. Therefore, a better approach was needed that would both yield comparable or lower root mean square errors than SWT and reduce the pseudo-Gibbs phenomena. To this end, a novel noise reduction algorithm was developed (see section 2). By averaging a set of recursively expanded wavelet coefficients, it is able to reduce much of the noise and thereby reveal important features of the recorded force spectra, which are initially hidden. For example, a tether onset force of human T lymphocytes could be made visible for the first time.

If a direct visual control of the rupture events is not necessary, the step detection can also be accomplished directly without increasing the SNR beforehand. This was put into practice by a novel step detection algorithm that is ideally suited to locate steplike features separating adjacent plateaus, even if they are smooth and hidden by noise (see section 3). It can be adjusted to detect very low or narrow steps that cannot be recognized by conventional methods.

1.3. Integrins

Both algorithms were deployed to analyze single-molecule force spectra of the interactions between integrins, force-exposed receptors located on the surface of biological cells, and their corresponding ligands. Cellular adhesion and migration across tissue boundaries, and the sensing of mechanical properties of the extracellular micro-environment are fundamental biological events. Integrins are the main receptors involved in these tasks. They are heterodimers composed of one α and one β subunit, each of which encompasses a large extracellular domain, a single-pass trans-membrane domain (TMD) and a short cytoplasmic tail⁹⁻¹¹ (fig. 3). Up to now, 24 different dimers have been identified in mammals, combined from 18 α and 8 β chains¹². By mediating attachment to the extracellular matrix (ECM) or to other cells, and by transducing signals, integrins control important cellular functions¹³⁻¹⁵. Signal transduction operates bidirectionally: Information about the ECM, such as the ligands' type and state, can be transmitted across the membrane, e.g. to regulate growth, proliferation, migration, apoptosis, or cell differentiation (outside-in signaling)^{16, 17}. Conversely, intracellular signals can also be forwarded to the cell's environment (inside-out signaling). By conformational activation of the receptors, their affinity can be increased¹⁸⁻²⁰. This is particularly important to control processes like inflammation and immune

response^{16, 21-23}. Complex cellular functions, such as those involved in hemostasis, rely on both directions of transmission²⁴.

In the ground state, integrins are non-adhesive, but they are activated (i.e. become binding-competent to ECM ligands) subsequent to receiving cellular signals²⁵. A putative signaling pathway leading to integrin activation is initiated by extracellular ligation, e.g. of the chemokine SDF-1 α (CXCL12), to a G-protein-coupled receptor (GPCR), such as CXCR4^{20, 26, 27} (fig. 4). Chemokines (chemotactic cytokines) are signaling proteins that induce chemotaxis, i.e. the directed movement of cells along a concentration gradient. GPCRs are integral membrane proteins consisting of an extracellular N terminus, seven trans-membrane α -helices connected by six intra- and extracellular loops, and an intracellular C terminus²⁸. The ligation induces a conformational change in the GPCR. As a consequence, the coupled G-protein trimer is activated and its α subunit dissociated. The remaining β - γ -complex activates phospholipase C (PLC), which cleaves phosphatidylinositol-4,5-bisphosphate (PtdIns(4,5)P₂) into inositol-1,4,5-trisphosphate (InsP₃) and diacylglycerol (DAG). InsP₃ triggers the release of Ca²⁺ ions from the endoplasmic reticulum (ER), which again increases calcium influx from the extracellular space through calcium-release activated calcium (CRAC) channels. Ca²⁺ and DAG are supposed to activate guanine-nucleotide-exchange factors (GEFs). Subsequently, guanosine triphosphate hydrolase enzymes (GTPases) may be triggered. They presumably induce the binding of adaptor proteins (e.g. talin) to the cytoplasmic tails of the integrins, which are thereby activated^{25, 29}.

Besides chemical stimulation (e.g. through chemokines), also mechanical stimulation is important for integrin activation³⁰. For example, SDF-1 α is a highly potent VLA-4 stimulatory chemokine³¹⁻³³, but fails to activate VLA-4 in the absence of external forces³⁴. The cytoskeletal attachment of integrins is known to determine their adhesiveness^{18, 35, 36}. Hence, the attachment state may play an important role in mechanically induced integrin activation.

Investigation of integrin-mediated cell signaling aimed at two objectives: First, the chemical and mechanical stimuli mandatory for firm T lymphocyte adhesion in the absence of external forces were investigated by studying the effect of SDF-1 α on the mechanics of the integrin $\alpha_4\beta_1$ environment (see section 4). Second, transient conformational states of the $\alpha_{IIb}\beta_3$ TMD were characterized by force spectroscopy measurements with integrin mutants functionally representing these states (see section 5).

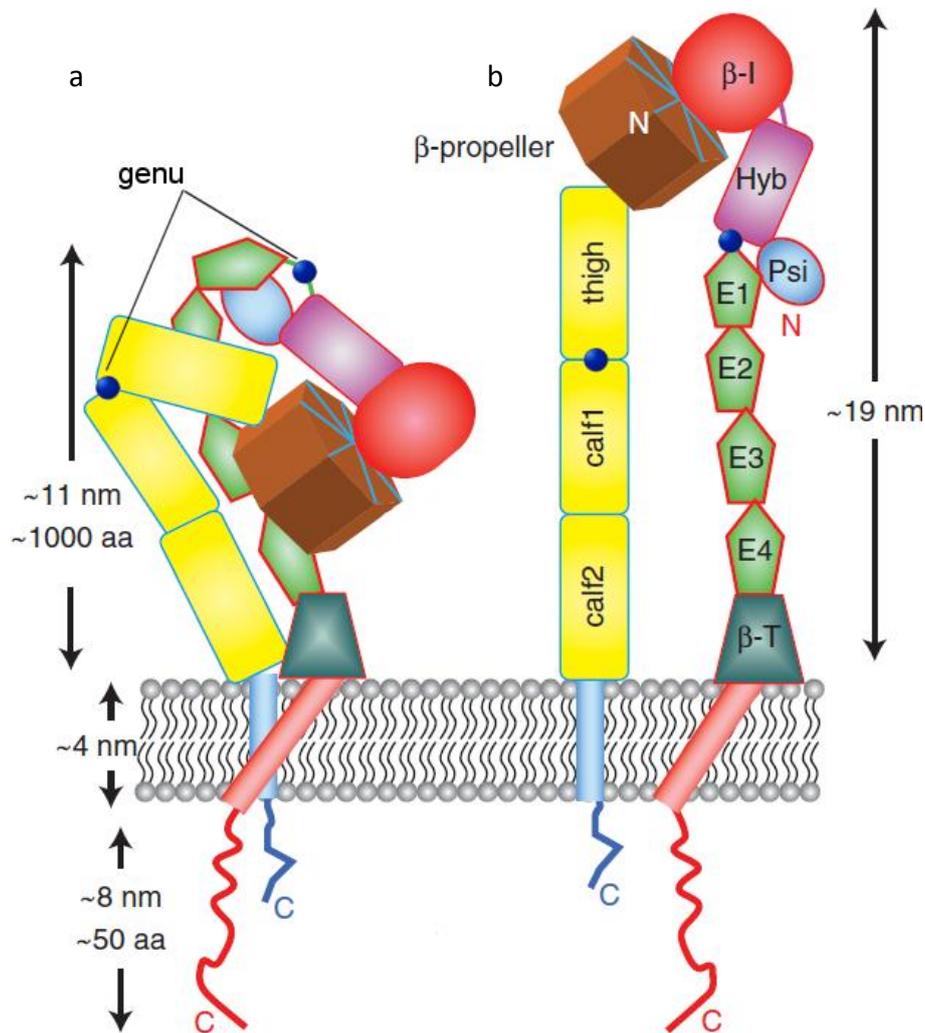


Fig. 3: Schematic structure of integrins like $\alpha_4\beta_1$ or $\alpha_v\beta_3$ lacking the α -I domain (modified from Campbell et Humphries³⁷). The α and β subunits are outlined in blue and red, respectively. a: In the crystal structure both dimers are bent over at the "genu" towards the C termini of the legs (marked with "C"). This is also supposed to be the inactive conformation of native integrins, which are embedded into the cell membrane with their legs. b: Model of the extended conformational state, which is thought to be a prerequisite for integrin-mediated cellular adhesion. The β -propeller presumably contains ligand and ion binding sites³⁸.

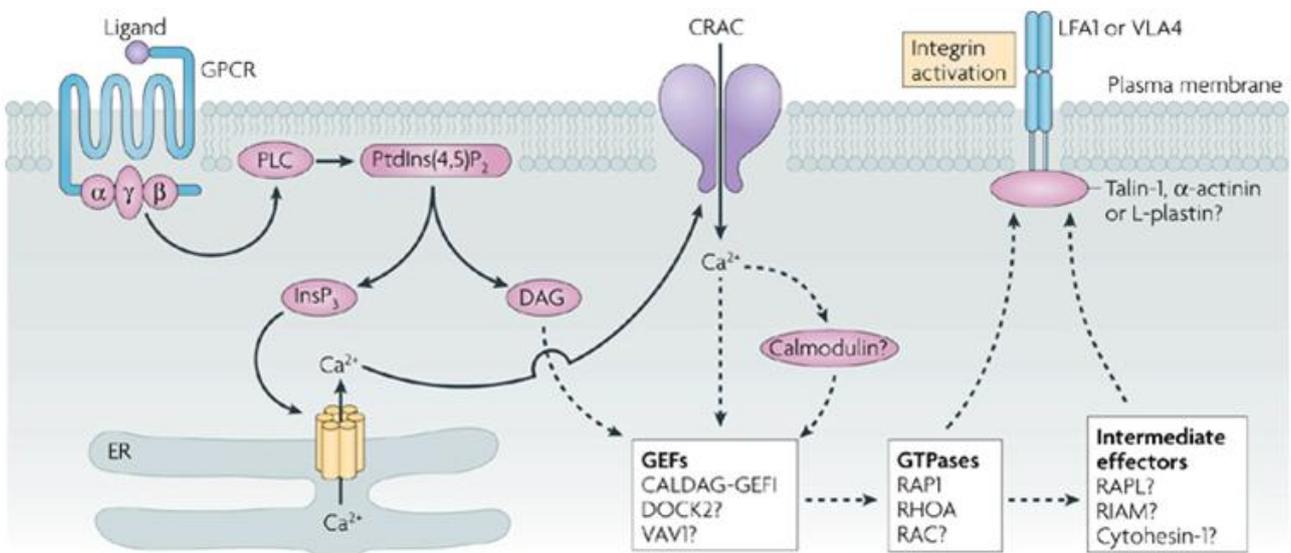


Fig. 4: Putative inside-out signaling cascade leading to integrin activation (modified from Ley et al.²⁷). The cascade is initiated by extracellular binding of a stimulatory ligand to the G-protein-coupled receptor (GPCR). Some steps of the pathway are still to be confirmed, as indicated by the dashed lines. Ultimately, the integrins are activated by adaptor proteins binding to their cytoplasmic tails.

2. Noise reduction

Experimental sciences and technologies concerning the measurement of weak signals share a fundamental problem: noise. Such signals can be e.g. physical observables, transmitted messages or any stored information. As it is inherently impossible to conduct noise-free measurements, the signal-to-noise ratio (SNR) is in many cases too low for direct extraction of the actual information and can neither be increased by improving the measuring principle nor by averaging over repeated runs (e.g. if the observation is very time-consuming or based on non-reproducible random effects). Yet, a post-processing technique may still reveal the concealed information. This is true for fields of applications as diverse as force spectroscopy of single-molecule interactions³⁹, fiber-optic communication systems⁴⁰, or the discovery of extrasolar planets⁴¹.

A measured signal can be represented by the vector y and is assumed to be a linear superposition

$$y_i = (s + n)_i, \quad i = 1, \dots, N \quad (1)$$

of the noise-free source s and a random noise n sampled at discrete intervals i . If the signal is smooth enough, simple methods, such as a moving average⁴², the Savitzky-Golay filter⁴³, or the convolution of y with a specially-designed low-pass filter kernel⁴⁴ are feasible to reduce the noise. However, if the signal is discontinuous (as for example the unfolding force of single molecules in two-state systems⁴⁵), other noise reduction methods are inevitable to preserve sudden transitions like steps or spikes. Fourier-based time-invariant filters fail to separate the high frequency components corresponding to these features from broadband noise. By contrast, a discrete wavelet transform (DWT) multilevel decomposition partially avoids this problem. It is frequently implemented in commercially available data analysis tools. The technique is outperformed by denoising based on stationary wavelet transform (SWT), which has been shown to provide a particularly good estimate \tilde{s} for virtually any kind of signal s ⁴⁶. It translates y from the time into a wavelet domain, where it is represented at different frequency scales, but still as a function of time (fig. 5). Coefficients with low amplitudes that do not impinge much on the signal are filtered out. Because the reverse transforms are slightly different, they can be averaged to reduce the noise. While still incorporating this concept, “recursive noise reduction” (ReNoiR) is based on two novel ideas (fig. 6): First, the wavelet transform is applied recursively to the original signal and to certain time-scale domains, resulting in a highly redundant expansion of y . As a consequence, an increased amount of reverse transforms can be averaged, so that the noise is reduced more

effectively. Second, small wavelet coefficients below a threshold are substituted by a de-noised version resulting from the next recursion level (see section 2.1.2). In contrast, conventional hard or soft thresholding⁴⁷ implies a higher loss of information, as these values are simply set to zero. ReNoiR differs from other techniques in superior signal recovery performance and preservation of discontinuities even at high noise levels. In the following, this is demonstrated on a variety of simulated and measured data.

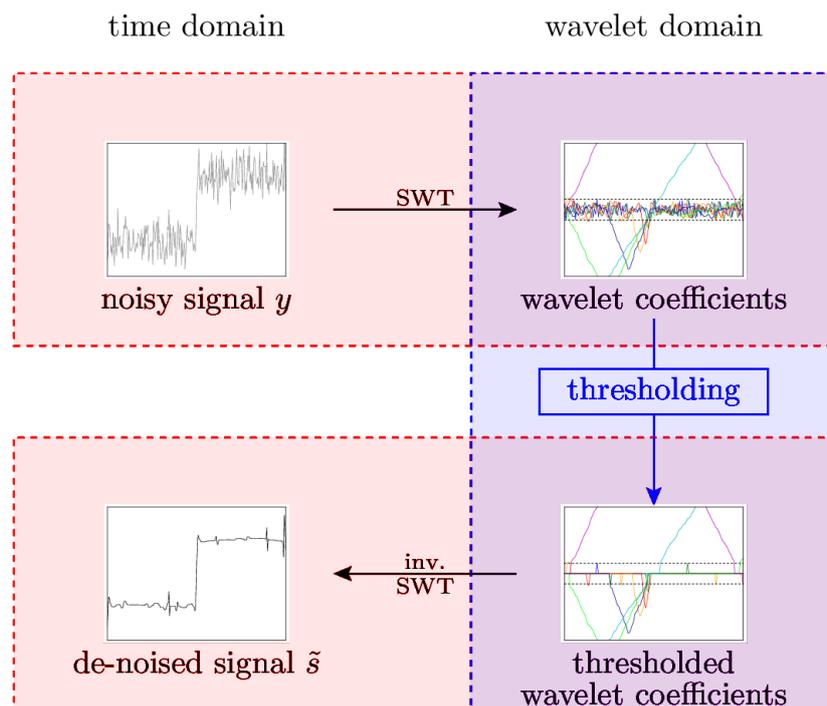


Fig. 5: Basic principle of SWT-based de-noising: A signal y is translated from the time into a wavelet domain by SWT (black arrows). After setting coefficients below a threshold (dashed lines) to zero, the reverse transforms are averaged to reconstruct the de-noised signal \tilde{s} .

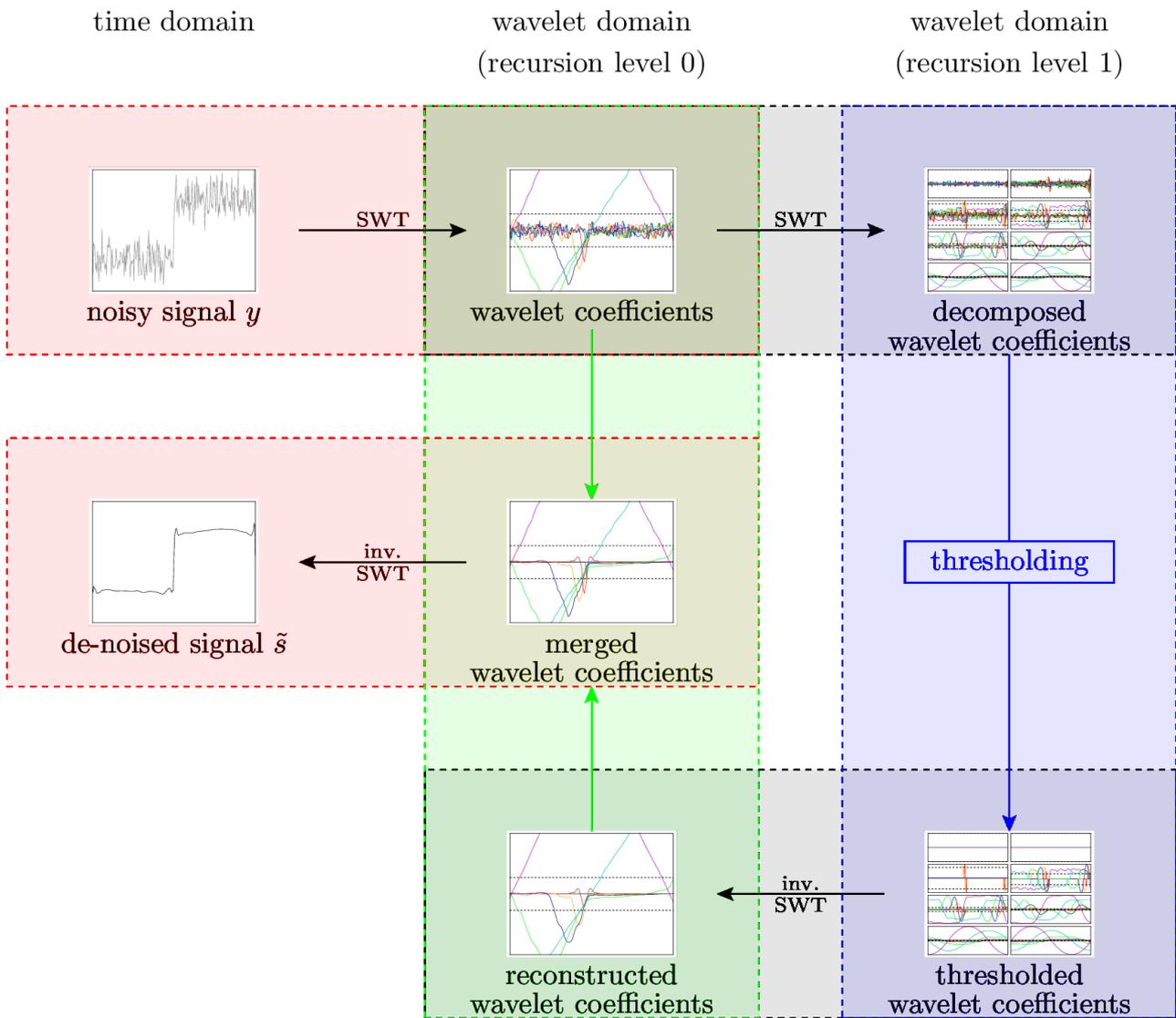


Fig. 6: Schematic representation of the ReNoiR algorithm: The wavelet coefficients of y are recursively decomposed by SWT (here only shown for one recursion) and reconstructed after conventional thresholding (blue). Thereby, an increased amount of reverse transforms can be averaged, so that the noise is reduced more effectively. Small wavelet coefficients of recursion level 0 are substituted by their reconstructed counterparts (green), because they are most affected by noise. Finally, \tilde{s} is recomposed by inverse SWT.

2.1. Materials and methods

2.1.1. Stationary wavelet transform

The stationary wavelet transform (SWT) is also known as "algorithme à trous"^{6, 7}. Briefly, the estimated source signal \tilde{s} is calculated by decomposing the noisy signal $y \equiv a_0^0$ of length N into scaling and wavelet coefficients a_1^0 and d_1^0 of the same lengths. These vectors contain the large-scale (approximation) and small-scale (detail) components of y , respectively. The decomposition is iteratively performed by convolution (denoted by $*$) of a_l^0 with the impulse responses g_l^{lo} and g_l^{hi} of a low- and high-pass filter until the maximum level of decomposition ($L \leq \log_2 N$) is reached ($l = 0, \dots, L - 1$):

$$a_{l+1}^0 = a_l^0 * g_l^{lo} \quad (2)$$

$$d_{l+1}^0 = a_l^0 * g_l^{hi} \quad (3)$$

Unless stated otherwise, $L = 5$ was chosen in this work, as in average best results were obtained with this setting. The decomposition filters g_0^{lo} and g_0^{hi} correspond to the Haar wavelet⁴⁸. They are up-sampled by a factor of 2 at the end of every iteration. After shrinking all of the L detail coefficients d_l^0 below a limit T_0 according to the rule $\hat{d}_l^0 = \text{sgn}(d_l^0) \cdot \max(0, |d_l^0| - T_0)$ (this is known as soft-thresholding⁴⁷), the inverse algorithm is used to recombine the estimate \tilde{s} . The amount of suppressed details is determined by T_0 .

SWT creates a $(L + 1)$ -fold redundant set of wavelet coefficients that correspond to the N possible translations of y : For each shift, y can be reconstructed from d_l^0 and a_L^0 . Thresholding of d_l^0 results in slightly different inverse transforms, which are averaged to reduce the noise. Thus, redundancy is the key concept of SWT. A more detailed description of the algorithm can be found in Fowler et al.⁴⁹.

2.1.2. The ReNoiR algorithm

ReNoiR is based on two ideas: First, redundancy of the representation of the noisy signal y in the wavelet domain is highly increased by recursively decomposing the detail coefficients d_l^0 of a SWT of y into approximation and detail coefficients $a_{lmn\dots}^r$ and $d_{lmn\dots}^r$, respectively (fig. 7). Second, $d_{lmn\dots}^r$ is merged with the filtered version $\hat{d}_{lmn\dots}^r$ resulting from the next recursion. Conventional hard-thresholding⁴⁷ of $d_{lmn\dots}^R$ is only performed in the last recursion level R . The merged small-

scale components $\tilde{d}_{lmn\dots}^r$ are recursively recomposed with the large-scale components $a_{lmn\dots}^r$ to $\hat{d}_{lmn\dots}^{r-1}$ ($r > 0$). Eventually, \tilde{d}_l^0 and a_l^0 are used to calculate an estimate \tilde{s} for the noise-free signal by inverse SWT.

This concept is implemented as follows: Initially, the vector $y \equiv a_0^0$ of size N is iteratively decomposed into $L = \log_2 N$ approximation and detail coefficients a_l^0 and d_l^0 ($l = 1, \dots, L$) using eq. (2) and (3) (fig. 8). After every iteration step, the low- and high-pass quadrature mirror filters

$$g_0^{\text{lo}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad (4)$$

$$g_0^{\text{hi}} = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix} \quad (5)$$

are up-sampled by a factor of 2. They are the finite impulse responses corresponding to the orthogonal Haar wavelet⁴⁸, which is particularly suited for the analysis of sudden transitions. It is defined by the wavelet function^{50, 51}

$$\psi(t) = \begin{cases} 1, & t \in [0, 1/2[\\ -1, & t \in [1/2, 1[\\ 0, & t \notin [0, 1[\end{cases} \quad (6)$$

and the scaling function

$$\phi(t) = \begin{cases} 1, & t \in [0, 1[\\ 0, & t \notin [0, 1[\end{cases} \quad (7)$$

To eliminate any dependency of the noise reduction efficiency on the chosen wavelet and thereby allow for an objective comparison between ReNoiR and the conventional SWT-based method, identical filters are used in both cases.

Next, the convolutions

$$a_{l,\dots,z+1}^{r+1}[i] = (a_{l,\dots,z}^r * g_z^{\text{lo}})[i] = \sum_{j=0}^{M-1} d_{l,\dots,z}^r \left[-\frac{M}{2} + i + j \right] g_z^{\text{lo}}[M - j], \quad i = 1, \dots, N \quad (8)$$

$$d_{l,\dots,z+1}^{r+1}[i] = (a_{l,\dots,z}^r * g_z^{\text{hi}})[i] = \sum_{j=0}^{M-1} d_{l,\dots,z}^r \left[-\frac{M}{2} + i + j \right] g_z^{\text{hi}}[M - j], \quad i = 1, \dots, N \quad (9)$$

are iteratively applied to the wavelet coefficients $d_l^0 \equiv d_{l,m=0,n=0,\dots,z=0}^0$, starting with $z = m$, $m = 0$ and $r = 0$ (fig. 8). The $R + 1$ indices l, m, n, \dots run from 0 to $L - 1$ and designate the iteration steps of each recursion level (z is only a substitute for the last index). M is the length of the filter vectors (here $M = 2$). $d_{l,\dots,z}^r$ must be periodically extended:

$$d_{l,\dots,z}^r[N + i] = d_{l,\dots,z}^r[i], \quad i = -\frac{M}{2} + 1, \dots, \frac{M}{2} - 1 \quad (10)$$

Again, the decomposition filters g_z^{lo} and g_z^{hi} are up-sampled by a factor of 2 after every iteration. This scheme is repeated recursively with $d_{lmn\dots}^{r+1}$ until the expansion is stopped at a specified level R . Totally, L^{R+1} detail and $\sum_{r=0}^R L^r = (L^{R+1} - 1)/(L - 1)$ approximation coefficients of size N are generated, so that y is represented by a $(L^{R+2} - 1)/(L - 1)$ -fold redundant set of wavelet coefficients at the end of the recursive expansion.

The small-scale components $d_{lmn\dots}^R$, whose amplitudes fall below a limit $T_R(l, m, n, \dots)$, are set to zero (this is known as hard-thresholding⁴⁷):

$$\hat{d}_{lmn\dots}^R[i] = \begin{cases} d_{lmn\dots}^R[i], & d_{lmn\dots}^R[i] \geq T_R(l, m, n, \dots) \\ 0, & d_{lmn\dots}^R[i] < T_R(l, m, n, \dots) \end{cases} \quad i = 1, \dots, N \quad (11)$$

Having reached this point, the recursive expansion must be completely reverted to reconstruct the signal. Therefore, the wavelet coefficients $\hat{d}_{lmn\dots}^R$ are recomposed with the scaling coefficients $a_{lmn\dots}^R$ by the inverse algorithm, i.e. both components are deconvolved by means of the reconstruction filters \hat{g}_z^{lo} and \hat{g}_z^{hi} :

$$\hat{d}_{l,\dots,z-1}^r = \frac{1}{2} (a_{l,\dots,z}^{r+1} * \hat{g}_z^{\text{lo}} + \hat{d}_{l,\dots,z}^{r+1} * \hat{g}_z^{\text{hi}}) \quad (12)$$

Eq. (12) is iteratively evaluated, starting with $l, m, n, \dots = L$ and $r = R - 1$. As g_z^{lo} and g_z^{hi} are conjugate quadrature mirror filters, \hat{g}_z^{lo} and \hat{g}_z^{hi} are simply the time reverses of the up-sampled decomposition filters g_{L-1}^{lo} and g_{L-1}^{hi} . They are down-sampled by a factor of 2 after every iteration.

For $0 \leq r < R$, the original detail coefficients $d_{lmn\dots}^r$ are then replaced with the filtered $\hat{d}_{lmn\dots}^r$ if their amplitudes are smaller than a threshold $T_r(l, m, n, \dots)$:

$$\tilde{d}_{lmn\dots}^r[i] = \begin{cases} d_{lmn\dots}^r[i], & d_{lmn\dots}^r[i] \geq T_r(l, m, n, \dots) \\ \hat{d}_{lmn\dots}^r[i], & d_{lmn\dots}^r[i] < T_r(l, m, n, \dots) \end{cases} \quad i = 1, \dots, N \quad (13)$$

To calculate the wavelet coefficients of the next recursion level $r - 1$, $a_{lmn\dots}^r$ and the merged small-scale components $\tilde{d}_{lmn\dots}^r$ are substituted into eq. (12). The whole procedure (reconstruction, down-sampling, and merging) is repeated until $r = 0$ ($\hat{d}_{l,m=0,n=0,\dots,z=0}^0 \equiv \hat{d}_l^0$). Ultimately, the estimate \tilde{s} is reconstructed from a_l^0 and \tilde{d}_l^0 by inverse SWT. In this work, the number of recursions was limited to $R = 1$ and the thresholds T_r were either optimized automatically (see section 2.1.3) or set to ideal values in the sense of minimum deviation from the clean signal (see section 2.2.3).

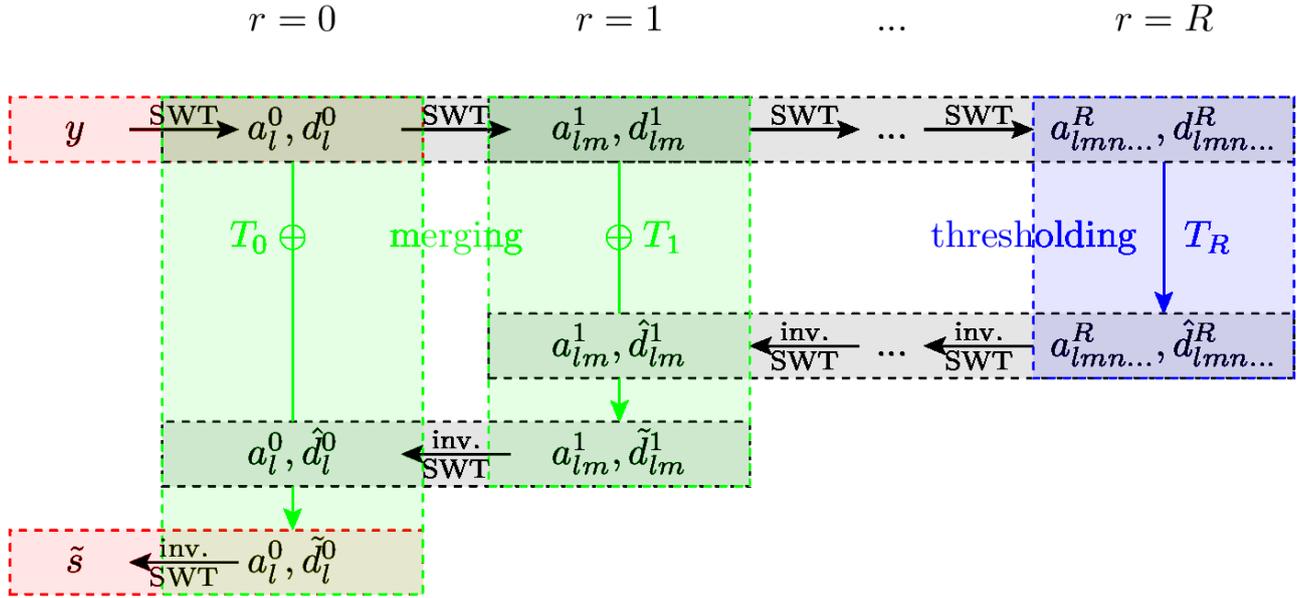


Fig. 7: Principle of the ReNoiR algorithm. The detail coefficients d_l^0 of the stationary wavelet transform of a noisy signal y are recursively decomposed into approximation and detail components $a_{lmn\dots}^r$ and $d_{lmn\dots}^r$, respectively (black arrows). For each $r < R$, the detail coefficients $d_{lmn\dots}^r$ are merged with the filtered $\hat{d}_{lmn\dots}^r$ obtained in the next recursion (symbolized by the operators \oplus). At the last recursion level $r = R$, the $d_{lmn\dots}^R$ are thresholded instead (blue arrow). In both cases, the result is recomposed with $a_{lmn\dots}^r$ to $\hat{d}_{lmn\dots}^{r-1}$ ($r > 0$). Eventually, \tilde{d}_l^0 and a_l^0 are used to calculate an estimate \tilde{s} for the noise-free signal by inverse SWT.

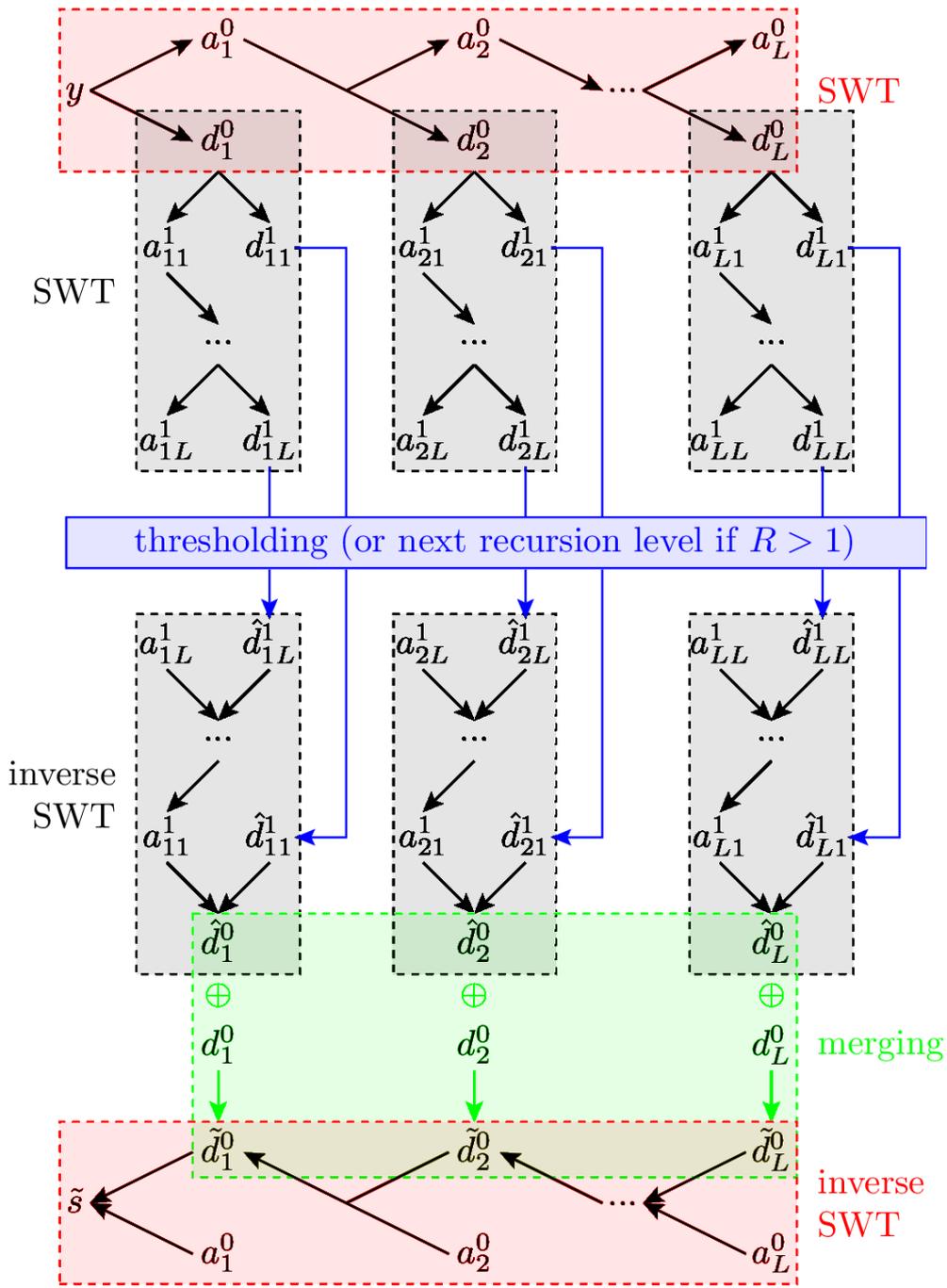


Fig. 8: Expanded view of the concept of ReNoiR for $R = 1$ showing individual iteration steps.

2.1.3. Automatic parameter optimization

The level-dependent thresholds T_0 for SWT were calculated by the VisuShrink⁴⁷, SUREShrink⁵², and minimax^{47, 53} algorithms. They require the knowledge of the standard deviation of the noise, which was estimated separately for each decomposition level l based on the median absolute deviation (MAD) of the wavelet coefficients⁴⁷:

$$\sigma_{\text{noise}}(l) = \text{median}(|d_l^0|)/\gamma \quad (14)$$

For white noise, the result is independent of l , and $\sigma_{\text{noise}}(0)$ was used for all levels. γ was calibrated by means of artificially generated noise with different frequency spectra (table 1).

For ReNoiR, a new method to optimize T_0 and T_1 automatically had to be developed: In a first step, for a known noise spectrum, the standard deviations σ of the wavelet coefficients d_l^r corresponding to the noise are determined by evaluation of artificial noise signals ($r \in [0,1]$; $0 \leq l < \log_2 N$). The level-dependent thresholds

$$T_r(l, m) = \sigma(d_l^r)[c_1(N) + c_2(\text{SNR})] \quad (15)$$

are chosen proportional to these “fingerprints”. c_1 and c_2 are correction terms compensating the effects of variable signal lengths and signal-to-noise ratios:

$$c_1(N) = aN^b \quad (16)$$

$$c_2(\text{SNR}) = c_2(\sigma_{\text{source}}/\sigma_{\text{noise}}) = A \exp[-c(\sigma_{\text{noise}}/\sigma_{\text{source}} - d)^2] \quad (17)$$

These relations and parameters were determined empirically (table 1) by analyzing a priori known ideal values in the sense of minimal root mean square (RMS) error for a large number of synthetic test signals with varied N and SNRs (see also section 2.2.3). σ_{noise} was estimated according to eq. (14) and σ_{source} by

$$\sigma_{\text{source}} = \sqrt{\max[\sigma^2(d_l^0)] - \sigma_{\text{noise}}^2} \quad (18)$$

noise type	γ	r	a	b	A	c	d
white	0.6745	0	2.43	0.068	-0.91	5.09	0.374
		1	3.03	0.052	-1.21	3.49	0.445
pink	0.3020	0	1.46	0.129	-2.10	150.00	0.050
		1	1.62	0.122	-2.10	150.00	0.050
Brownian	0.0227	0	0.05	0.553	-2.44	18.22	0.017
		1	0.04	0.577	-2.10	100.56	0.118
AFM	0.4320	0	2.65	0.060	-1.07	47.20	0.179
		1	2.76	0.061	-1.12	16.64	0.174
optical tweezers	0.4610	0	2.08	0.088	-1.26	42.17	0.139
		1	2.74	0.062	-1.40	26.92	0.140

Table 1: Calibration parameters for different types of noise

2.1.4. Evaluation of the signal recovery performance

Each noise reduction method was analyzed by means of synthetic test signals y consisting of $N = 4096$ samples superimposed with additive noise of different amplitudes and frequency spectra. White, pink, Brownian, and instrumental noise was generated by inverse Fourier transform of a given spectral distribution $H(\omega)$. For white noise $H(\omega) = 1$, for pink noise $H(\omega) = \omega^{-1/2}$ and for Brownian noise $H(\omega) = \omega^{-1}$. Noise spectra of an atomic force microscope and of optical tweezers were obtained by recording time series free of sample-specific effects. The data were de-trended by subtraction of a linear baseline and scaled to standard deviation 1. $H(\omega)$ was calculated by fast Fourier transform using a block size of 2^8 and a Hann window to reduce spectral leakage (fig. 9).

The SNRs of y and \tilde{s}

$$\text{input SNR} = \text{SNR}(y) := \frac{\max(s) - \min(s)}{\sigma(y - s)} \quad (19)$$

$$\text{output SNR} = \text{SNR}(\tilde{s}) := \frac{\max(s) - \min(s)}{\sigma(\tilde{s} - s)} \quad (20)$$

are here defined as the ratio between the amplitude range of the source and the standard deviation σ of the noise to serve as a measure of their fidelities ($\text{SNR}(s) = \infty$). The output SNR is commonly used to evaluate the signal recovery performance of noise reduction algorithms⁴⁶ and can be related to the input SNR to express the relative gain.

Four standard test signals⁴⁷ with 101 logarithmically distributed input SNRs ranging from 10^{-2} to 10^2 and 2840 variations of these basic forms were evaluated (see section 2.1.5). Python was used for generation of the test signals and data analysis (Python Software Foundation, Wolfeboro Falls, NH).

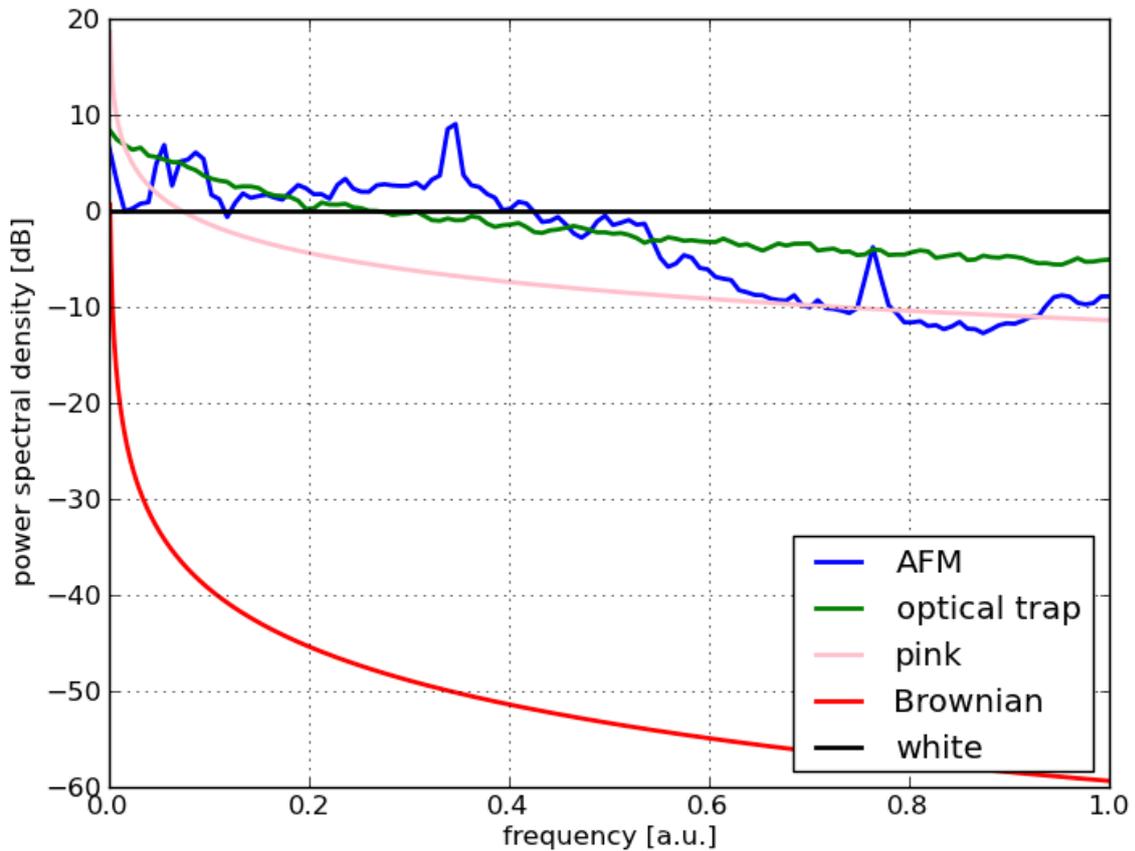


Fig. 9: Power spectral densities $|H(\omega)|^2$ of different types of noise

2.1.5. Estimate of the noise reduction performance for fundamental signal forms

To provide a quantitative prediction of the efficiencies attainable by SWT and ReNoiR, 2840 variations of four fundamental signal forms⁴⁷ frequently encountered in experimental data were analyzed (steps, peaks, damped oscillations, and a combination of sinusoidal oscillations and steps – see insets of fig. 11): generalized “Blocks” with variable number of steps and step widths, “Bumps” with variable number of peaks and peak widths at random positions, “Dopplers” with

variable number of oscillations and phase shifts, as well as “Heavisines” with variable number of oscillations, step heights and random step positions/directions (see table 2). These generalized test signals are defined by the following formulas:

- Blocks (l steps of width w):

$$s_i = \sum_{j=0}^{l-1} (-1)^j \Theta \left(i - \frac{N - lw}{2} - jw \right) \quad (21)$$

- Bumps (l peaks of width w at l random positions $p_j \in [0, N]$):

$$s_i = \sum_{j=0}^{l-1} \left(1 + \left| \frac{i - p_j}{w} \right| \right)^{-4} \quad (22)$$

- Dopplers (mean oscillation time τ , phase shift φ):

$$s_i = \sqrt{\frac{i}{N-1-i}} \sin \left(2\pi \frac{\tau + 1.000001}{\frac{i}{N-1} + \tau + 0.000001} + \varphi \right) \quad (23)$$

- Heavisines (l full oscillations with one step of height $d_j h$ and random $d_j \in \{-1, +1\}$ each at random positions $p_j \in [jN/l, (j+1)N/l]$):

$$s_i = \frac{1}{2} \sin \left(2\pi l \frac{i}{N-1} \right) - \sum_{j=0}^{l-1} d_j h [2\Theta(i - p_j) - 1] \quad (24)$$

Θ denotes the Heaviside step function.

The SNR was varied using a variable magnitude of additive white Gaussian noise corresponding to input SNRs of 0.1, 1.0, 5.0, 25.0, and 100.0. Some randomly selected examples are shown in fig. 10. Filter parameters were optimized automatically (see section 2.1.3).

The performance of the compared algorithms was evaluated by systematically varying all signal parameters except one at a time and averaging the relative gains $\text{SNR}(\hat{s})/\text{SNR}(y)$ obtained for the signals generated with the constant parameter. For example, the results for the “Bumps” signal

were averaged over different peak widths and random peak positions for each number of peaks. Subsequently, the gains were averaged over the numbers of peaks and peak positions for each peak width. That way it is possible to estimate the noise reduction efficiency for arbitrary variations of basic signal forms as a function of one characteristic parameter (e.g. the peak width).

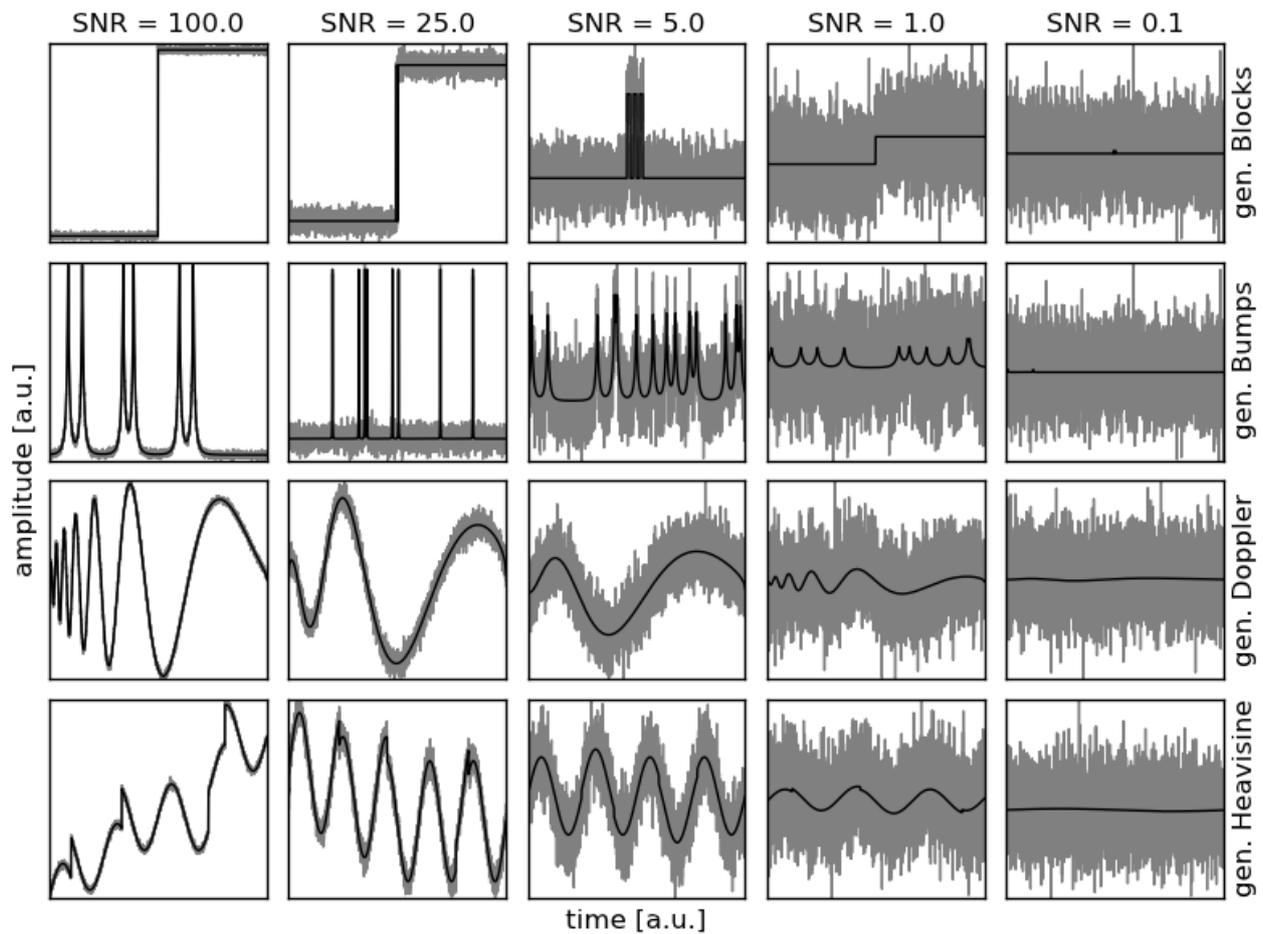


Fig. 10: Randomly selected examples out of 2840 test signals (gray) with systematically varied characteristics (such as the number of peaks and their widths) and SNRs, which were analyzed using different noise reduction algorithms. The noise-free data (black) were combined with additive white Gaussian noise.

gen. Blocks	gen. Bumps	gen. Doppler	gen. Heavisine
number of steps (15 values from 1 to 20)	number of peaks (11 values from 1 to 20)	mean oscillation time (40 values from 0 to 1)	number of full oscillations (1, 2, 3, 4, 5)
step width (20 values from 2 to 1000 data points)	peak width (13 values from 1 to 400 data points)	phase shift (10 values from 0 to π)	step height (15 values from 0.0 to 0.5)
	peak positions (10 random curves)		random step positions and directions (10 random curves)

Table 2: Varied parameters characterizing the generalized test signals

2.1.6. Analysis of the tether onset force

A Nanowizard II atomic force microscope (JPK, Berlin, Germany) was deployed to measure force-distance curves of β_1 integrin-deficient Jurkat A1 lymphocytes with re-substituted β_1 integrin⁵⁴ interacting with the VLA-4 ligand VCAM-1 as described by Schmitz et al.⁵⁵. A VCAM-1 concentration of 2 $\mu\text{g}/\text{ml}$ and a constant approach/retract velocity of 3.4 $\mu\text{m}/\text{s}$ were used. ReNoiR was applied to reveal hidden steps in the acquired data. Curves showing at least two steps were extracted. The first one was considered a force barrier for tether formation if meeting the following criteria: It appears within a retraction distance of 500 nm and is clearly separated by not less than 50 data points from its successive step that is at least 10 pN higher than the first one. Step heights were determined as indicated by the red lines in fig. 16b.

2.1.7. Analysis of the SERS spectra

SERS data measured with a Raman spectroscope⁵⁶ was processed by ReNoiR to reduce the noise using the filter settings $R = 1$, $T_0 = \infty$ and $T_1 = 10$. Local maxima were detected as peaks if showing a strictly monotonically increasing flank on the left and a strictly monotonically decreasing flank on the right consisting of 4 data points each.

2.1.8. Analysis of the optical tweezers measurements

Histograms over force vs. time records of leucine zipper constructs measured with optical tweezers⁵⁷ were calculated after reducing the noise using the ReNoiR algorithm (fig. 17a). Three

distributions corresponding to the unfolded and two intermediate states could be clearly resolved. The numerical derivative of the de-noised signal was thresholded to locate possible folding/unfolding transitions (fig. 17b) and the previously determined force levels were fitted to the data between these transitions (fig. 17c).

2.2. Results

2.2.1. Synthetic data

The ability of ReNoiR to recover a clean source s hidden in a noisy signal y with SWT-based de-noising was compared using automatically optimized thresholds. For SWT, multiple techniques exist to choose these parameters without any prior knowledge about the clean signal. The VisuShrink⁴⁷, SUREShrink⁵², and minimax^{47, 53} algorithms were compared. For ReNoiR, a new method had to be developed. It aims at thresholding wavelet coefficients adaptively to the mean amplitudes expected for the noise of a known spectrum at each recursion and decomposition level (see section 2.1.3).

Frequently used synthetic test signals superimposed with noise of different amplitudes and spectral distributions were analyzed, as they reflect common properties of real experimental data⁴⁷ (insets of fig. 11). The relative gain of the mean SNR of 100 estimates \tilde{s} $\langle \text{SNR}(\tilde{s}) \rangle / \text{SNR}(y)$ was evaluated for input SNRs of the noisy data y in the range from 10^{-2} to 10^2 . For all investigated test signals and most input SNRs, ReNoiR yields higher signal fidelities than the SWT-based techniques, both for additive white Gaussian noise (fig. 11, 12) and for different types of colored noise (fig. 13). With decreasing $\text{SNR}(y)$, the correlation between \tilde{s} and s is gradually lost. At very high noise levels, the best estimate for s ultimately becomes a straight line (i.e. $\tilde{s} = 0$). Consequently, the output SNRs of all noise reduction algorithms approach $\text{SNR}(\tilde{s} = 0)$ for input SNRs around or below 0.5 (dashed gray lines in fig. 11). Recovery performance for low SNRs above this limit is particularly relevant for practical applications (fig. 12). The relative gains averaged between 0.5 and 2.5, which is also the range shown in fig. 12, are largest for ReNoiR, except for the “Bumps” signal contaminated with white noise of intermediate amplitudes (table 3). The recursive approach generates less spike-shaped artifacts (known as pseudo-Gibbs phenomena) than SWT and reproduces sharp transitions more clearly (fig. 14). For too low input SNRs to yield any information correlated to the clean signal (such as in the beginning of the “Doppler” example, fig.

14c), ReNoiR tends to produce a noise-free constant signal with zero amplitude, whereas a considerable amount of noise still passes the SWT-based filters.

To provide a quantitative prediction of the noise reduction performance for the basic signal forms of the four standard test signals (peaks, steps, oscillations, and a combination of the latter two), their characteristics were systematically varied (such as the peak heights and widths) and the output SNR was evaluated as a function of these parameters (see section 2.1.5). The analysis revealed that ReNoiR in average shows significantly higher gains than SWT for the four basic signal forms and all noise levels, independently of the chosen parameter optimization algorithm (fig. 15 and table 4). Averaged over all test signals of the four base classes, an improvement between 17% and 857% was attained.

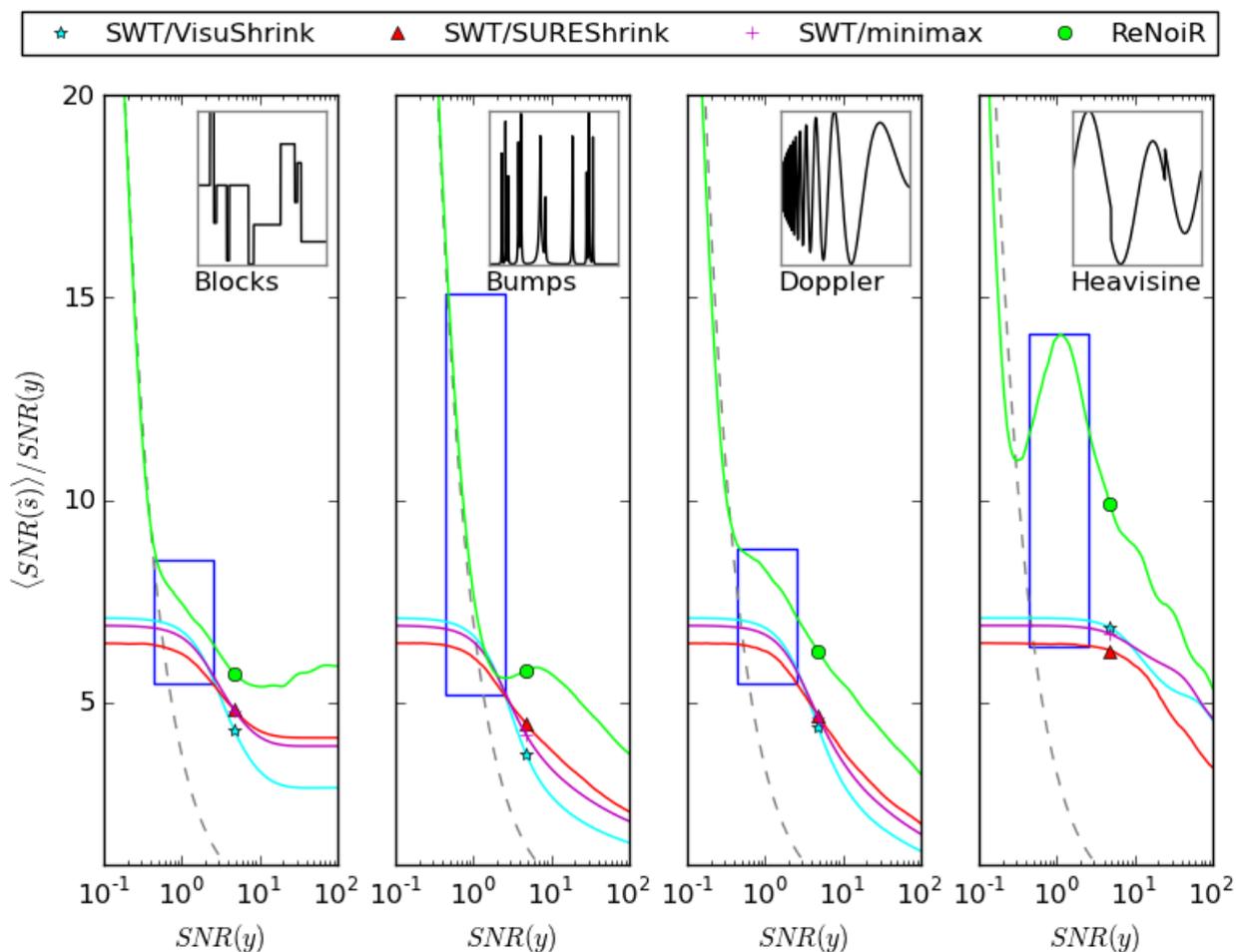


Fig. 11: Signal recovery performance of different filter methods. The relative gain of the mean SNR after noise reduction is plotted vs. the SNR of the noisy test signals y , which are composed of the sample curves s shown in the insets and additive white Gaussian noise of varied amplitudes. Each data point represents an average value obtained from 100 curves with distinct random noise. Dashed gray lines: output SNRs for $\tilde{s} = 0$. The markers are drawn for clarity only.

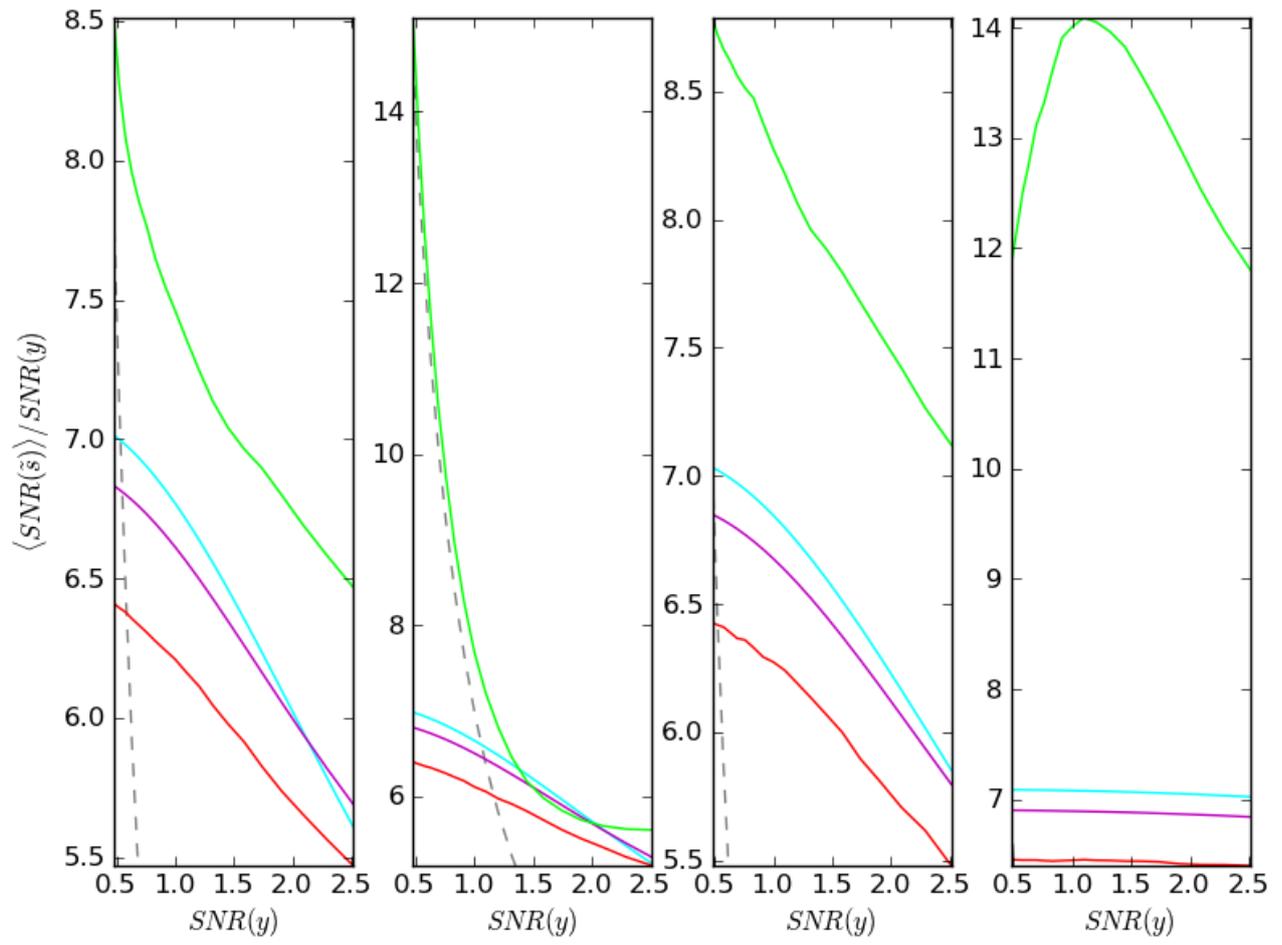


Fig. 12: Zoomed views of the areas marked by the blue rectangles in fig. 11.

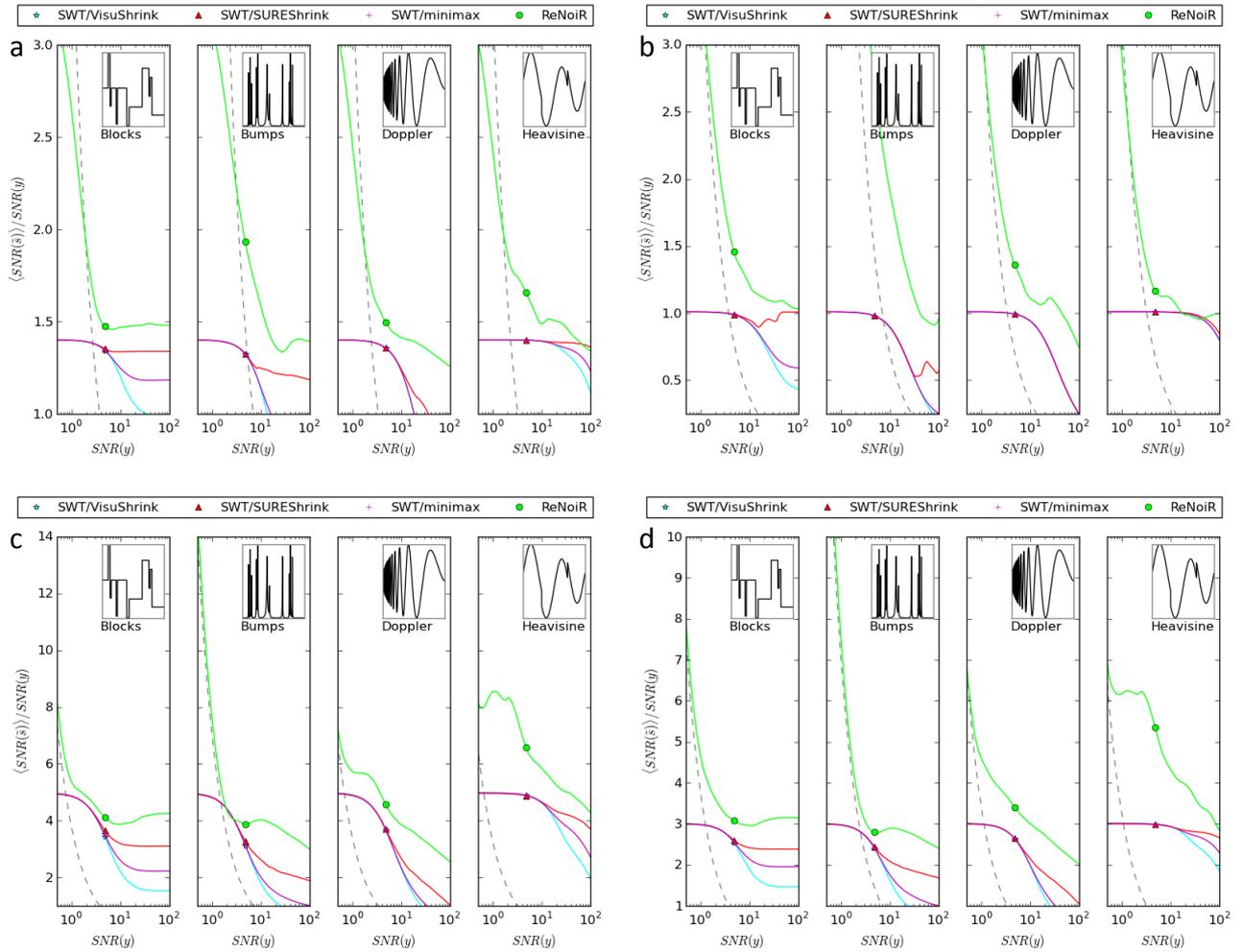


Fig. 13: Signal recovery performance of different filter methods with automatically optimized parameters applied to test signals composed of the same noise-free curves as in fig. 11 and additive correlated noise of varied amplitudes and frequency spectra. The relative gain of the mean SNR after noise reduction is plotted vs. the SNR of the noisy signals y contaminated with pink (a) and Brownian (b) noise, as well as instrumental noise of an atomic force microscope (c) and of optical tweezers (d). Each data point represents an average value obtained from 100 curves with distinct random noise. The markers are drawn for clarity only.

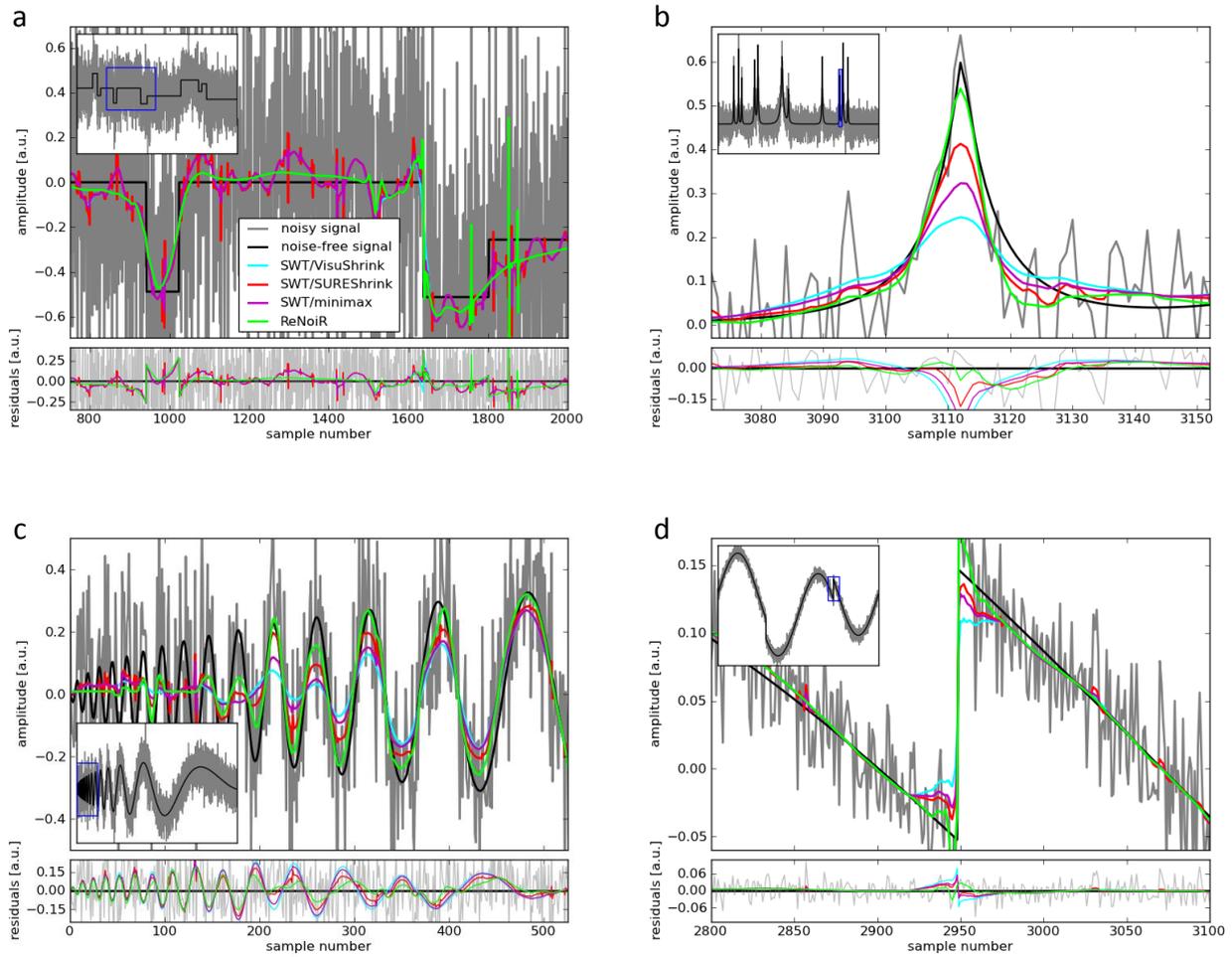


Fig. 14: Test signals (a: “Bumps”, b: “Blocks”, c: “Doppler”, d: “Heavisine”) contaminated with white Gaussian noise after filtering with different techniques (top) and residuals showing the deviation from the noise-free signal (bottom). The magnified sections are indicated by rectangles in the insets. The input and output SNRs are listed in table 5.

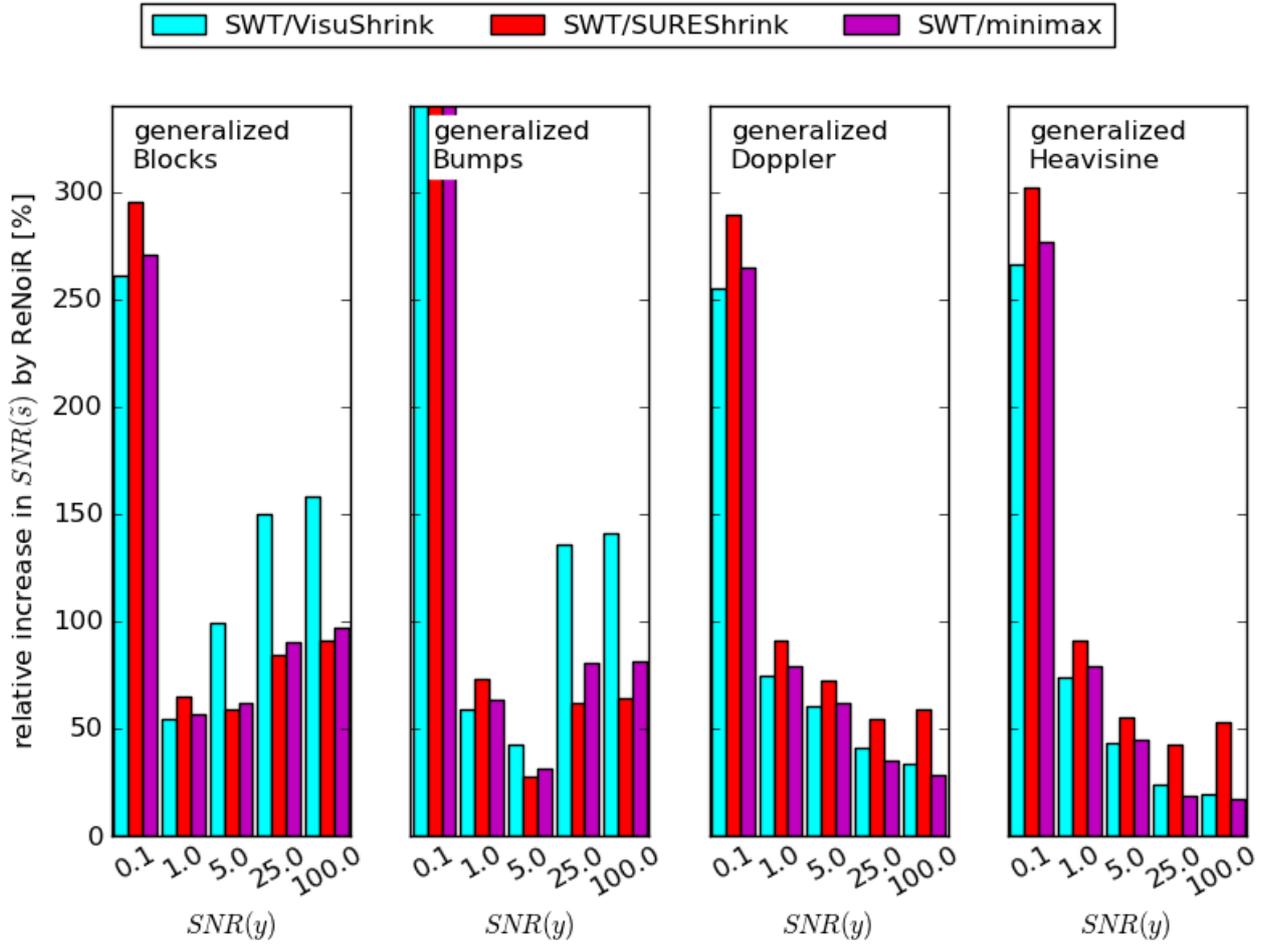


Fig. 15: Signal recovery performance of ReNoiR relative to SWT obtained with generalized test signals (see examples in fig. 10) and automatically optimized thresholds. The relative gain in output SNR $\langle \text{SNR}(\hat{s}_{\text{ReNoiR}}) / \text{SNR}(\hat{s}) \rangle - 1$ averaged over varied signal characteristics (table 2) is shown for the four fundamental signal forms and various input SNRs. ReNoiR always produced higher average output SNRs than the SWT-based methods, so that no negative gains occur.

de-noising method	threshold optimization method	$\langle\langle \text{SNR}(\hat{s}) \rangle\rangle / \text{SNR}(y)$			
		Blocks	Bumps	Doppler	Heavisine
SWT	VisuShrink	6.56	6.39	6.67	7.07
SWT	SUREShrink	6.07	5.95	6.13	6.44
SWT	minimax	6.44	6.28	6.52	6.88
ReNoiR ($R = 1$)	adaptive	7.38	8.39	8.10	13.13

Table 3: Relative gain of the signal-to-noise ratio (SNR) achieved with various noise reduction methods and automatically optimized thresholds applied to the test signals plotted in the insets of fig. 11. Each value represents the average gain over the range of input SNRs shown in fig. 12 (0.5 to 2.5).

compared threshold optimization method	SNR(y)	relative increase in SNR(\hat{s}) by ReNoiR [%]			
		generalized Blocks	generalized Bumps	generalized Doppler	generalized Heavisine
VisuShrink	0.1	261.1	772.9	255.0	266.5
	1.0	54.3	59.2	74.6	74.0
	5.0	98.9	42.4	60.4	43.5
	25.0	149.7	136.0	40.7	23.8
	100.0	158.2	140.7	33.3	19.5
SUREShrink	0.1	295.5	856.5	289.0	301.7
	1.0	64.6	73.0	91.2	90.7
	5.0	58.7	27.6	72.1	55.1
	25.0	84.6	61.6	54.1	42.8
	100.0	90.9	64.0	59.1	52.8
minimax	0.1	270.8	796.5	264.6	276.5
	1.0	56.4	63.0	79.3	78.7
	5.0	62.1	31.1	62.2	45.0
	25.0	90.1	80.3	35.0	18.3
	100.0	96.6	81.2	28.3	16.9

Table 4: Signal recovery performance of ReNoiR relative to SWT obtained with generalized test signals and automatically optimized thresholds. The table lists the values shown in fig. 15.

de-noising method	threshold optimization method	SNR			
		Blocks	Bumps	Doppler	Heavisine
none (noisy signal)		2.0	10.0	5.0	30.0
SWT	VisuShrink	12.0	25.4	20.8	156.9
SWT	SUREShrink	11.5	39.4	23.1	171.6
SWT	minimax	12.1	32.1	21.7	178.2
ReNoiR ($R = 1$)	adaptive	14.3	57.6	31.1	225.6

Table 5: Signal-to-noise ratio (SNR) achieved with various noise reduction methods and automatically optimized thresholds applied to the test signals plotted in fig. 14.

2.2.2. Experimental data

ReNoiR was further analyzed using experimental data to demonstrate its benefit for practical applications. To this end, the adhesion force of single tethers (membrane tubes) formed by lymphocytes when interacting with the integrin VLA-4 ligand VCAM-1 was measured. Force-distance curves were obtained by atomic force spectroscopy measurements with membrane tethers pulled from living human T lymphocytes to test the existence of a force barrier for tether formation (fig. 16a). The phenomenon has been described theoretically and was measured to be in the low pN range for membrane vesicles at the used retract velocities and contact areas⁵⁸. A tether onset force of hundreds of pN was found using optical tweezers on outer hair cells⁵⁹, but on lymphocytes the barrier could not be revealed directly in AFM spectra, because it is hidden by the noise. The property of ReNoiR to preserve sudden transitions made it possible to observe the tether onset force on lymphocytes for the first time (fig. 16b) and determine it to be (10.4 ± 0.5) pN.

By applying the same method to the detected position of a bead trapped by magnetic tweezers and bound to a DNA molecule⁶⁰, the distance between the microsphere and the lens focus could be resolved down to less than 1 nm (fig. 16c, d). Similarly, by de-noising force vs. time records of leucine zipper constructs measured with optical trap experiments⁵⁷, two intermediate folded states could be clearly distinguished and the transitions between them could be resolved (fig. 17). In contrast to hidden Markov models, no estimates are necessary, e.g. about transition probabilities. Hence, an unbiased data analysis is possible. The individual states are less marked in the histograms of the raw data and of the same force trace de-noised by SWT. ReNoiR increases

the SNR, so that the distributions are sharpened. Thereby, peaks initially blurred by noise become visible. This is particularly important for systems with unknown and hardly separated states.

ReNoiR can also be deployed for locating very low peaks, as for example in a surface-enhanced Raman spectroscopy (SERS) spectrum⁵⁶ (fig. 16e). Besides offering faster data acquisition, the technique allows for the recognition of scattering species at small concentrations that are initially below the detection limit.

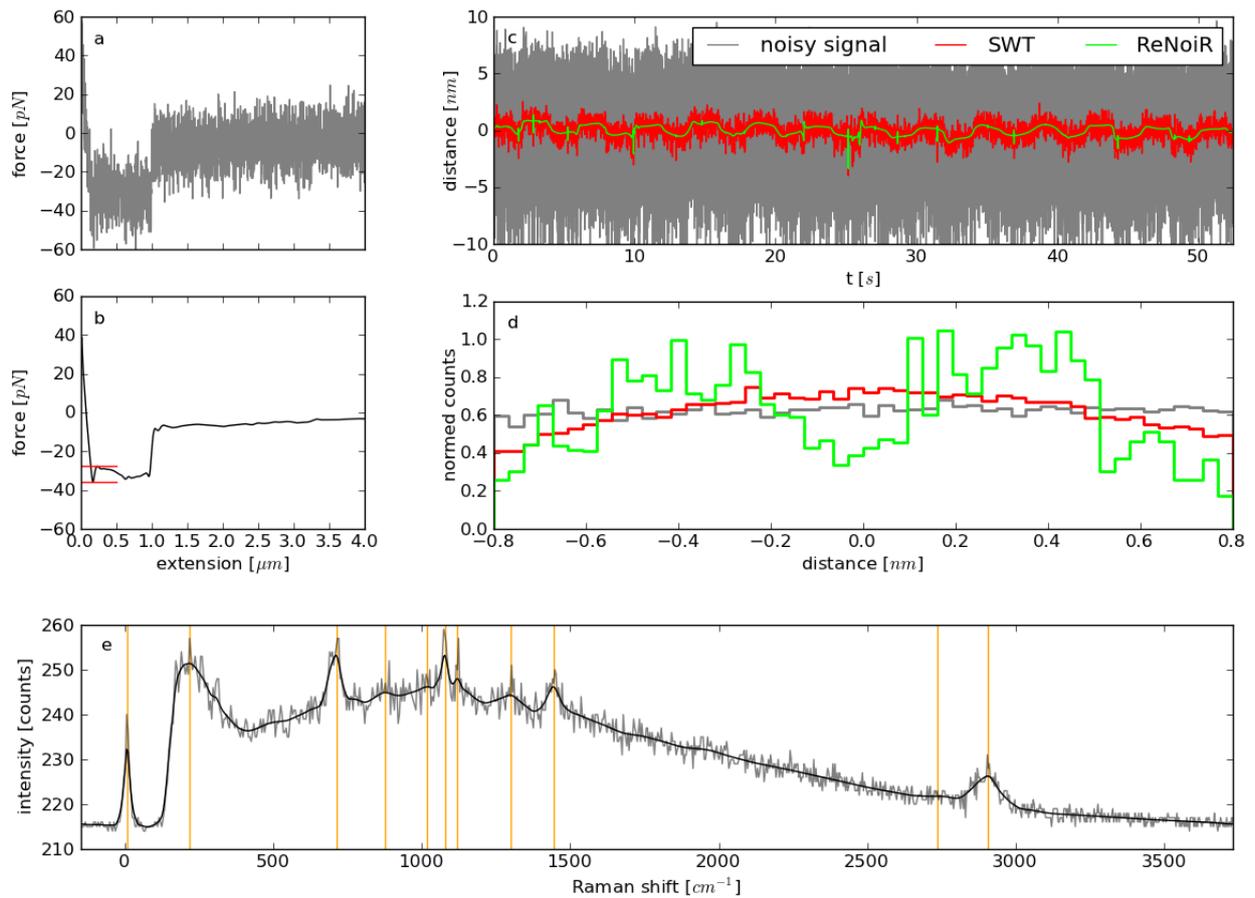


Fig. 16: Application of the ReNoiR algorithm to experimental data. a: Force-distance curve obtained by atomic force spectroscopy measurements of the interaction of human T lymphocytes with the integrin VLA-4 ligand VCAM-1. The curve shows the typical signature of membrane tether formation. b: A force barrier can be observed after noise reduction and was determined to be (10.4 ± 0.5) pN (indicated by the red lines). c: Position of a magnetic bead attached to a DNA molecule measured with magnetic tweezers. The bead is fixed and the focus is shifted bidirectionally by about 0.8 nm. After noise reduction by SWT/SUREShrink or ReNoiR, the relative movement becomes visible. d: A histogram over the data filtered by ReNoiR reveals the two focus positions. Without noise reduction or with SWT, the distributions cannot be distinguished. e: Surface-enhanced Raman spectroscopy (SERS) spectrum of a silver substrate covered by dodecanethiol. Some peak locations (orange) only become visible after application of the ReNoiR algorithm (black) to the raw data (gray).

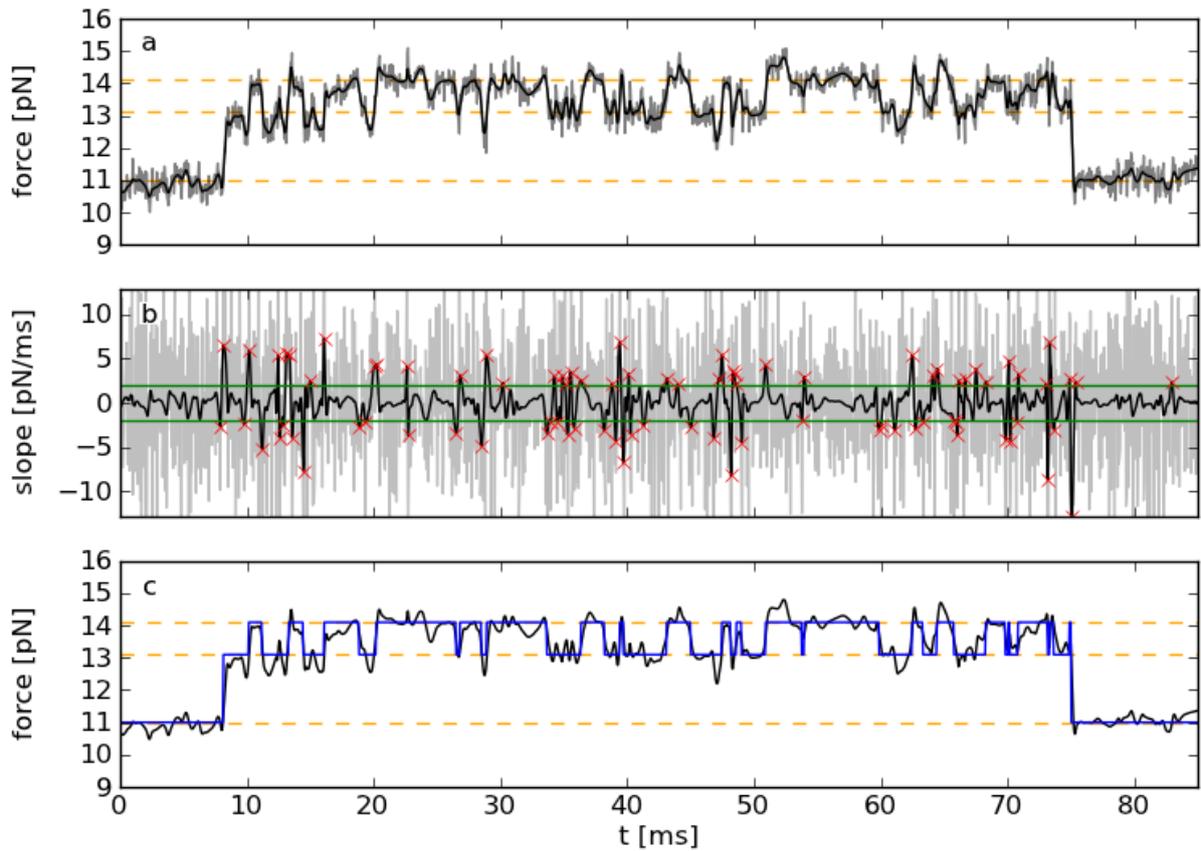


Fig. 17: a: Force vs. time trace of a leucine zipper held at a pretension of 14.1 pN by two optical tweezers at constant trap separation before (gray) and after noise reduction (black). b: The numerical derivative of the de-noised data was thresholded (green) to locate possible folding/unfolding transitions (red). c: By calculating a histogram over the de-noised data, the unfolded and two intermediate states (11.0, 13.1, and 14.1 pN, respectively) could be resolved (orange). The blue curve shows the fits to these states.

2.2.3. Ideal filter parameters

Automatic threshold selection techniques are far from a (theoretical) optimum, because they can only be grounded on estimates about the root mean square error

$$\text{RMS}(\tilde{s}) = \sqrt{(\tilde{s} - s)^2 / N} \quad (25)$$

corresponding to a choice, because s is a priori unknown. However, in some applications other selection criteria exist, so that no automatic parameter optimization is necessary. For instance, if s

reflects a number of discrete states concealed by large noise amplitudes, the filter settings can be optimized to obtain sharp peaks in a histogram of \tilde{s} (fig. 16d). In such cases, the theoretical optimum, i.e. the signal recovery performance attainable by ideal filter settings in the sense of minimum RMS, is more relevant. Further, a comparison of maximum attainable noise reduction efficiencies is only possible with optimal settings, because the efficiency of an automatic threshold optimization algorithm would influence the result.

For these reasons, ReNoiR was compared with SWT using ideal thresholds, too. These parameters were obtained for both methods, every signal, and every noise amplitude by minimizing the deviation between the clean synthetic test signal s and the estimate \tilde{s} , which was quantified by the RMS. With ideal parameters, ReNoiR cannot perform worse than SWT, as the recursive expansion is effectively omitted for $R = 1$ and $T_1 \rightarrow \infty$ (see section 2.1.2). In this case, both algorithms are equivalent. For most test signals and noise levels, the recursive approach showed better results (fig. 18 and 19). If the merging step is skipped (by setting $T_0 \rightarrow \infty$) and only T_1 is optimized, the attained output SNRs are significantly lower. This indicates that both the recursive expansion and the merging process are responsible for the increased performance of ReNoiR.

Two classical smoothing techniques were also included in the comparison: The Savitzky-Golay method consists in convolving a signal with a polynomial kernel⁴³. Here, a polynomial of degree 5 was used for the kernel with varying length w . Gaussian smoothing is another common method to reduce noise⁴⁴. It can be expressed as a convolution with the normal distribution

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right) \quad (26)$$

characterized by the standard deviation σ . Again, the parameters w and σ were selected so as to minimize the RMS. Gaussian smoothing attained the same or better output SNRs than the wavelet-based methods at very high noise levels – given that an ideal σ can also be found without the prior knowledge of the clean signal s .

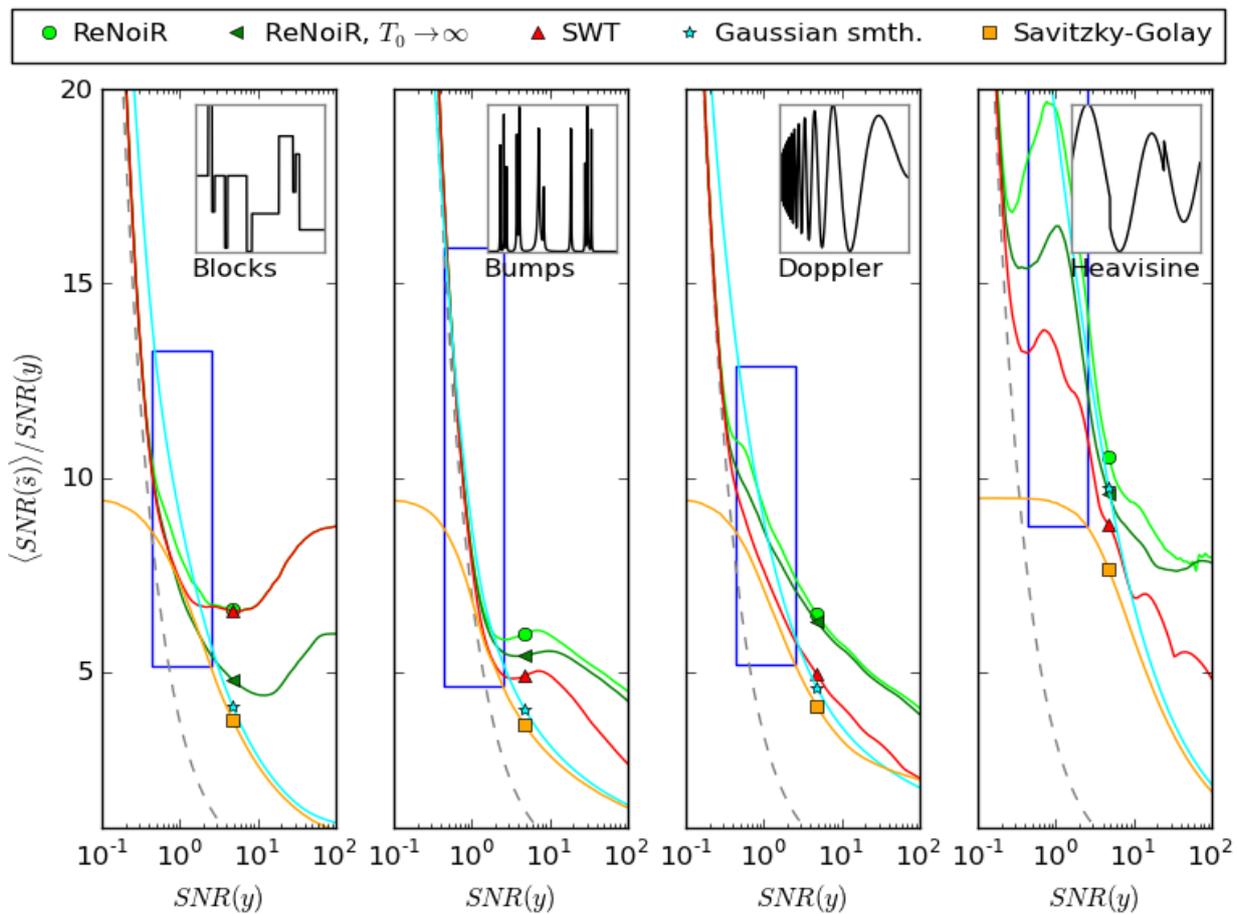


Fig. 18: Signal recovery performance obtained with ideal thresholds in the sense of minimal RMS error. In the limit $T_0 \rightarrow \infty$, the ReNoiR merging process is omitted and the output SNRs are significantly decreased. The Savitzky-Golay algorithm is not suited for de-noising. Gaussian smoothing yields good results for very high noise levels. As in fig. 11, the four standard test signals shown in the insets are contaminated with additive white Gaussian noise. Each data point represents an average value obtained from 100 curves with distinct random noise.

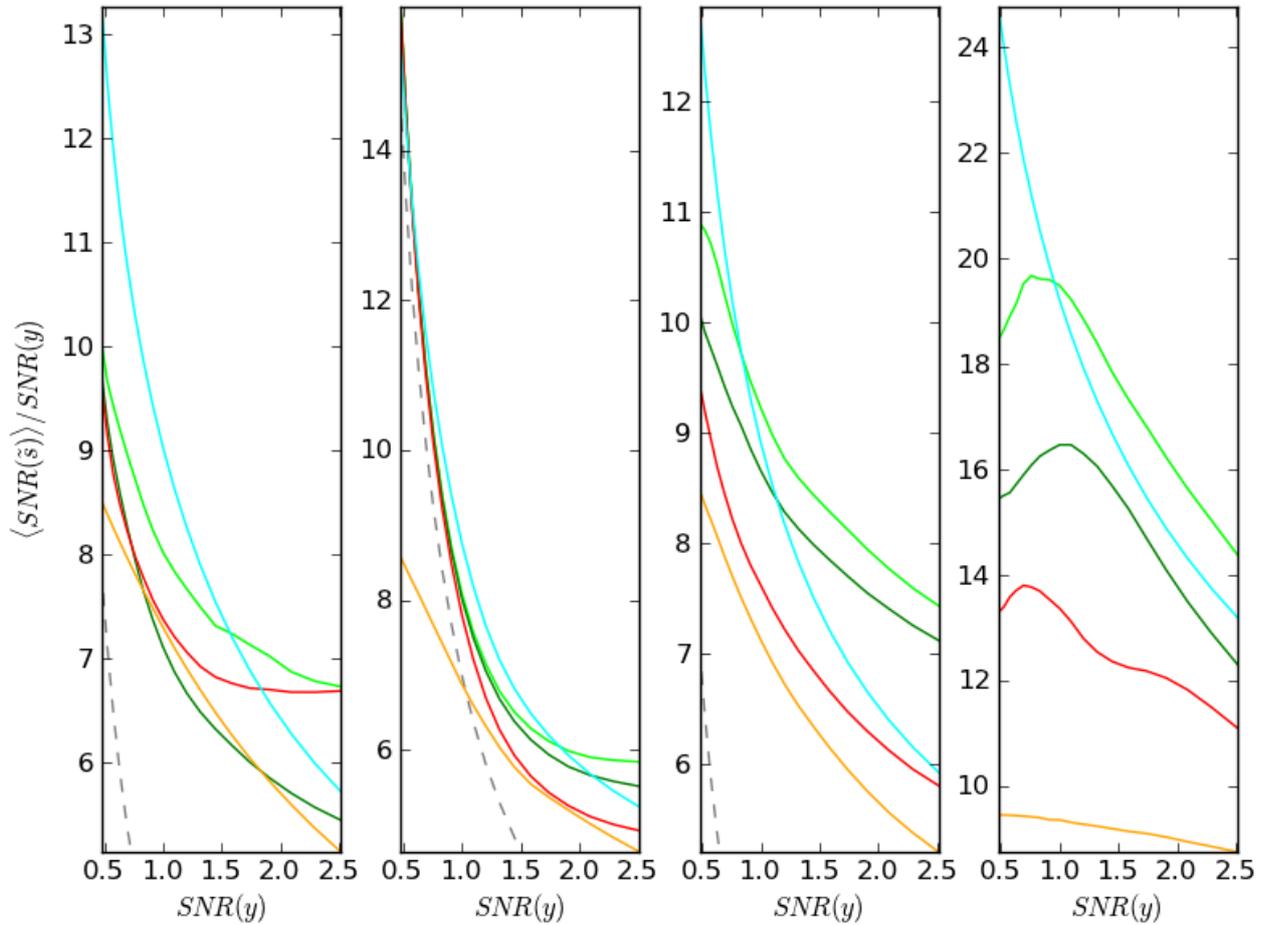


Fig. 19: Zoomed views of the areas marked by the blue rectangles in fig. 18.

2.3. Discussion

ReNoiR aims at averaging out uncorrelated fluctuations. This is also the fundamental concept of SWT, but the increased redundancy originating from the recursive wavelet expansion provides a more efficient way to suppress these fluctuations. By merging the wavelet coefficients resulting from different recursion levels (fig. 6), errors induced by conventional thresholding are reduced. This contrasts with de-noising based on conventional SWT, where small-scale components below a threshold are set to zero (fig. 5). As a consequence, important information is lost, signal recovery is impaired and pseudo-Gibbs artifacts can arise in the reconstructed signal. These effects are reduced by replacing values below the threshold with the filtered coefficients emerging from recursion level $r + 1$ (fig. 20). Higher values are in general less affected by noise and can be left unchanged, as they mainly originate from the deeper levels of decomposition and thus comprise less detail information. If the merging is omitted by setting $T_0 \rightarrow \infty$ (i.e. only the recursive

expansion is performed; see section 2.1.2), noise is reduced less efficiently (fig. 18 and 19), indicating the importance of the procedure.

Conventional SWT-based de-noising can be regarded as a special case of ReNoiR, as both algorithms are equivalent if the recursive expansion is omitted by setting $R = 0$. Effectively, only the red and blue elements marked in fig. 7 and 8 are used under this condition.

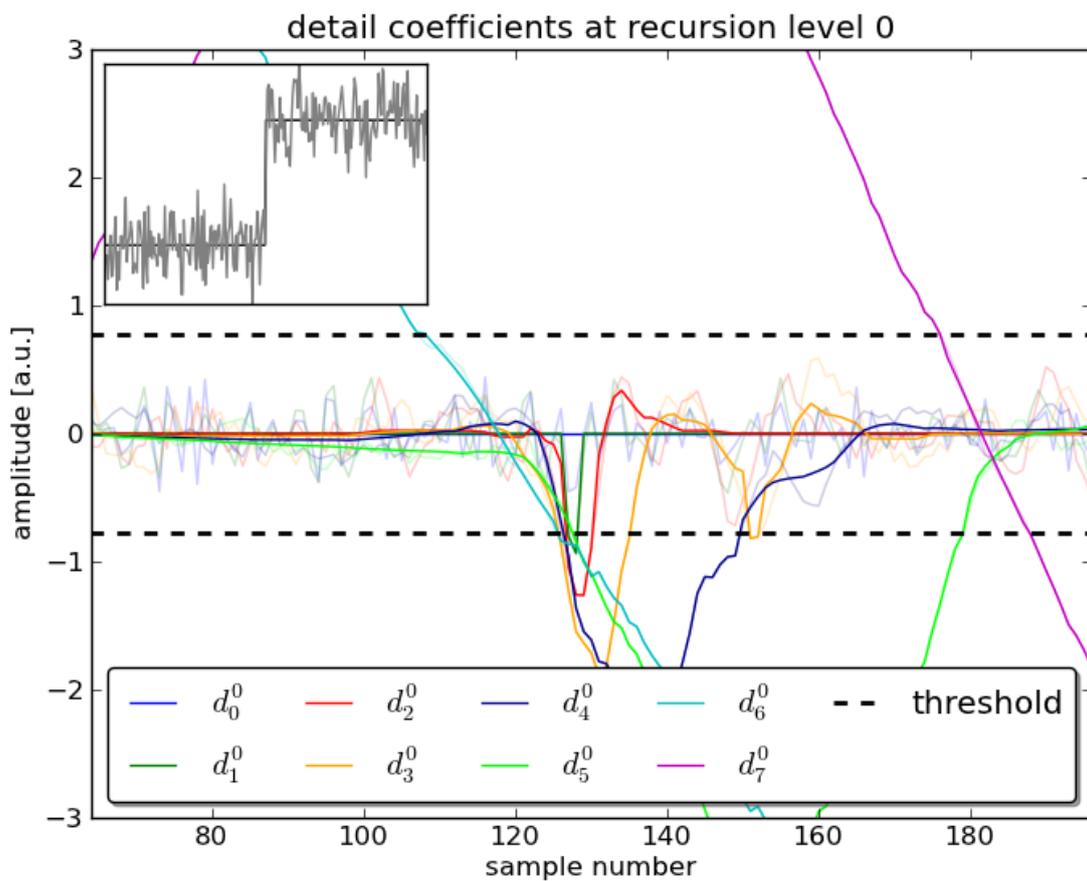


Fig. 20: Detail coefficients obtained from the decomposition of the noisy step signal shown in the inset before (light colors) and after the merging process (dark colors). Values below the threshold (dashed black lines) are replaced with a de-noised version resulting from the next recursion level.

3. Step detection

Experimental data obtained by observing a biological system at microscopic scales often reflects multiple discrete states, e.g. the disruption of intermolecular bonds^{61, 62}, the unfolding of proteins^{45, 63}, or the stepwise movement of a molecular motor⁶⁴⁻⁶⁶. The study of such phenomena requires highly sensitive single-molecule techniques like optical/magnetic tweezers or atomic force microscopy. They are common tools to examine the kinetics of molecular bonds or enzymatic activity⁶⁷, because their spatial and force resolution is sufficient to analyze events on a nm and pN scale, respectively. In contrast to ensemble measurements, they allow revealing the mechanical properties of individual proteins and resolving conformational changes. For example, observation of the unbinding of membrane tubes pulled from living cells would not be possible without single-molecule techniques. However, the measured signal is impaired by thermal fluctuations, electronic noise, and vibrations, as the disturbances are generally of the same order of magnitude as the investigated events. To analyze such data, an automated method to identify the steps marking the transitions between the discrete states of the investigated system is necessary. In a comparison of existing detection algorithms, an iterative fit procedure (“ χ^2 method”) proposed by Kerssemakers et al.⁶⁸ combined with a moving average filter showed the best overall performance⁶⁹. The window size of the mean filter can be optimized for the types of steps to be recognized.

It is obvious that any information about the steps, such as their average width or signal-to-noise ratio (SNR), i.e. the ratio between height and the standard deviation of the noise, can be used to increase the probability of successful detection. In practical applications, these properties are often very similar for all steps, and their approximate heights and widths are usually known. The noise level can generally be determined from the measured data. Here, it is demonstrated that significantly higher detection rates can be obtained by a novel moving step fit (MSF) algorithm, which makes use of this information. In contrast to other methods optimized to identify changes between a small number of identical states⁷⁰, MSF is intended to reveal transitions between arbitrary discrete states. It allows the identification of low steps hidden in experimental data, which have been unrecognized before due to very low SNRs. By adjusting its fit window size, steps can be detected within two extreme cases: low and wide steps with heights far below the noise amplitude, as well as higher, but very narrow steps separated by only a few data points.

3.1. Materials and methods

3.1.1. Detection method of Kerssemakers et al.

In contrast to the original approach⁶⁸, a windowed mean filter replacing each data point in the middle of $2\lambda + 1$ consecutive points with their average value was applied to the noisy data beforehand, because that has been shown to increase the detection rate⁶⁹. After adding the mean filter, the Matlab implementation (The Mathworks, Natick, MA) kindly provided by J. Kerssemakers was used. Briefly, a single step is fitted to every possible position of the time trace, and the data is partitioned at the location corresponding to the smallest χ^2 sum. The procedure is repeated iteratively with both resulting parts until the χ^2 sum has been determined for every data point. Values below a threshold correspond to possible step positions.

3.1.2. MSF algorithm

Initially, the noisy data y_i sampled at discrete time or space intervals x_i is pre-processed by convolution with a Gaussian kernel⁴⁴ with standard deviation σ . Thereby, both the source signal and the noise are smoothed, but continuous parts, such as plateaus between steps, are preserved. Then, a step of height h_i is fitted at position i in the middle of a moving window of size $2w$ ($w \leq i < N - w$). Here, a piecewise linear fit function

$$f_i(x_j) = \begin{cases} m_i x_j + t_i^l, & i - w \leq j < i \\ m_i x_j + t_i^r, & i \leq j < i + w \end{cases} \quad (27)$$

with uniform slope but different constant terms was used. A second, continuous function

$$g_i(x_j) = m_i^0 x_j + t_i^0, \quad i - w \leq j < i + w \quad (28)$$

is fitted to the whole window. In both cases, the global optima of the free parameters are obtained analytically:

$$m_i = \frac{\frac{N}{2} \sum_{i-w}^{i+w-1} x_j y_j - (\sum_{i-w}^{i-1} x_j)(\sum_{i-w}^{i-1} y_j) - (\sum_i^{i+w-1} x_j)(\sum_i^{i+w-1} y_j)}{\frac{N}{2} \sum_{i-w}^{i+w-1} x_j^2 - (\sum_{i-w}^{i-1} x_j)^2 - (\sum_i^{i+w-1} x_j)^2} \quad (29)$$

$$t_i^l = \frac{2}{N} \left(\sum_{i-w}^{i-1} y_j - m_i \sum_{i-w}^{i-1} x_j \right) \quad (30)$$

$$t_i^r = \frac{2}{N} \left(\sum_i^{i+w-1} y_j - m_i \sum_i^{i+w-1} x_j \right) \quad (31)$$

$$m_i^0 = \frac{N \sum_{i-w}^{i+w-1} x_j y_j - (\sum_{i-w}^{i+w-1} x_j)(\sum_{i-w}^{i+w-1} y_j)}{N \sum_{i-w}^{i+w-1} x_j^2 - (\sum_{i-w}^{i+w-1} x_j)^2} \quad (32)$$

$$t_i^0 = \frac{1}{N} \left(\sum_{i-w}^{i+w-1} y_j - m_i^0 \sum_{i-w}^{i+w-1} x_j \right) \quad (33)$$

At each position the residual sum of squares (RSS) is calculated for f_i and g_i . The term

$$\vartheta_i = [\text{RSS}(g_i) - \text{RSS}(f_i)]h_i = \left(\sum_{i-w}^{i+w-1} [g_i(x_i) - y_i]^2 - \sum_{i-w}^{i+w-1} [f_i(x_i) - y_i]^2 \right) (t_i^r - t_i^l) \quad (34)$$

only takes on high values if the step function f_i fits better than the continuous function g_i . Multiplication by the step height h_i is optional and assures that large steps are more likely to be detected. Therefore, ϑ_i is an indicator for the probability of i to be a potential step position. Consequently, local maxima exceeding a threshold significantly above the statistical fluctuations of ϑ can be regarded as steps. If the number of steps S is a priori known, the S highest local maxima define the step positions instead.

3.1.3. Generation of the test signals

Different types of test signals were used to evaluate the χ^2 and MSF method (see table 6 and fig. 21):

Data set A consists of curves with $S = 2$ steps of height one separated by a variable distance (see example in fig. 22a).

For data set B, artificial force spectra were generated by Monte Carlo simulations (see example in fig. 22b). They contain a given number S of steps with randomly determined positions and heights.

In brief, for $N = 8192$ tether extensions z_i sampled at small time intervals $\Delta t = z_N/Nv$, the rupture probability $p_i = k_{\text{off}} \cdot \Delta t$ was calculated from the force-dependent off-rate⁷¹

$$k_{\text{off}}(F_i) = k_{\text{off}}^0 \exp\left(\frac{F_i w}{kT}\right) \quad (35)$$

using the force

$$F_i = F(z_i) = k_1 z_i + \eta v \left[1 - \exp\left(-\frac{k_2 z_i}{\eta v}\right)\right] \quad (36)$$

exerted by a Kelvin body representing a single tether⁵⁵. Each p_i was compared with a random number r_i in the range $[0,1[$ and the first occurrence of $p_i > r_i$ was considered a rupture event, i.e. F_j was set to zero for $i < j < N$. To obtain curves with multiple steps, the procedure was repeated and the forces F_i were summed up. The experiment-specific parameters were chosen to mimic real single-molecule force spectroscopy data obtained with biological cells^{55, 72}: $k_1 = 1.6$ pN/ μm , $k_2 = 260$ pN/ μm , $\eta = 5.9$ $\mu\text{Ns/m}$, $k_{\text{off}}^0 = 0.1$ s^{-1} , $w = 1.8$ \AA , $T = 36$ $^\circ\text{C}$, $z_N = 16.0$ μm , $v = 3.4$ $\mu\text{m/s}$. As it is common practice, forces are plotted with reversed sign.

To allow for a quantification of force resolution, data set C was designed to contain steps with discrete heights – as opposed to the continuous distribution of step heights resulting from the Monte Carlo simulations. Exactly four steps with randomly chosen heights (5, 10, 20, or 40 pN) were placed at fixed positions (2, 6, 10, and 14 μm). To resemble data set B, linear plateaus with slope k_1 were created between the steps.

The signals of data set A were contaminated with additive white Gaussian noise. The artificial force curves of data sets B were superimposed by normal-distributed random noise with a frequency spectrum measured by a Nanowizard II atomic force microscope (JPK, Berlin, Germany), and both types of noise were applied to data set C.

White Gaussian noise was created by a Box-Muller transform⁷³ of uniformly distributed random numbers generated by the Mersenne Twister algorithm⁷⁴. To reproduce unlimited amounts of the instrumental noise, force signals free of sample-specific effects were recorded with the atomic force microscope and de-trended by subtraction of a linear baseline. Their average spectral distribution was calculated by fast Fourier transform using a block size of 2^8 and a Hann window to

reduce spectral leakage. Noise signals of various amplitudes were generated by inverse Fourier transforms of this spectrum with uniformly distributed random phase shifts.

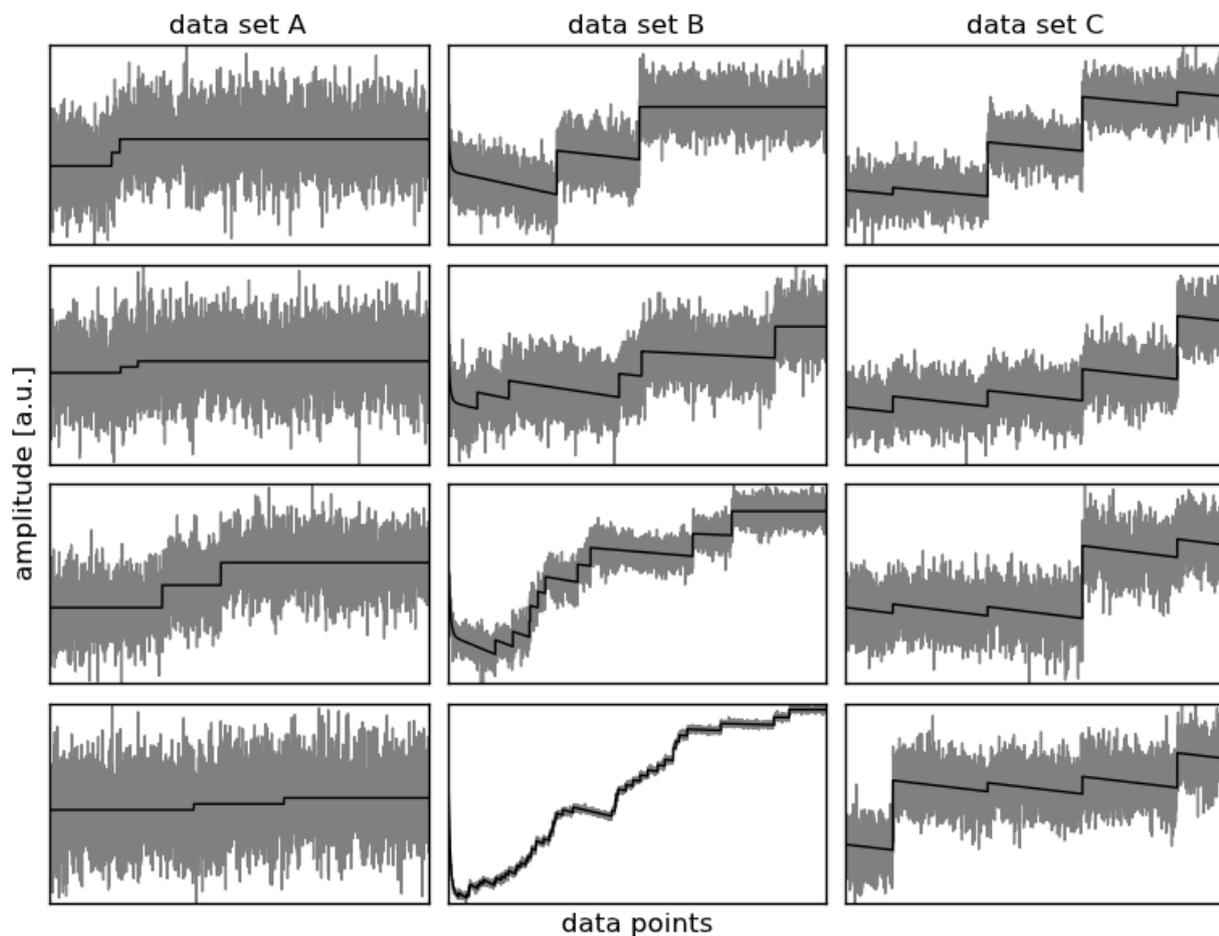


Fig. 21: Examples of the three types of test signals deployed for data analysis (A: constant plateaus separated by two steps of height one at a variable distance contaminated by additive white Gaussian noise, B: artificial force-distance curves mimicking single-molecule force spectroscopy experiments with living cells superimposed by AFM noise, C: like B, but exactly 4 steps at 2, 6, 10, and 14 μm with discrete heights randomly chosen from 5, 10, 20, and 40 pN).

data set	description
A	2 steps with constant plateaus separated by a variable distance superimposed by additive white Gaussian noise of variable amplitudes (100 curves for each step distance and noise level with N = 4200 data points)
B	Monte Carlo simulations of force curves obtained by atomic force spectroscopy with a variable number of steps (1 to 100) at random positions superimposed by instrumental noise of variable amplitudes (100 curves for each number of steps and noise level with N = 8192 data points)
C	artificial force curves containing 4 steps at fixed positions (2, 6, 10, and 14 μm) with randomly selected heights (5, 10, 20, or 40 pN) superimposed by white or instrumental noise with a standard deviation of 10 pN (1000 curves with N = 8192 data points)

Table 6: Test signals used for the evaluation of the step detection algorithms

3.1.4. Evaluation of the step detection algorithms

To evaluate the step detection performance of the algorithms described above, their efficiencies to localize the steps contained in the noisy test signals were quantified. The problem of finding an estimate for the true number of steps S was excluded from the analysis, because a wrong number would affect the results, so that an unbiased comparison of the actual detection performance would not be possible.

Both detection methods generate a measure for the estimated probability (“significance”) of any data point to be a step. The S highest local maxima of this indicator define the (potentially false) identified step positions (fig. 22e and f). A detected step was rated a false-positive if the deviation from its true position was greater than ± 4 data points, and a true step was rated a false-negative if the deviation from its detected position was greater than ± 4 data points. Since S step candidates were tested, each missed true step implied the false detection of a non-existent step and vice versa. Hence, the numbers of false-positives and false-negatives are equal for each test signal. The numbers of false detections were recorded and the rate of successful detections was calculated according to the formula

$$\text{detection rate} := \frac{S - \text{number of false detections}}{S} \quad (37)$$

For each noise-free test signal, detection method, and noise level, the evaluation was repeated at least 100 times with distinct random noise.

The method-specific parameters were chosen to maximize detection rates for the average SNRs and widths of the steps contained in the test signals. This was only possible, because these properties were a priori known. In practical applications, optimal settings must be determined either manually or calibrated by simulated data (see section 3.2.8). Data evaluation was done with Matlab and Python.

3.1.5. Calculation of the step heights

Linear fits

$$f_{\text{left}}(x) = m_{\text{left}}x + t_{\text{left}} \quad (38)$$

$$f_{\text{right}}(x) = m_{\text{right}}x + t_{\text{right}} \quad (39)$$

were performed over up to 2048 data points to the left and to the right of the identified step positions x_{step} , but no further than to the neighboring detected steps. The step heights were determined from the difference of the values of both fit functions at the positions of the steps:

$$h = f_{\text{right}}(x_{\text{step}}) - f_{\text{left}}(x_{\text{step}}) = (m_{\text{right}} - m_{\text{left}})x_{\text{step}} + t_{\text{right}} - t_{\text{left}} \quad (40)$$

3.1.6. Recording of the AFM spectra

A Nanowizard II atomic force microscope (JPK, Berlin, Germany) was deployed to measure force-distance curves of β_1 integrin-deficient Jurkat A1 lymphocytes with re-substituted β_1 integrin⁵⁴ interacting with the VLA-4 ligand VCAM-1 as described by Schmitz et al.⁵⁵. A VCAM-1 concentration of 2 $\mu\text{g}/\text{ml}$ and a constant approach/retract velocity of 3.4 $\mu\text{m}/\text{s}$ were used.

3.2. Results and discussion

3.2.1. Step detection performance

To render a quantitative evaluation of correct and false detections possible, the step positions must be a priori known. For that reason, the step detection performance of MSF was compared with the χ^2 method⁶⁸ by means of synthetic test signals (fig. 21). The χ^2 method was chosen as reference, because it has been shown to perform best among other highly efficient techniques⁶⁹.

Two different types of signals were analyzed (table 6): First, simple curves with two steps of height one separated by a variable distance were deployed to study the influence of the SNR and of the distance between successive steps on the detection rates (data set A, see example in fig. 22a, c). The clean signals were contaminated by additive white Gaussian noise. Second, artificial force-distance curves mimicking single-molecule force spectroscopy experiments with living cells were created by Monte Carlo simulations and superimposed by the characteristic instrumental noise of a JPK Nanowizard II AFM (data set B, see example in fig. 22b, d). Such experiments are highly relevant to understand cell-surface or cell-cell adhesion and cellular force sensing⁵⁵. Every simulated curve contains a predefined number S of steps at random positions. Since the χ^2 method requires a manual selection of the number of steps to be detected, a comparison how accurately S can be determined is not possible. Thus, the number of steps was assumed to be known, i.e. both algorithms were configured to detect the S most significant steps. As a consequence, the number of false-positives and false-negatives is equal and needs not be compared separately. For practical applications, an automatic selection of S might be required. Therefore, the MSF algorithm can also be deployed with a given detection sensitivity, i.e. any steps with a significance exceeding a given threshold are detected (see section 3.1.2).

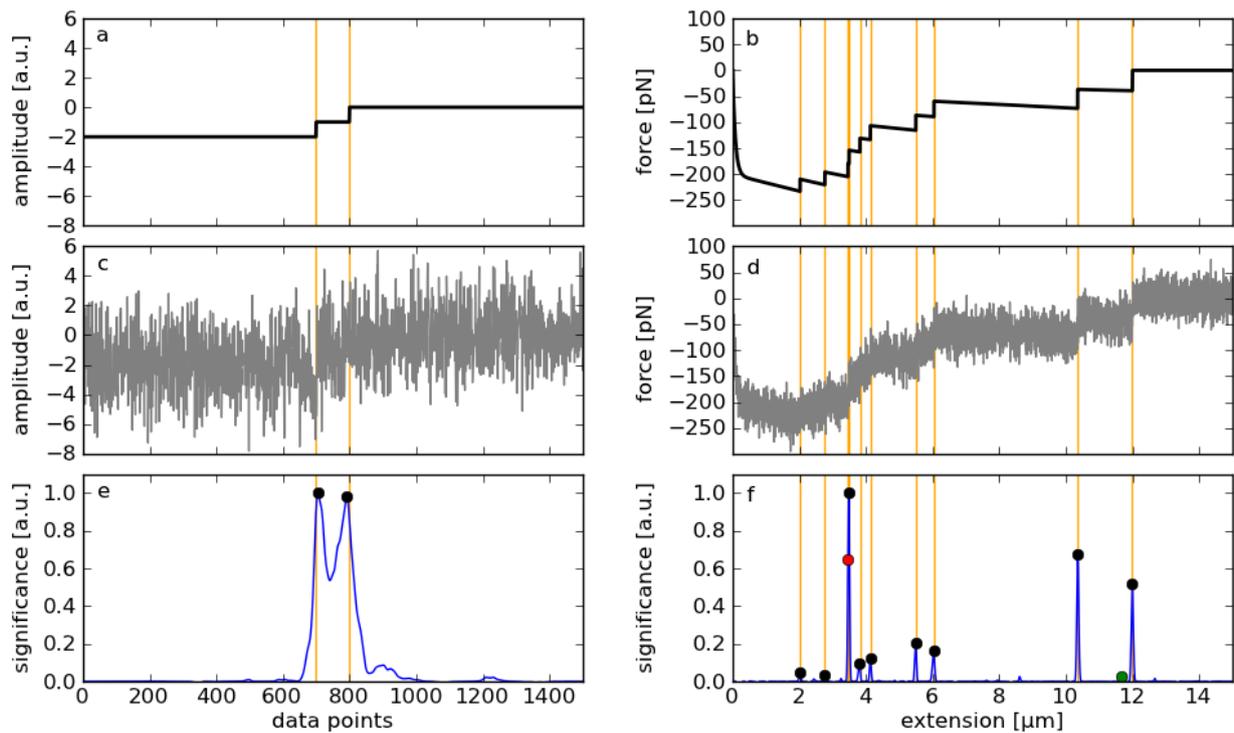
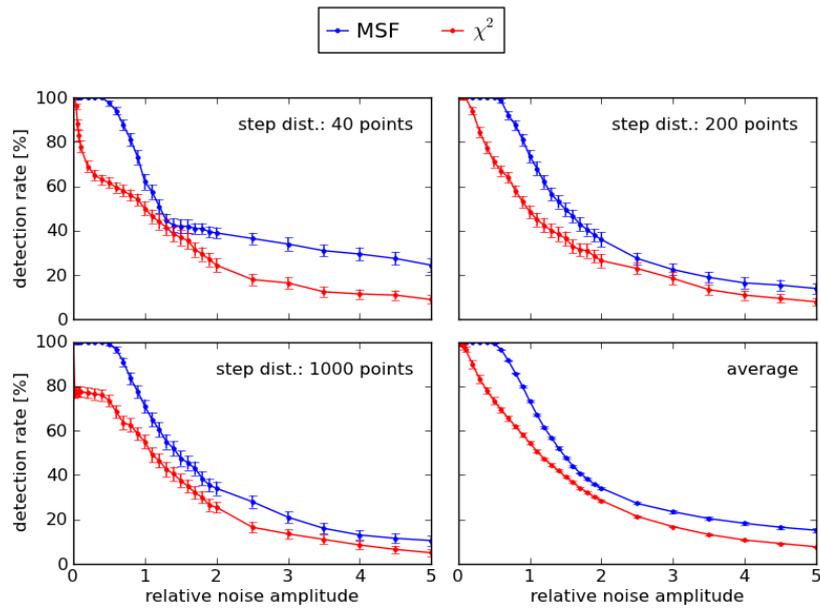


Fig. 22: Evaluation of the MSF algorithm. a: Synthetic test signal containing two steps of height one separated by a defined distance (data set A). b: Artificial AFM spectrum generated by Monte Carlo simulations mimicking an idealized (noise-free) force-distance curve typically obtained by cell adhesion measurements (data set B). c: Clean signal superimposed by normal-distributed white noise. d: Clean signal superimposed by random AFM noise with a standard deviation of 20 pN. e, f: Indicators of possible step positions calculated from the noisy signals by the MSF algorithm (blue). Local maxima were used to identify the steps. Orange lines mark the true step positions, black dots correct detections, the green dot a false-positive, and the red dot a false-negative.

The efficiencies of both step detection methods depend not only on the signal and noise characteristics, but also on the choice of parameters: The χ^2 algorithm can be optimized by varying the window size λ of the moving average filter and MSF by varying the width of the smoothing kernel σ and the half window size w (see sections 3.1.1 and 3.1.2). As optimal parameters in the sense of maximum detection rates depend on the SNRs and widths of the steps in a complex way, they were determined numerically by evaluating test signals with pre-defined characteristics and a priori known step positions (see section 3.2.8). In doing so, four scenarios were considered, each for data set A and B: First, the average SNR and either the average width (for set A) or the number of the steps (for set B) were assumed to be known and fairly constant, i.e. both facts were used for parameter selection (fig. 23a and b). In the second case, very different step widths (or numbers of steps) can occur, i.e. parameters were optimized for each noise level (i.e. the constant SNR for data set A and an average SNR for data set B) and for a broad range of step distances (fig. 24a and b). Third, parameters were chosen yielding highest detection rates for an approximately constant SNR and variable step distances (fig. 24c and d). If neither the SNRs nor the widths of the steps can be narrowed down, optimization must be performed for arbitrary step characteristics within a reasonable range (fig. 24e and f), resulting in constant parameters for all test signals ($\sigma = 2$, $w = 30$, and $\lambda = 4$).

Generally, attainable detection rates depend on the step characteristics and on the type of noise. If either the SNRs or the widths of the steps are similar, detection rates can be highly improved by MSF in comparison to the χ^2 method. MSF is particularly effective for curves with many or narrow steps. Even if both the heights and the widths of the steps vary, it yields higher detection rates in many cases for both types of test signals and shows the best overall performance with about 20% higher detection rates than the χ^2 algorithm for data set A and about 30% for data set B (table 7). The Kerssemakers method combined with the moving average filter works well with curves containing few steps. It is less effective in general, because the χ^2 sum calculated at a potential step position is increased by other steps, making it a less sensitive indicator. Only curves with at least 2 steps could be included in the analysis, as the algorithm fails if the number of steps to be detected is set to 1. Further, the decay in the beginning of the artificial force curves of data set B impairs the method and results in some false-positive detections.

a



b

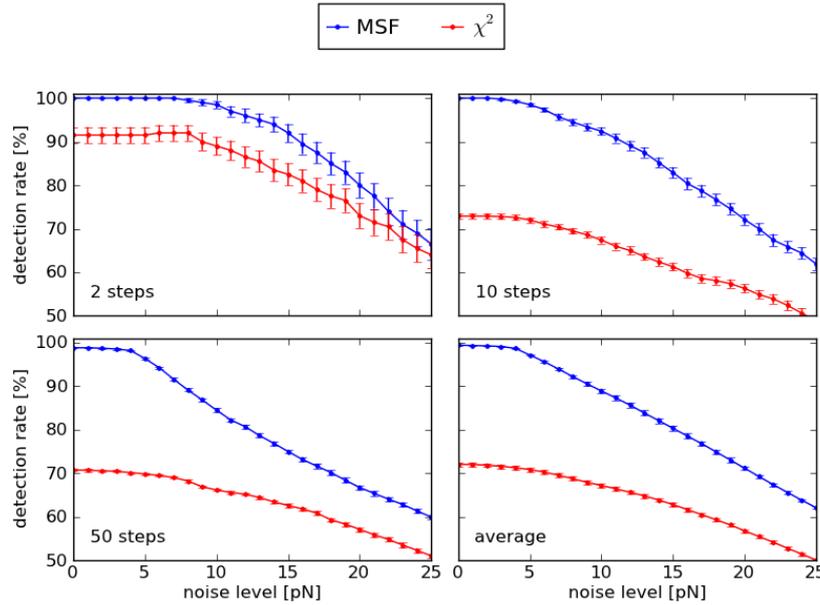


Fig. 23: Detection rates vs. noise amplitudes of the MSF and the χ^2 method applied to synthetic test signals. Every marker represents the average rate over 100 curves with distinct random noise, with error bars indicating the standard errors. a: Two steps of height one separated by a variable distance (data set A, see example in fig. 22c). The last plot shows the average detection rates over 33 step distances between 40 and 1000 data points. Detection parameters were optimized for the constant SNRs and distances of the steps. b: Artificial force curves generated by Monte Carlo simulations with varied number of steps (data set B, see example in fig. 22d). Parameter optimization was performed for each noise level and number of steps. The last plot shows the average detection rates over the range of 2, 3, ..., 50 steps.

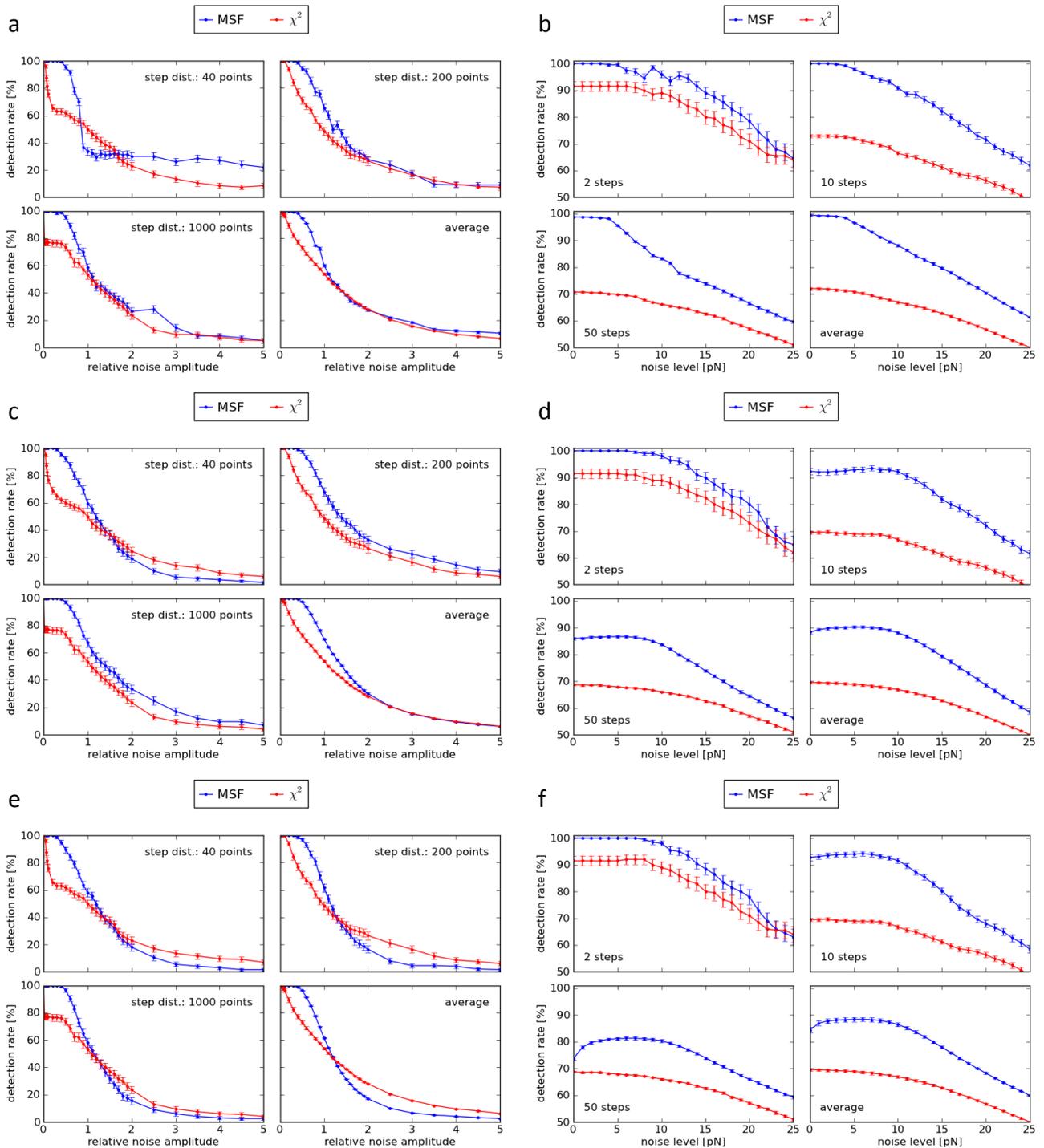


Figure 24: Detection rates vs. noise amplitudes of the MSF and the χ^2 algorithm applied to synthetic test signals for different optimization methods. Every marker represents the average rate over 100 curves with distinct random noise, with error bars indicating the standard errors. a, c, e: Results for data set A with parameters optimized for (a) each SNR of the steps and variable width, (c) variable SNR and each width, and (e) for variable SNR and width. The last plot shows the average detection rates over 33 step distances between 40 and 1000 data points. b, d, f: Results for data set B with parameters optimized for (b) each noise level and a variable number of steps, (d)

variable noise level and each number of steps, and (f) for variable noise level and number of steps. The last plot shows the average detection rates over the range of 2, 3, ..., 50 steps.

step detection method	data set A		data set B	
	absolute detection rate	relative detection rate	absolute detection rate	relative detection rate
MSF	65.4 %	119.9 %	83.6 %	131.3 %
χ^2	54.5 %	100.0 %	63.7 %	100.0 %

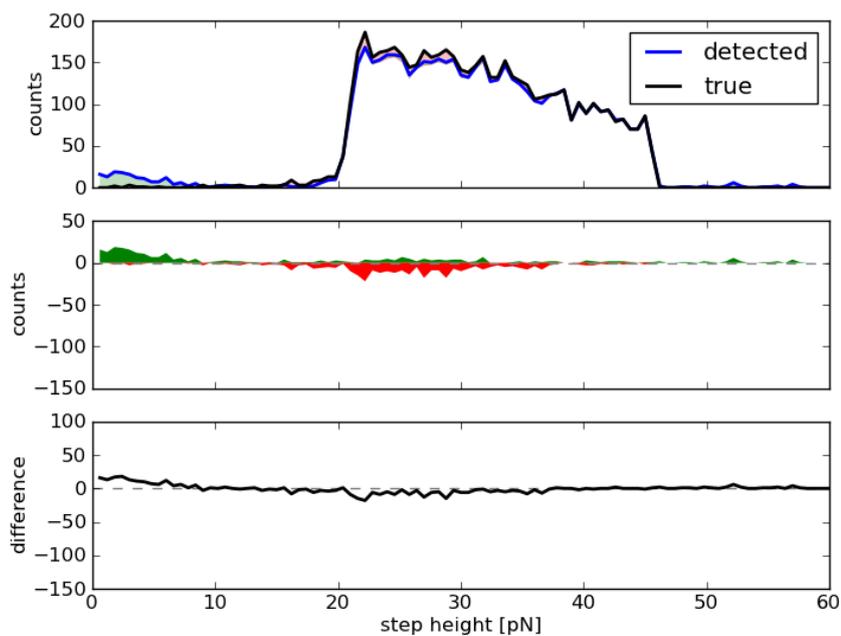
Table 7: Total average detection rates for the data shown in fig. 23. Filter settings were individually optimized for every noise amplitude and step width (data set A) / number of steps (data set B). Relative rates are related to the results obtained by the χ^2 method combined with the windowed mean filter.

3.2.2. Distribution of false-positives and -negatives

The real AFM measurements modeled by the simulated force curves typically show a noise level of about 10 pN and no more than 10 steps. Therefore, any further analysis of data set B was restricted to curves with 2 to 10 steps and 10 pN AFM noise. On these conditions, optimal detection rates were obtained by $\sigma = 3.5$, $w = 30$, $\lambda = 1$. If not stated otherwise, these parameters were used in the following.

A detected/unrecognized step was rated a false-positive/false-negative if the deviation from the nearest true position was greater than ± 4 data points. Otherwise it is correct by definition. Both false-positive and false-negative detections decrease with increasing step heights, as higher steps can be identified more reliably. The number of false-negatives is much lower for MSF than for the χ^2 method and the number of false-positives is similar (fig. 25).

a



b

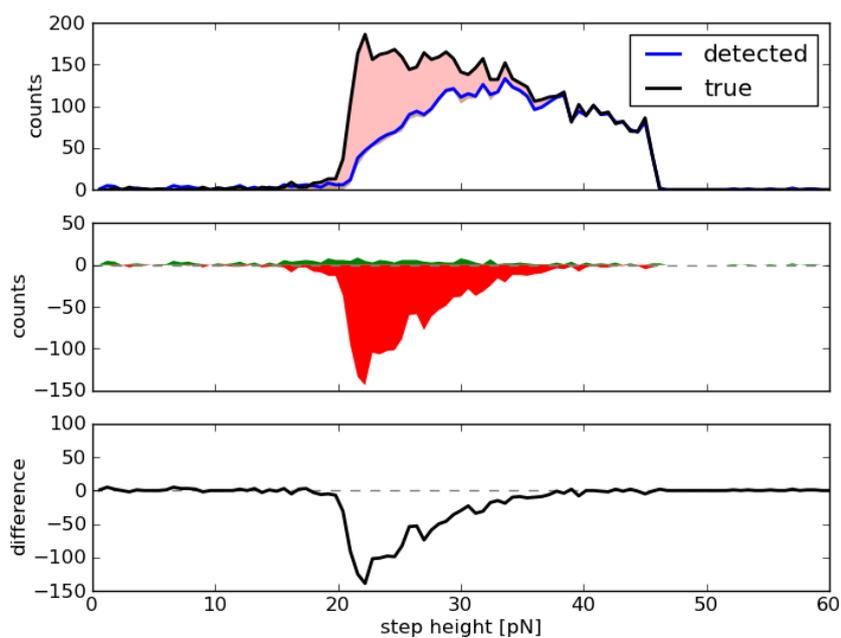


Fig. 25: True (black) and detected (blue) step positions found in data set B by MSF (a) and the χ^2 method (b). The numbers of false-positives (green) and false-negatives (red) are significantly lower for MSF.

3.2.3. Height resolution

To analyze real data, generally not only the step positions, but also their heights must be determined. The latter process depends on the former, and both are error-prone. The precision of the height estimation achievable with MSF and the χ^2 method was quantified by the example of data set C, which also models AFM force curves, but contains exactly 4 steps at 2, 6, 10, and 14 μm with discrete heights randomly chosen from 5, 10, 20, and 40 pN (see table 6 and examples in fig. 21). These modifications rendered it possible to determine the height resolution limit. Again, the signals were contaminated by AFM noise of a single amplitude ($\sigma_{\text{noise}} = 10$ pN) and the method-specific parameters were chosen to maximize the total average detection rates ($\sigma = 4.7$, $w = 40$, $\lambda = 4$).

To resolve the heights of the steps, they must be detected in the first place. MSF yields more false-positives and less false-negatives than the χ^2 method, which does not reproduce the 5 pN peak at all (fig. 26).

The test signals consist of linear plateaus, so that the step heights can be calculated from adjacent linear fits of these plateaus left and right of the identified steps. The calculations were performed according to eq. (38) - (40) with m_{left} and m_{right} fixed to the constant slope of the plateaus (1.6 pN/ μm). The total average deviation between true and estimated heights is (-0.22 ± 1.20) pN, i.e. systematic errors (e.g. arising from false-negative detections within the fit range) are much smaller than the statistical errors resulting from the noise (fig. 27). Step heights determined by the χ^2 algorithm deviate from the true values by (-7.45 ± 6.87) pN. They are considerably underestimated, because they are calculated from the difference of the mean force of the left and right edge, and not by linear fits.

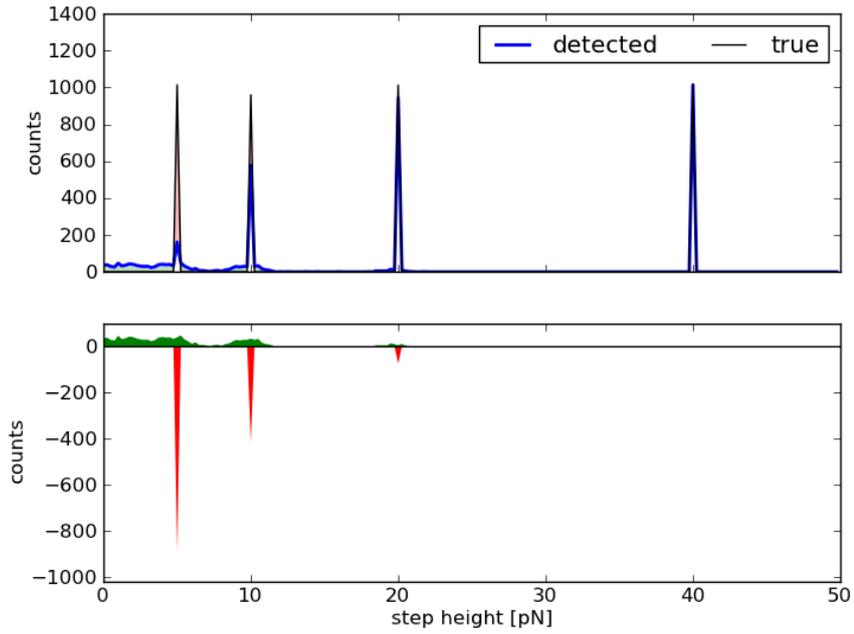
As a consequence of the noise-induced errors, the discrete force distributions are blurred, i.e. the force resolution is reduced. In case of white noise, the standard error σ_{height} of the heights determined from the fits decreases with the square root of the fit length l :

$$\sigma_{\text{height}} = \sqrt{2/l} \cdot \sigma_{\text{noise}} \quad (41)$$

Thus, the distributions of calculated step heights are expected to be Gaussians with standard deviation σ_{height} . If their amplitudes are weighted by the corresponding detection rates, a prediction for the resulting histogram is possible (gray line in fig. 28). False-positive detections

cause the wrong step heights below the 5 pN peak and a reduction of l (see section 3.1.2), so that the actual peaks (green) are slightly broader than predicted ($\sigma_{\text{height}} \approx 0.3$ pN). As eq. (41) is not valid for AFM noise, the resulting peaks (blue) are also wider than expected ($\sigma_{\text{height}} \approx 0.8$ pN). Hence, the height resolution for this kind of test signals is of the order of a pN. As a consequence of the underestimated step heights, the histogram obtained by the χ^2 method (red) is shifted to lower forces by about 5 pN.

a



b

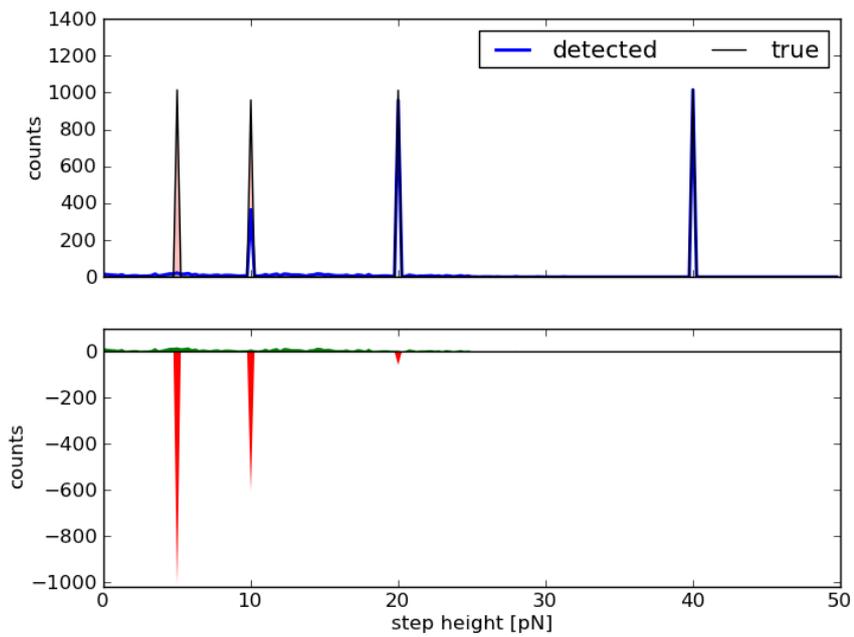


Fig. 26: True (black) and detected (blue) steps as a function of their true heights resulting from application of the MSF (a) and the χ^2 method (b) on data set C. The numbers of false-negatives (red) are significantly higher for the χ^2 method. In contrast to MSF, it yields very few false-positives (green), but also does not reproduce the 5 pN peak at all.

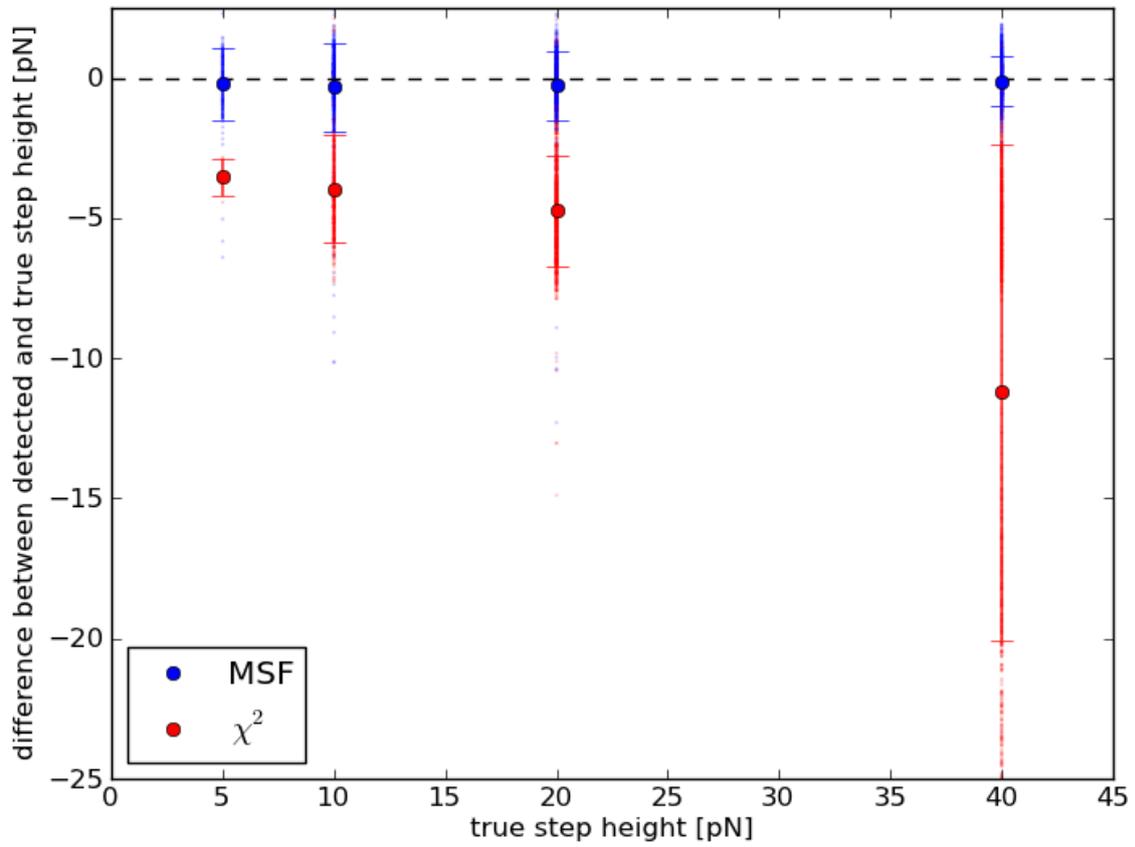


Fig. 27: Mean deviations between calculated and true step heights of data set C. Only steps correctly identified by MSF (blue) or the χ^2 method (red) are included. The error bars indicate the standard deviations.

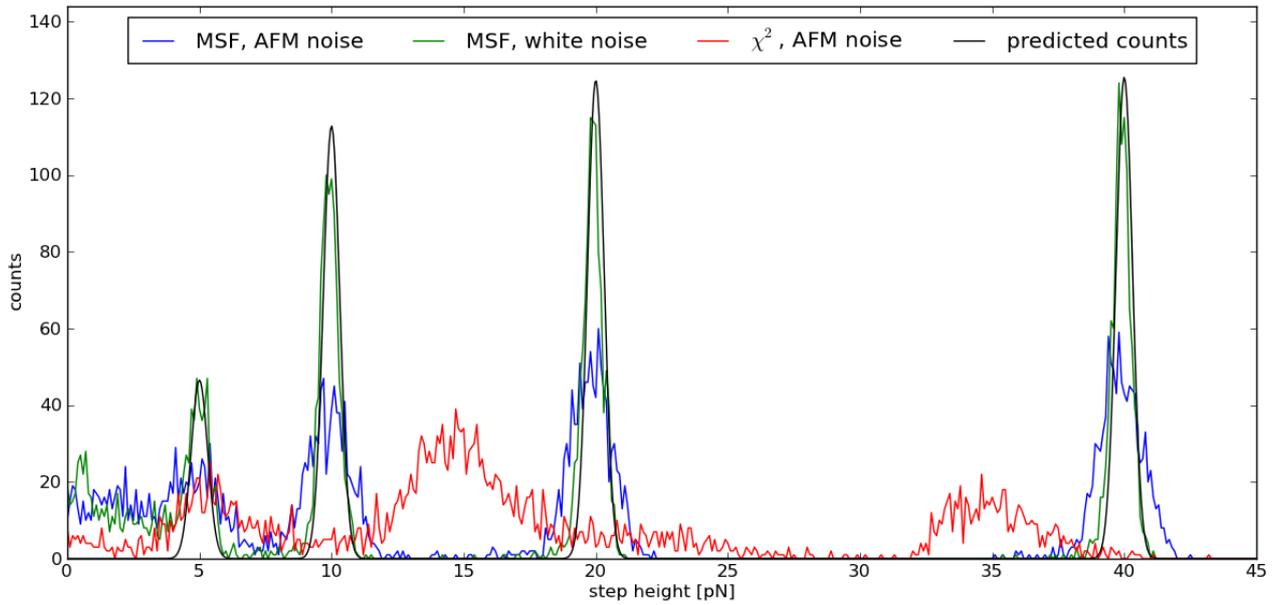
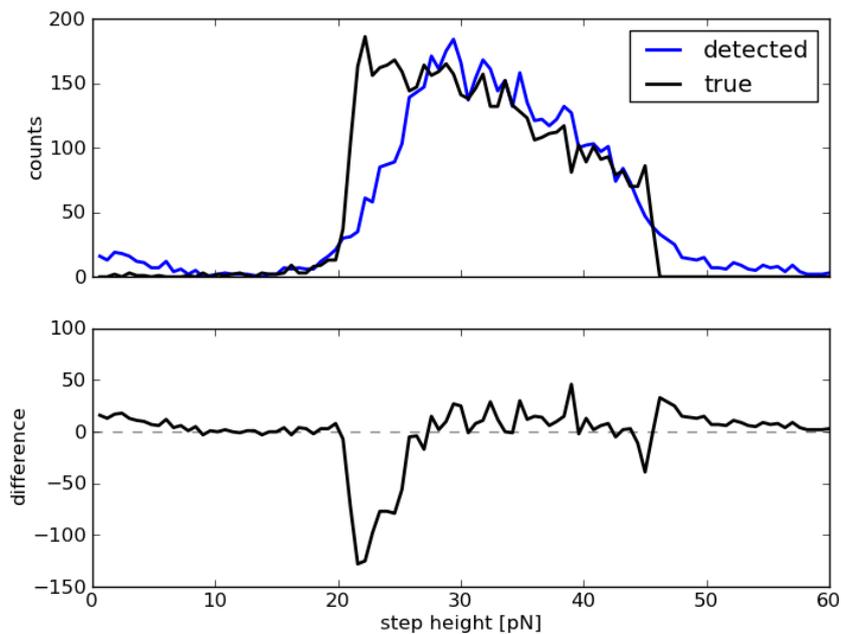


Fig. 28: Number of steps detected by MSF vs. calculated step heights for white noise (green) and AFM noise (blue). The results for the χ^2 method are drawn in red and the predicted histogram for white noise in gray. Independently of the step height, individual distributions can be clearly verified if they are sufficiently far away from each other (less than 5 pN for both types of noise). The colored AFM noise results in broader peaks.

3.2.4. Reproduction of continuous height distributions

Steps heights encountered in real data are generally not restricted to discrete values, but are continuously distributed. In practical applications, the recovery of these distributions can be highly relevant, e.g. for the analysis of force spectroscopy data. To this end, test signals of data set B with 2 to 10 steps (100 each) superimposed with AFM noise (standard deviation 10 pN) were analyzed and the step heights obtained by the χ^2 method were compared with an approach based on linear fits left and right of the step position (see section 3.1.5). Noise-induced errors impair both techniques, so that the calculated heights differ significantly from the true values (fig. 29). The linear fits reproduce the continuous height distribution well for step heights above 25 pN. The χ^2 method underestimates all heights, and the shape of the resulting distribution does not resemble the actual one.

a



b

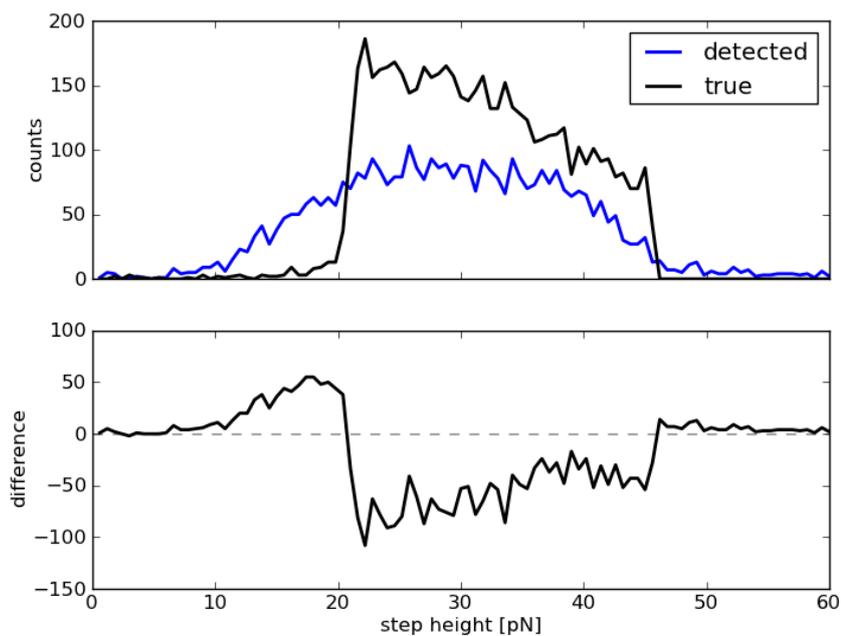


Fig. 29: True (black) and calculated (blue) step heights obtained from data set B by linear fits (a) and by the χ^2 method (b). The former does not reproduce low steps, the latter underestimates all heights.

3.2.5. Computational cost

Detection of steps in data with $N = 8192$ samples using a C++ implementation of the MSF algorithm with $w = 100$ requires a computation time Δt of the order of a millisecond on a current personal computer. This allows for automated processing of large data sets. Δt rises linearly with N . The χ^2 method is about 7000 times slower for the same number of samples, to some extent because it is based on more complex calculations. It performs linear fits over comparatively large intervals, partially including the same data points repeatedly⁶⁸. The relative difference in computation time increases with N ($\Delta t \sim N^{1.4}$).

3.2.6. Analysis of AFM force spectra

Force-distance curves were obtained by atomic force spectroscopy measurements with membrane tethers pulled from living human T lymphocytes. The adhesion force of single tubes formed by the cell membranes when interacting with the integrin VLA-4 ligand VCAM-1 was measured as described by Schmitz et al.⁵⁵. Bond rupture results in abrupt changes of the force exerted on the cantilever. As a consequence, discrete force states were recorded (see example in fig. 30a). The steps marking the transitions between these states were detected by the MSF algorithm with manually optimized parameters ($\sigma = 3.0$, $w = 200$) and a constant threshold for the significance of 10000 (blue vertical lines in fig. 30b and c; see section 3.1.2). Both MSF parameters are higher than those resulting from the optimization based on data set B to suppress oscillations contained in the force signals, which are not modeled by the artificial AFM noise. For comparison, the Kerssemakers algorithm was also applied to the example with $\lambda = 0$. If configured to detect the same number of steps, it does not identify the first one at $\approx 0.25 \mu\text{m}$ with the lowest significance (red lines in fig. 30c). However, the fit indicates that it is correct.

The resulting step heights found within a maximum pulling extension of $1.5 \mu\text{m}$ show a very symmetric distribution (fig. 31). The modal of about 23 pN corresponds well to previous AFM measurements of the same cell type under comparable conditions³⁹. As proven by the evaluation of simulated data, the MSF algorithm allows recovering steps, which are below the detection limit for conventional techniques (≈ 10 pN). In fact, the unimodal symmetric shape of the histogram provides strong evidence that it is not substantially distorted by false-positives. An automated analysis of these data with the Kerssemakers method is not possible, because the number of steps must be specified manually for each curve.

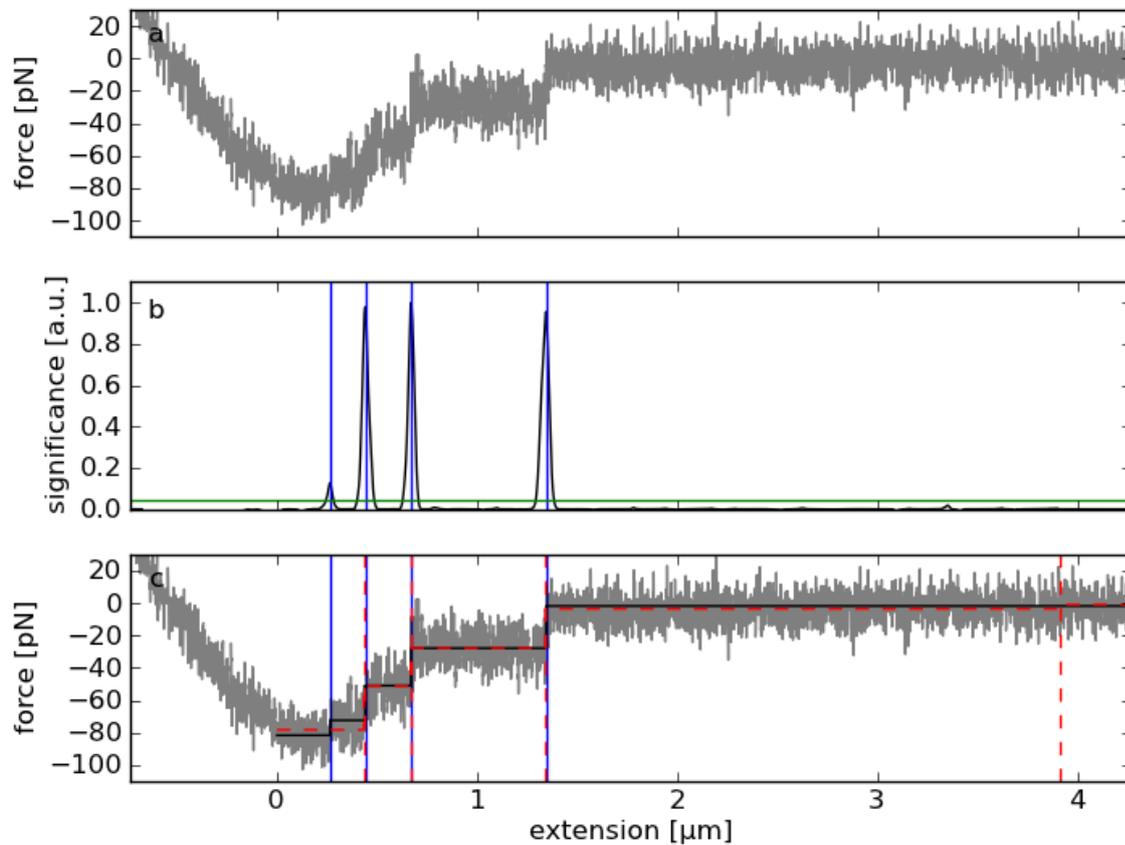


Fig. 30: a: Force-distance curve measured by atomic force spectroscopy of the interaction of human T lymphocytes with the integrin VLA-4 ligand VCAM-1. The curve shows the typical signature of membrane tether formation. b: Local maxima of the calculated significance exceeding a threshold (green horizontal line) indicate the rupture of the tethers (blue vertical lines). c: Fitting constant plateaus piecewise to the sections between the steps yields a clean force-distance curve (black). The Kerssemakers method identifies different step positions (red).

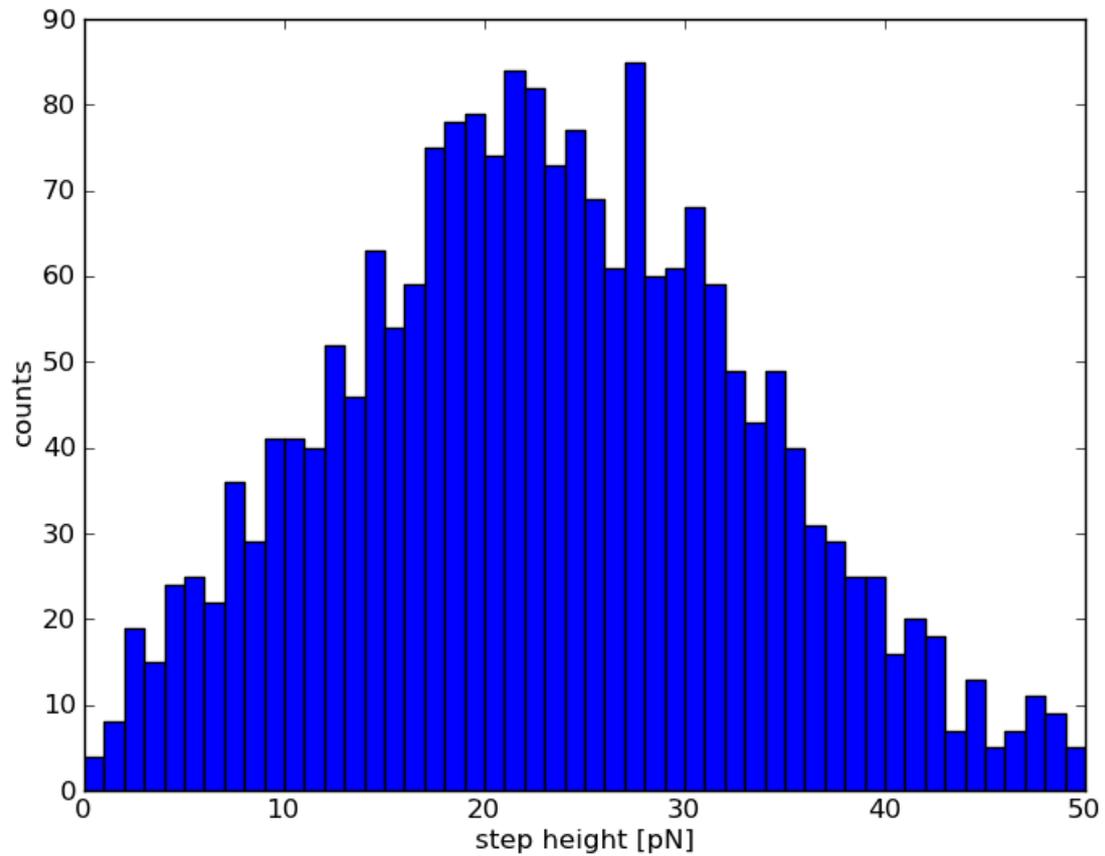


Fig. 31: Distribution of the calculated step heights resulting from the analysis of about 4200 force curves by MSF.

3.2.7. Analysis of kinesin motor experiments

Molecular motors constitute another example for a biological system showing discrete states. Both methods were applied to resolve the step-like movement of kinesin-2 along microtubules⁷⁵ with manually optimized parameters (MSF: $\sigma = 100$, $w = 500$, indicator threshold = 22000; Kerssemakers: $\lambda = 0$, 20 steps). They show similar results, but the Kerssemakers algorithm does not detect the two potential steps at ≈ 2.9 s, which correspond to the two lowest maxima in the MSF indicator (fig. 32). However, the rising flank suggests that at least one of them is actually correct.

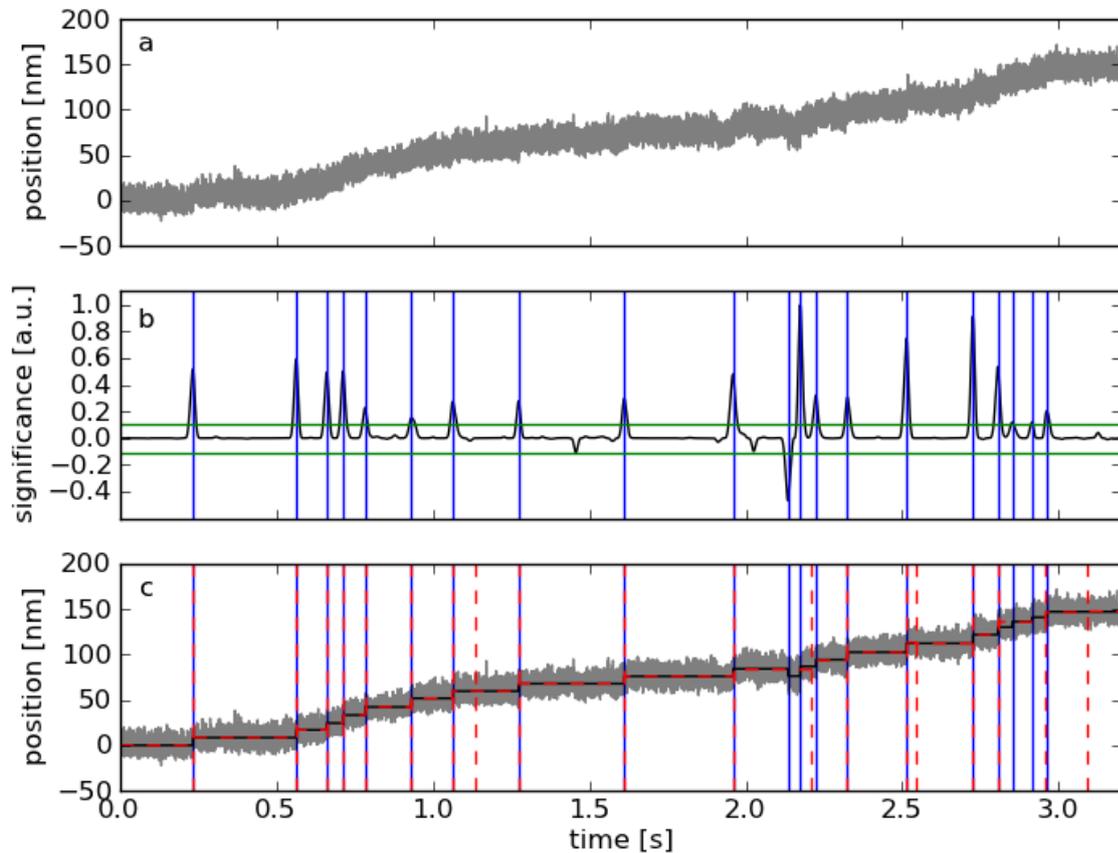


Fig. 32: Application of the MSF method to kinesin motor data. a: Distance vs. time trace obtained by optical tweezers. A polystyrene bead decorated with kinesin-2 proteins is held in an optical trap at a constant pretension of 1.4 pN, while one of the motor proteins moves along a surface-attached microtubule. b: The MSF indicator is thresholded (green) to locate the positions of the motor steps (blue). c: The mean values between these step positions are used to reconstruct the movement of the bead (black). The Kerssemakers method shows similar results, but does not detect the potential steps at ≈ 2.9 s, which correspond to the two lowest maxima in the MSF indicator.

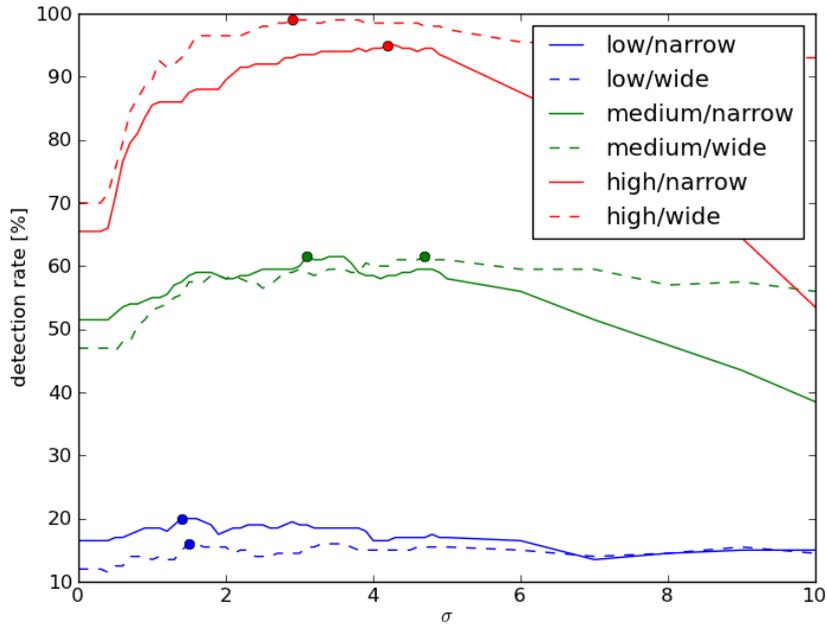
3.2.8. Parameter optimization

Independently of a particular detection algorithm, steps with arbitrarily low SNRs can be recognized if they are wide enough, and narrow steps if their SNR is high enough. No simple relation exists between the method-specific parameters and the lateral or height resolution. Optimal values in terms of efficient detection of steps depend on their widths and heights (fig. 33).

However, for sufficiently wide and high steps, MSF yields good results for $\sigma = 2$ and $w = 30$, and the χ^2 method for $\lambda = 4$. Whereas varying λ has only a moderate effect on the results, optimizing the width of the Gaussian kernel σ and the half window size w can highly improve or impair the detection efficiency, particularly for narrow steps. Ideal values of w in the sense of maximum detection rates are approximately inversely proportional to the SNR for large and intermediate heights ($w_{\text{opt}} \approx 50/\text{SNR}$). The optimal σ depends on w and on the step characteristics in a complex way. As a rule of thumb, a higher w and a lower σ increases the height resolution, and therefore increases detection rates for low steps. If steps are lying too close together to be separated, w must be decreased. If false-positives appear within the flanks of the indicator peaks of correctly identified steps, σ must be increased. Thereby, minor peaks with low prominences are eliminated.

The parameters can also be fine-tuned by comparing actual and detected steps in simulated signals mimicking the characteristics of real experimental data (such as noise amplitude or step heights and widths). By systematically varying the detection parameters, optimal values can be identified. If the signal characteristics are not constant, parameters resulting in maximum average detection rates can be determined by evaluating multiple test signals. This approach is illustrated using the example of single molecule force spectroscopy data obtained with living lymphocytes: The number of steps and their positions are determined by random statistic processes, which can be mimicked by Monte Carlo simulations. To create a realistic reproduction, the model parameters must be fitted to the experimental data⁷². In the next step, the model is used to create a set of artificial curves with known steps and random noise (in this example data set B). By comparing actual and detected step positions for a large batch of simulated data, optimal parameters for the simulated data can be identified. A similar procedure can be performed with data set A. The optimized parameters are only valid for the chosen model. An inapplicable model results in sub-optimal parameters.

a



b

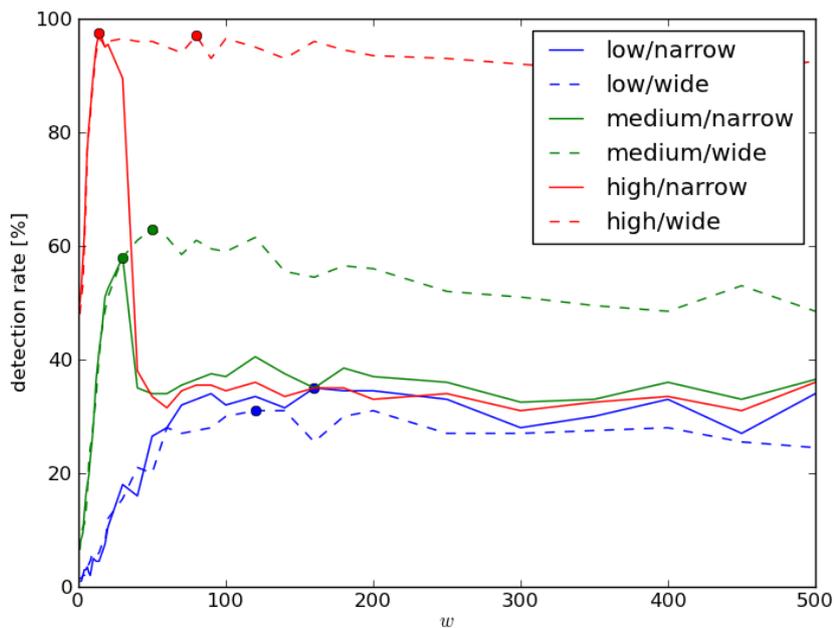


Fig. 33: Influence of the SNRs and widths of steps on the detection rates and optimal parameters for (a) the size of the Gaussian kernel σ and (b) the half width of the fit window w . Only one parameter was varied at a time, the other was held constant (either $w = 30$ or $\sigma = 2$). Each point represents the detection rate averaged over 100 curves of data set A with a SNR of 0.5, 1.0, or 2.0, and a step distance of 40 or 1000 data points. Optimal values for the varied parameter are marked by the circles.

4. Effect of SDF-1 α on the mechanics of the integrin $\alpha_4\beta_1$ environment

T lymphocytes are important components of the adaptive, cell-mediated immune response. They circulate in the blood stream until they detect a site of inflammation. There they strongly adhere against the shear force of the blood stream and extravasate into the tissue. The adhesion is mediated by the family of integrins. For the lymphocyte integrin LFA-1, conformational inside-out activation has been shown to be triggered by chemokines^{19, 76}. In contrast, the affinity of the integrin VLA-4 to its ligand VCAM-1 is not affected by chemokine stimulation alone: No activation epitopes were detectable after exposure of cells to chemokines, and VLA-4 affinity to soluble VCAM-1 was not altered^{77, 78}. However, despite the lack of detectable affinity increase, the chemokine SDF-1 α , the ligand of the G-protein-coupled receptor CXCR4, is a highly potent VLA-4 stimulatory chemokine³¹⁻³³. Yet, even SDF-1 α and other chemokines failed to induce firm T lymphocyte adhesion in the absence of external forces³⁴. Therefore, an additional mechanical stimulus seems to be necessary. To test this hypothesis, the effects of SDF-1 α on the properties of the VLA-4/VCAM-1 bond were investigated by single-molecule atomic force microscopy (AFM) on living T lymphocytes. The results show that SDF-1 α increases the strength of the interaction of an individual VLA-4/VCAM-1 binding site. The increase in strength is associated with a stiffening of the micro-environment of the integrins, indicating a supportive role of the cytoskeleton. Deleting the binding site to the talin head group in the VLA-4 cytoplasmic tail suppresses lymphocyte binding in the absence of SDF. In the presence of SDF, no effect of the deletion can be observed. Hence, the binding of the talin head group to VLA-4 is not required for SDF-induced adhesion strengthening.

4.1. Materials and methods

4.1.1. Reagents

BSA (fraction V), HSA (fraction V) and HBSS (without calcium/magnesium) were purchased from Sigma-Aldrich. Recombinant human VCAM-1 (CD106), SDF-1 α (CXCL12) and anti-human CD43 mAb were purchased from R&D Systems. HP1/2 was a gift from Francisco Sánchez-Madrid (Hospital Universitario de la Princesa, Madrid, Spain).

4.1.2. Lymphocytes

Transfectants of the β_1 integrin-deficient Jurkat T-cell line A1 were used, which lacks expression of

the β_1 integrin subunit^{54, 79}. Wild-type β_1 integrin and β_1 constructs with mutations targeting the membrane-proximal NPIY motifs were re-substituted (table 8). All cells were kindly provided by M. A. Rosenthal-Allieri, Department of Immunology, Archet Hospital CHU de Nice, and cultured in VLE RPMI 1640 (Biochrom) supplemented with 10% FCS (Biochrom), 10 mM HEPES (Biochrom), 2 mM L-glutamine (Biochrom), 1 mM sodium pyruvate (Gibco) and 4.5 g/l D-Glucose (Sigma) in 5% CO₂ at 37 °C. The cells were transferred to binding medium (HBSS with 2 mg/ml BSA, 1 mM CaCl₂, 1 mM MgCl₂, 10 mM HEPES) immediately before the experiments.

4.1.3. Substrate preparation

2 μ g/ml VCAM-1 (corresponding to a surface density of 4000 sites/ μ m²) was incubated over night at 4 °C on the lid of a polystyrene petri dish together with the carrier protein HSA (2 μ g/ml). The spots were washed four times with PBS and either blocked with 2% HSA in PBS for > 1 h at 4 °C or incubated with SDF-1 α (2 μ g/ml in PBS) for 3 h at 4°C. After SDF adsorption, the spots were washed four times with PBS and blocked with 2% HSA in PBS for > 1 h at 4 °C.

4.1.4. AFM measurements

Tipless silicon cantilevers (Arrow TL2, Nanoworld, Neuchâtel, Switzerland) were used. Spring constants were determined by thermal noise analysis. The cantilever was functionalized with 0.1 mg/ml anti-CD43 mAb for 30 minutes at room temperature. The force spectroscopy experiments were conducted at (36 \pm 1) °C in binding medium with a Nanowizard II AFM (JPK, Berlin, Germany). Immediately before the experiment, a single lymphocyte was immobilized at the cantilever. The piezo was extended and retracted periodically with a velocity of 3.4 μ m/s. A maximum indentation force of 100 - 150 pN was applied for 100 ms. Typically, 200 force-distance curves were recorded per cell (see example in fig. 34). For control, the same experiments were performed on VCAM-1 substrates after blocking with HP1/2 at a concentration of about 2 mg/ml. Further, measurements of cells were conducted that were treated with 100 ng/ml pertussis toxin (PTX) for 15 hours.

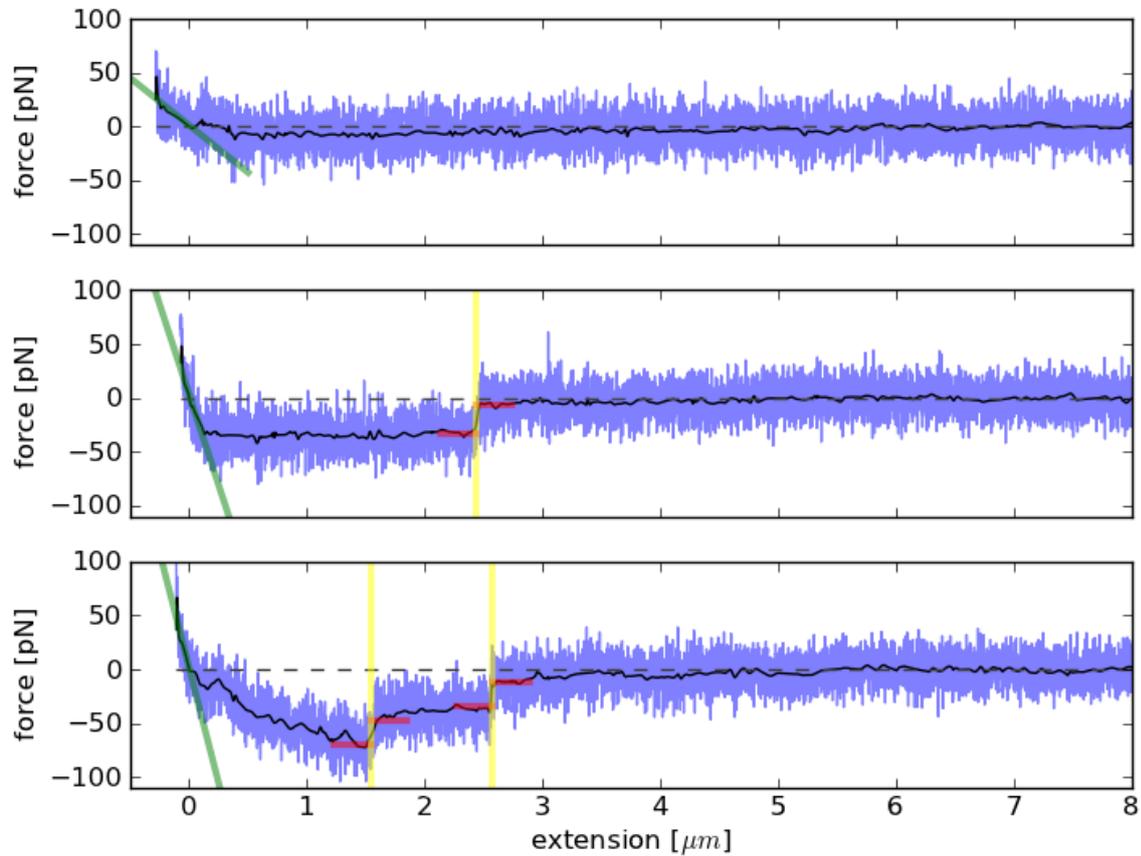


Fig. 34: Three example retrace curves. Raw force curves are marked in blue, de-noised curves in black, step positions in yellow, step heights in red, and the indentation slope (i.e. the slope of the retraction curve at the point where the force exerted on the cantilever is zero) in green.

4.1.5. Data analysis

To allow for a completely objective evaluation, recorded data were analyzed automatically using custom-designed software written in Python and C++ (see sections 7.1 and 7.2). Instrumental drift was corrected by baseline subtraction. The ReNoiR algorithm was deployed for noise reduction (see section 2.1.2). The zero point of the measured distance (contact point) was determined by intersecting the baseline with the indentation part of the force curves. Subsequently, the positions and heights of steps in the force-distance curves were detected by the MSF algorithm (see section 3.1.2).

4.1.6. FACS analysis

Immunocytochemical staining of CD49d, CD29, and CXCR4 was conducted on viable cells using FITC-conjugated monoclonal antibodies as previously outlined^{80, 81}. After washing with 1% BSA, the cells were incubated with Fc block (Dianova 009-000-008) for 30 min at room temperature and with the antibodies (CD49d: Southern Biotech 9431-02, CD29: EXBIO 1F-219-T025, CXCR4: R & D Systems FAB170F) for 45 min at 4 °C.

construct	sequence
A ₁ β ₁	KLLMIIHDRREFAKFEKEKMNAKWDTGE NPIY KSAVTTVV NPKY EGK
A ₁ β ₁ -ΔNPIY	KLLMIIHDRREFAKFEKEKMNAKWDTGE ---- KSAVTTVV NPKY EGK

Table 8: Integrin constructs

4.2. Results

Three different cell types were compared: A1 lymphocytes not expressing β₁ integrins, A1β₁ lymphocytes expressing wild-type β₁ integrin, and A1β₁-ΔNPIY lymphocytes expressing a β₁ integrin with a deleted NPIY motif. The interaction of these cells with surfaces bearing the VLA-4 (α₄β₁) ligand VCAM-1 was analyzed. Since endothelial chemokines are physiologically presented in juxtaposition to integrin ligands³², the chemokine SDF-1α was co-immobilized with VCAM-1 for comparison with both substrates bearing VCAM-1 alone (VCAM-1 density was kept constant) or bearing SDF-1α alone.

4.2.1. VLA-4 expression

Using FACS analysis, A1β₁ and A1β₁-ΔNPIY lymphocytes were found to express very similar levels of VLA-4 (fig. 35), whereas the β₁ integrin-deficient A1 lymphocytes only showed expression of CD49d, the α₄ subunit of VLA-4. All cell types expressed CXCR4, the receptor for SDF-1. Staining procedures in the absence of any antibody did result in significantly lower fluorescence signals.

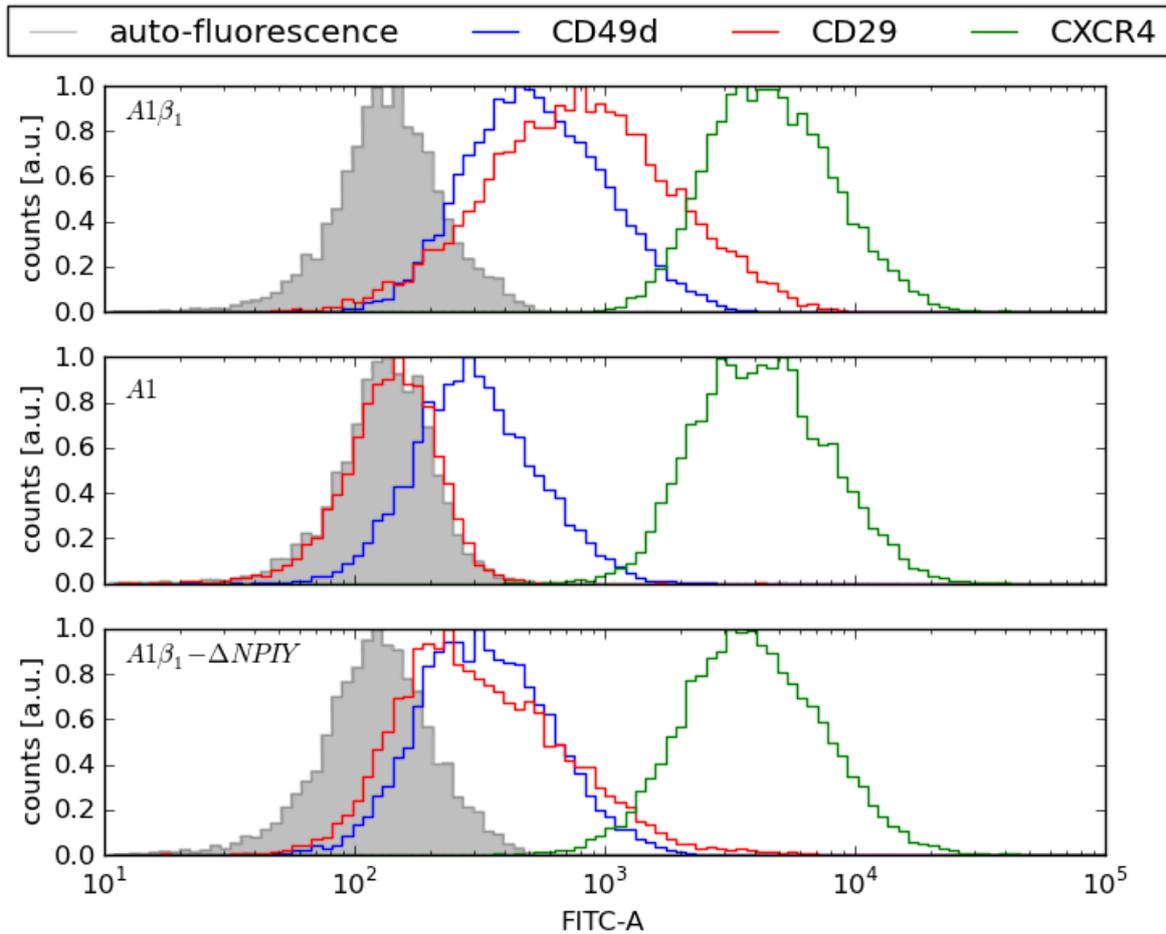


Fig. 35: FACS emission spectra of cells labeled with antibodies against CD49d (α_4 subunit of VLA-4), CD29 (β_1 subunit of VLA-4), and CXCR4 (SDF1-specific receptor). Counts are normed to the modals of the individual distributions.

4.2.2. Adhesion rates

4.2.2.1. Multiple adhesions

Binding of the cells to the functionalized surfaces leads to steps in the force-distance curves. Each force step represents the breaking of an individual cell-to-surface adhesive site. How many integrin-ligand interactions take part in this individual binding site is unknown. The average number of observed steps for $A1\beta_1$ lymphocytes encountered with VCAM-functionalized surfaces depends on the VCAM-1 coating concentration, indicating cooperative binding of integrins at higher ligand densities.

4.2.2.2. Individual binding sites

To investigate the modulation of the properties of individual cellular binding sites by the chemokine SDF-1 α , a VCAM coating concentration of 2.0 $\mu\text{g/ml}$ was used for further experiments. This yields adhesion rates of less than 30% and ensures a high probability of individual bindings. At this coating density, the histograms of adhesions of the lymphocytes fit to the Poisson distribution $f(n) = \lambda^n e^{-\lambda} / n!$, as expected for individual, mutually independent bindings (fig. 36). If no SDF was present, A1 lymphocytes not expressing the integrin $\alpha_4\beta_1$ showed an adhesion rate of (8.2 ± 1.3) % on VCAM-coated surfaces (fig. 37). With A1 β_1 lymphocytes that do express $\alpha_4\beta_1$, a rate of (29.1 ± 4.6) % was measured. Adhesion of A1 β_1 lymphocytes was reduced by the VLA-4-specific monoclonal antibody HP1/2 to a level of (7.4 ± 0.2) %. SDF-1 α lowered the adhesion rate of A1 β_1 lymphocytes. When co-immobilized with VCAM-1, the rate was (19.1 ± 4.8) %. A mutation in the membrane-proximal NPIY motif of VLA-4 reduced cellular adhesion to an unspecific level of (10.4 ± 1.3) % in absence of SDF. Surprisingly, SDF-1 α rescued the effect of the mutation: A1 β_1 - Δ NPIY lymphocytes showed an adhesion rate of (28.2 ± 3.6) %, the same as do lymphocytes bearing the wild-type integrin when SDF-1 α is juxtaposed to VCAM-1. SDF-1 α alone did not induce specific adhesion: Surfaces coated with SDF-1 α , but without VCAM, resulted in an adhesion rate of (11.3 ± 2.8) % for A1 β_1 and (10.6 ± 3.6) % for A1 β_1 - Δ NPIY lymphocytes, i.e. interactions were indistinguishable from unspecific adhesion.

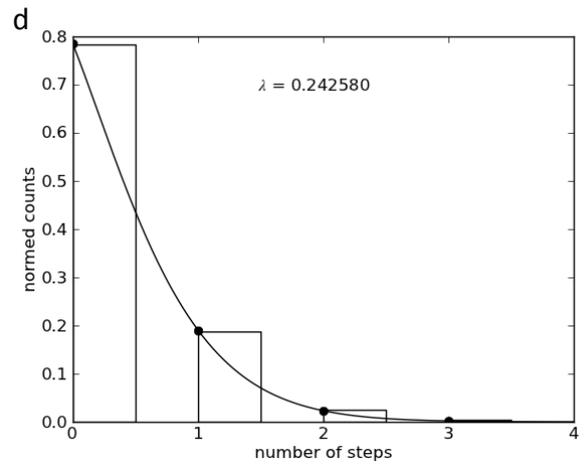
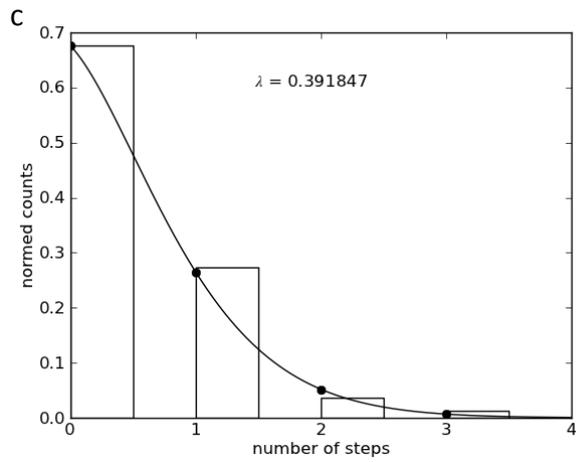
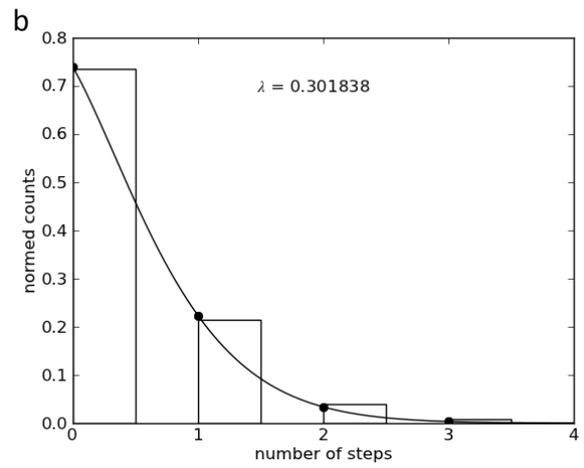
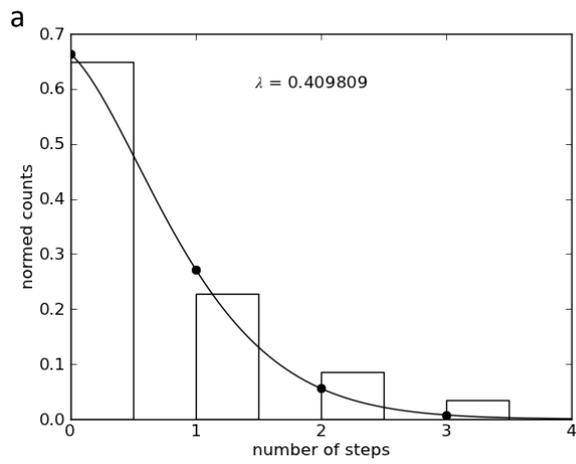


Figure 36: The numbers of steps follow Poisson distributions. a: $A1\beta_1 - SDF$, b: $A1\beta_1 + SDF$, c: $A1\beta_1 - \Delta NPIY + SDF$, d: PTX

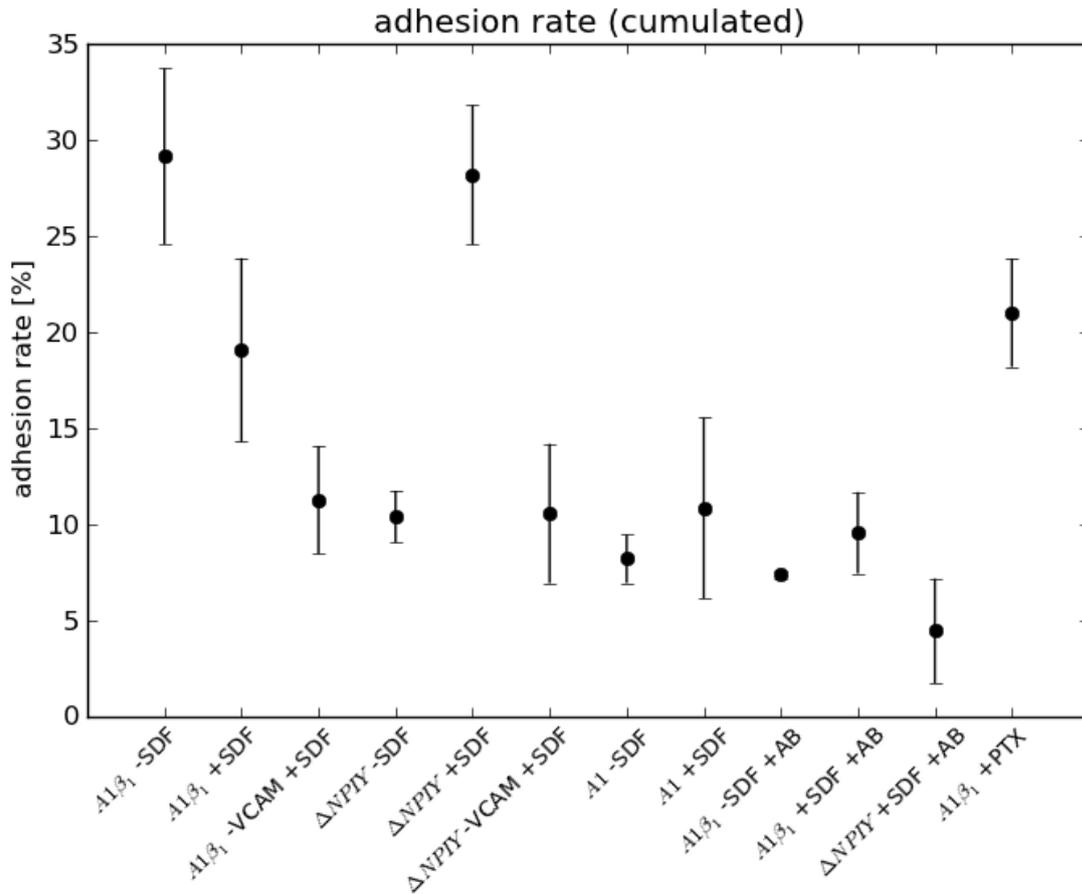


Fig. 37: Adhesion rates (i.e. the ratio of the numbers of adhesive and all curves)

4.2.3. Forces

Unbinding forces were evaluated only under the experimental conditions for which specific interactions had been observed. SDF-1 α led to a small, but significant increase in the step height of individual cell-to-surface binding sites from (21.0 ± 0.3) to (26.0 ± 0.3) pN (fig. 38). When pre-incubated with pertussis toxin (PTX), the adhesion strengthening was inhibited. The median step height of (19.7 ± 0.4) pN is comparable to untreated cells encountering VCAM-1 alone.

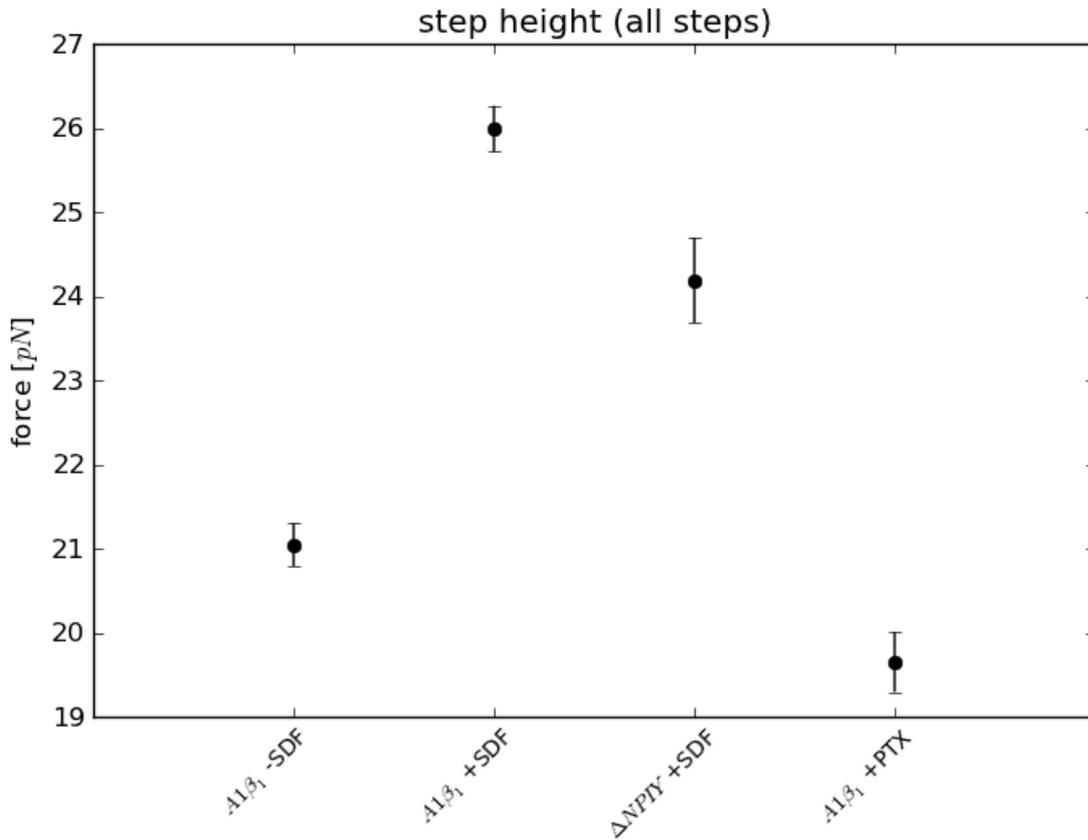


Fig. 38: Medians of the heights of the detected steps (depicted by the red lines in fig. 34)

4.2.4. Step positions

Step positions were also evaluated only for experimental conditions that resulted in significant specific adhesions (see fig. 37). They were determined relative to the contact point (see section 4.1.5). The median of the step positions of the A1β₁ cells interacting with a VCAM-functionalized surface was $(0.76 \pm 0.03) \mu\text{m}$ (fig. 39). SDF induced faster bond breakage, leading to a median step position of $(0.35 \pm 0.02) \mu\text{m}$. Since β₁ integrins with the NPIY deletion showed no significant interactions in the absence of SDF-1α, the step positions of the A1β₁-ΔNPIY cells were evaluated only in experiments, in which SDF-1α was co-immobilized with VCAM-1. Under these conditions, the step positions were comparable to those of A1β₁ cells interacting with a surface coated with both VCAM-1 and SDF-1α, and hence shorter than the step positions of A1β₁ cells interacting with surfaces coated with VCAM-1 alone. PTX acting on A1β₁ cells did shorten the steps even further.

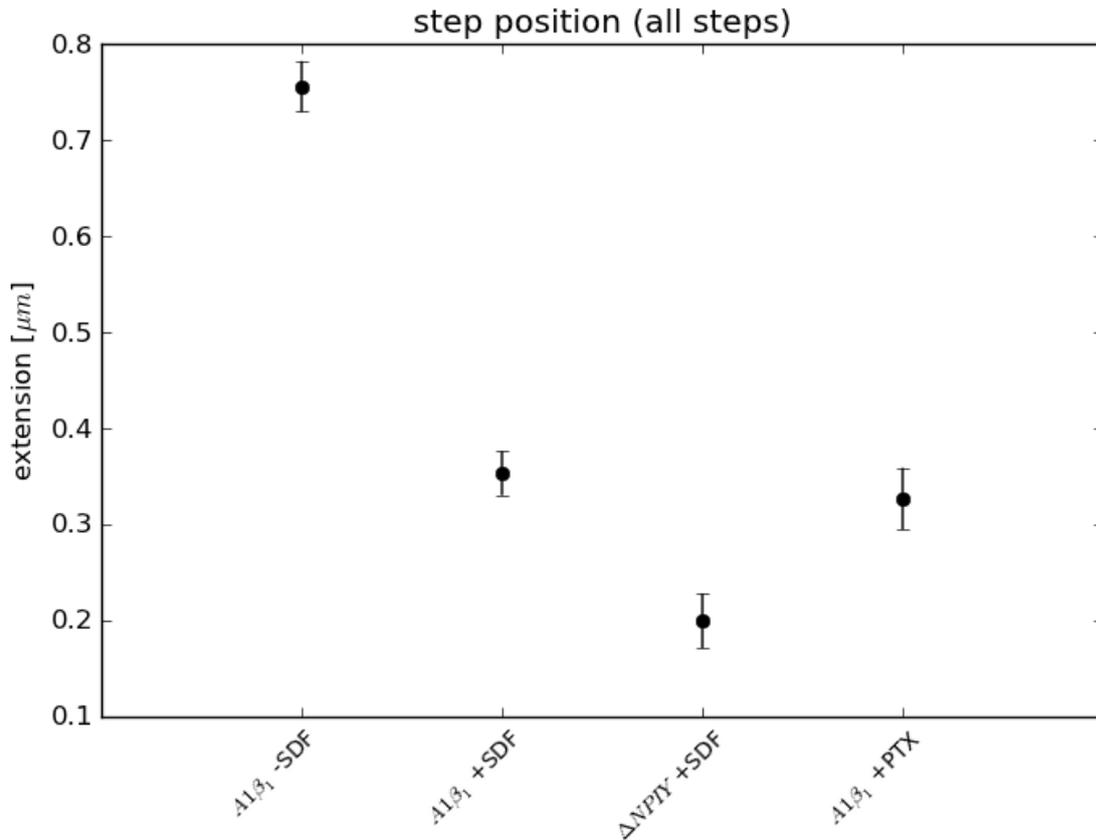


Fig. 39: Medians of the positions of the detected steps (depicted by the yellow lines in fig. 34)

4.2.5. Lymphocyte stiffness

After approaching a lymphocyte to the VCAM-1 substrate, it was further indented until a setpoint of 100 - 150 pN was reached. As the surface is much harder than the lymphocyte, the slope of a force-distance curve in the regime of positive indentation represents a measure for the stiffness of the lymphocyte, independently of possible subsequent lymphocyte adhesion. The slope of the force-distance retrace curve was evaluated within a relative cantilever-surface distance of 20 nm from the contact point. In this regime, the force-distance curves are approximately linear. A1 β ₁ lymphocytes opposed to VCAM-1 showed an indentation slope of $-(127 \pm 2)$ pN/ μm (fig. 40). When the cells encountered SDF-1 α juxtaposed to VCAM-1, a strong stiffening was observed with a slope of $-(170 \pm 4)$ pN/ μm . Stiffening was inhibited by pertussis toxin (PTX), a toxin that inhibits G-protein-coupled signaling⁷⁶. Lymphocytes facing SDF-1 α alone showed a slope of $-(120 \pm 4)$ pN/ μm ; they did not alter their mechanical properties compared to lymphocytes opposed to VCAM-1 only.

The effect was most pronounced in lymphocytes expressing the NPIY mutated VLA-4. In the presence of VCAM-1 only, NPIY lymphocytes were softer than cells bearing wild-type integrin with a slope of $-(84 \pm 4)$ pN/ μm . When encountering SDF-1 α juxtaposed to VCAM-1, they were stiffer than the lymphocytes expressing wild-type VLA-4 with a slope of $-(185 \pm 7)$ pN/ μm . SDF-1 α juxtaposed to VCAM-1 had no stiffening effect on A1 lymphocytes not expressing VLA-4. The stiffening was only observed when cells encountered both VCAM-1 and SDF-1. A surface coated only with SDF-1 did not induce any cell stiffening.

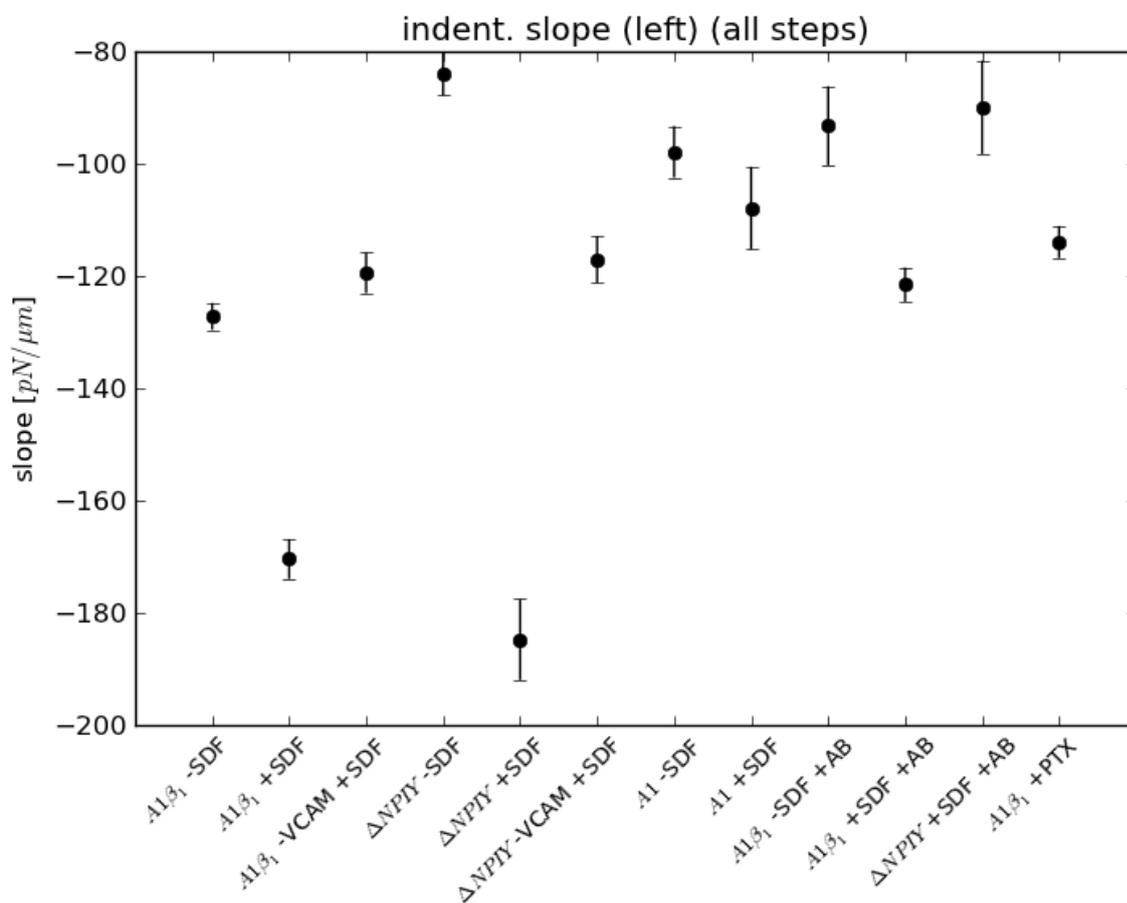


Fig. 40: Slopes of the retrace curves at the contact points (depicted by the green lines in fig. 34)

4.3. Discussion

At least two integrin subtypes expressed on lymphocytes are strongly affected by SDF-1 α : LFA-1 and VLA-4³¹. For LFA-1, it was demonstrated that SDF-1 α leads to conformational changes in the integrin's ecto-domain exposing neo-epitopes and increasing the affinity of LFA-1 to its ligand, ICAM^{19, 76}. The effect of SDF-1 α on the VLA-4/VCAM-1 interaction, however, is not understood to the same extent. While early on it has been shown that SDF-1 α is a strong modulator of VLA-4-mediated lymphocyte adhesion and migration^{31-33, 82, 83}, structural changes or affinity modulation caused by soluble SDF-1 α were not observable in cellular assays⁷⁷. However, immobilized SDF-1 α did induce a high-affinity $\alpha_4\beta_1$ conformation in parallel flow chamber assays^{33, 84}. For a deeper understanding of the effect of SDF-1 α on $\alpha_4\beta_1$, it is crucial to analyze the effects of SDF-1 α on the VLA-4/VCAM-1 interaction down to the molecular level. Analyzing the physiological activation of VLA-4 by the chemo-attractant SDF-1 α on the level of individual adhesion sites is a daunting task. First, it requires living lymphocytes with intact signaling pathways. Second, previous studies suggested that force is a mandatory co-signal for this activation^{20, 34, 84-86}. Hence, a force-based analytical technique is asked for that is capable of working with living lymphocytes and at the same time sensitive enough to measure the interaction forces of individual adhesion sites in the low pN range. As shown in the following, single-cell atomic force spectroscopy is perfectly suited for this task.

In this study, the interaction of lymphocyte-expressed VLA-4 with surface-bound VCAM-1 and its modulation by the chemokine SDF-1 α was measured down to the level of single adhesion sites using AFM. Each step in the force-distance curves signifies the breakage of such a single adhesion site, which may either be a single integrin or a cluster of integrins.

At adhesion rates below 30%, the histogram of the number of simultaneously observed adhesive interactions of the lymphocytes with the surface follows a Poisson distribution, which arises from a memoryless stochastic process counting the number of random, independent events in a given interval⁸⁷. Therefore, the histogram argues that the observed binding sites are linearly independent at low ligand densities. In order to scrutinize individual, independent binding sites, the effect of SDF on the binding properties of the lymphocytes to VCAM-1 covered surfaces was analyzed only at this low ligand density. However, this does not prove that in fact interactions of single molecules are measured. These individual binding sites may be composed of a small number

of integrins acting cooperatively.

The ability to block the measured adhesions by monoclonal antibodies as well as the lack of adhesion above a level of 10% of lymphocytes not bearing VLA-4 or of lymphocytes approaching surfaces without VCAM-1 demonstrates the VLA-4/VCAM-1 specificity of the here measured adhesions.

As shown above, SDF-1 α strengthens the VLA-4-mediated adhesions down to the level of individual binding sites. Treating the lymphocytes with PTX, a drug known to inhibit G-protein signaling⁷⁶, can reverse the effect. Hence, the SDF-1 α -triggered effects are caused by G-protein signaling.

On the here evaluated molecular level, no significant increase in the adhesion rate caused by SDF-1 α was observed. Since the interaction is not ligand-limited, this argues that SDF-1 α does not prominently activate previously inactivated VLA-4 on the cellular surface. Previous studies have shown that no neo-epitopes indicative of an activated VLA-4 state have been found on lymphocytes after exposure to SDF-1 α ^{77, 78, 88}. In addition, no effect of SDF on the affinity of VLA-4 to soluble VCAM-1 was detected^{69, 70}. Further, it was shown that strong VLA-4-mediated lymphocyte adhesion on surfaces bearing SDF-1 α is only achieved if the lymphocytes are exposed to shear stress^{20, 34, 84}. Hence, SDF-1 α alone in a force-free environment does not seem to induce a high-affinity VLA-4 state before ligand binding. This is in accord with the results presented here: SDF-1 α does not significantly increase the adhesion rate of the lymphocytes. It does, however, strengthen the adhesions. This indicates that the effect of SDF-1 α is to facilitate force-induced changes after ligand binding. The importance of ligand binding is further stressed by the fact that SDF-1 α only causes cellular stiffening when juxtaposed to VCAM-1. Hence, occupancy of both CXCR4 and $\alpha_4\beta_1$ by the respective ligands is important for the effect of SDF-1 α on the cell mechanics.

It has been shown previously that a swing-out of the integrin leg domain – a prerequisite for a high-affinity conformation – is strongly force-assisted: Small forces applied to integrin-bound ligands speed up this swing-out dramatically^{89, 90}. Efficient force transduction requires a stiff environment of the integrin $\alpha_4\beta_1$. Therefore, the ligand was bound to a polystyrene surface, i.e. a force-bearing, rigid structure. VLA-4, by contrast, is enriched at microvilli, which are easily extensible over many micrometers and therefore not well suited to bear and transmit forces.

However, previous experiments suggest that a force-mediated polymerization and hardening of the actin cytoskeleton occurs^{91, 92}. Further, the AFM measurements clearly show that SDF-1 α stiffens the cells, when juxtaposed to VLA-4. The ruptures of the VLA-4/VCAM-1 bond occur earlier and at higher forces under the influence of SDF-1 α – another indication of a rigidification of the integrin's environment. Early on, it has been demonstrated that SDF-1 α leads to a rapid and strong increase in F-actin^{82, 93-97}. SDF-1 α may therefore facilitate post-ligand binding force-assisted conformational changes of VLA-4 by stiffening the integrin's micro-environment.

The main adaptor protein attaching integrins to the cytoskeleton is talin^{16, 98-100}. Therefore, also the effect of SDF-1 α on lymphocytes was tested that bear a mutated integrin, Δ NPIY. The mutation abolishes binding of the talin head group to the integrin^{54, 101, 102}. Talin head group binding is thought to be essential for integrin activation^{29, 86}. Hence, a strong inhibition of cell attachment has to be expected by this mutation. Indeed, the mutation inhibited binding of the un-stimulated resting lymphocytes to surfaces coated with VCAM-1 alone despite strong integrin expression on the cellular surface, demonstrating the need of talin head group binding for basal integrin activation. Surprisingly, SDF-1 α rescued the effect of the mutation. When interacting with a surface bearing SDF-1 α in close proximity to VLA-4, the Δ NPIY lymphocytes were nearly indistinguishable from lymphocytes bearing wild-type VLA-4. This indicates that binding of the talin head group is not essential for VLA-4 activation by SDF-1 α . Earlier studies, though, pointed at an important involvement of talin in SDF-1 α -triggered VLA-4 activation: The knock-down of talin in lymphocytes abolished any effects of SDF-1 α on VLA-4-/VCAM-1-mediated adhesion of Jurkat lymphocytes⁷⁸. In contrast to this study, recent results could not observe any effect of talin-1 on SDF-1 α -induced $\alpha_4\beta_1$ affinity up-regulation in the U937 monocytic cells¹⁰³, which is in line with the presented results. In this study, however, talin is not knocked down. The examined mutation does only disturb the binding of the talin head group to the integrin β_1 tail. Recently, it has been shown that not the talin head group, but a second integrin binding site on the talin rod domain is crucial for the attachment of integrins to the intracellular adhesion complex^{104, 105}. This second binding site is independent of the NPIY motif. Attaching $\alpha_4\beta_1$ to the intracellular adhesion complex via talin's second binding site would cause a stiffening of the integrins' micro-environment, in agreement with the results presented here. The exact location of the interaction of VLA-4 with the second integrin binding site of talin is unknown to date, but suggested to be membrane-proximal of the NPIY motif¹⁰⁴. This region is still intact in the mutant cells.

In summary, results show that SDF-1 α signaling leads to a strengthened VLA-4-mediated lymphocyte adhesion independent of the NPIY motif, the binding site of the talin head group. The observed stiffening and the shorter ruptures demonstrate a rigidified mechanical cellular environment of VLA-4 after SDF-1 α signaling. This indicates an attachment of the integrin to the intracellular adhesion complex. The attachment may be caused by talin binding VLA-4 with its second, NPIY-independent binding site, linking it to F-actin, which is strongly and rapidly increased after SDF-1 α signaling. This then facilitates a post-ligand binding, force-assisted conformational change leading to a high-affinity VLA-4 conformation. Further studies will be aimed at identifying residues on the integrin VLA-4 that inhibit SDF-1 α signaling.

5. Role of the trans-membrane domains in integrin-mediated adhesion and signaling

Integrins adopt different conformational states to transmit signals bidirectionally across cell membranes; many of these are transient. The capture of transient states within experimental approaches is a major challenge in protein science. Therefore, model systems freezing intermediates by manipulating the protein sequence are often used to gain mechanistic insights into such dynamic protein systems. In this study¹⁰⁶, which is part of a project cooperation, we were interested in the functional consequences of different conformations of the integrin trans-membrane domain (TMD). The structure of the helical TMD dimer in the resting state of the integrin $\alpha_{IIB}\beta_3$ has recently been solved by NMR spectroscopy and is in excellent agreement with earlier computational studies¹⁰⁷⁻¹⁰⁹. In the fully activated state, the TMDs are separated^{110, 111}. Intermediate states like that after inside-out, but prior to outside-in signaling are so far ill-characterized. Here, we aimed at creating integrin mutants functionally representing these intermediate states. Computational studies showed that a GpA-like conformation is energetically favored for integrin TMD sequences. This conformation has further been postulated as an intermediate in forming focal adhesion¹¹². Therefore, we mutated the integrin $\alpha_v\beta_3$ TMD sequence to a GpA TMD sequence, thus enforcing a very stable GpA TMD conformation¹¹³ in the context of an $\alpha_v\beta_3$ integrin. In a second $\alpha_v\beta_3$ /GpA chimera we mutated the central GxxxG dimerization motif to GxxxI (TMD-GpA-I), a mutation known to abrogate TMD dimerization. This serves as a model for an $\alpha_v\beta_3$ integrin with constitutively dissociated TMD¹¹⁴. After expression of these constructs in a cellular system, we analyzed the effects of the two chimera with respect to cell adhesive strength to vitronectin (VN), the major ligand of $\alpha_v\beta_3$.

We found that expression of both TMD-GpA and $\alpha_v\beta_3$ -TMD-GpA-I provoked strong cell adhesion. These results correspond well to a three-state model of integrin activation, where a resting state is activated by intracellular ligands without TMD separation. The activation leads to an increase in adhesive strength. TMD-GpA mimics this intermediate state. Subsequent binding of ECM ligands then, in turn, provokes TMD separation and full integrin activation. In this state, integrin-triggered intracellular signaling events occur¹¹⁵ as well as linkage to cytoskeletal components. This enables cell migration²⁵. Hence, TMD-GpA-I mimics a constitutively activated integrin state.

Our results underline that integrin-mediated cell adhesion is decoupled from cell migration and

signaling. Both are influenced by the TMD sequences in an allosteric fashion.

5.1. Materials and methods

5.1.1. In vitro site-directed mutagenesis

Exchange of the α_v - and β_3 -TMD by the complete GpA-TMD was conducted by sequential in vitro site-directed mutagenesis as described earlier⁸⁰.

5.1.2. Cell culture

The origin, cultivation, and stable transfection of human ovarian OV-MZ-6 cancer cell clones has been described previously¹¹⁶.

5.1.3. Atomic force spectroscopy

All force measurements were conducted using a NanoWizard II atomic force microscope (JPK Instruments AG, Berlin, Germany) mounted on an inverted optical microscope (Axiovert 200, Zeiss). Tipless silicon cantilevers were used (Arrow TL2, NanoWorld, Neuchâtel, Switzerland). Cells in a low density were allowed to attach to poly-D-lysine-coated cover slips for single-cell measurements. After 24 h of incubation, cells were washed in PBS, followed by incubation in PBS, 1 mM CaCl₂, and 1 mM MgCl₂, at 37°C within a BioCell sample holder (JPK Instruments). VN was immobilized on silicon cantilever tips and the spring constants of the cantilevers were determined via thermal noise calibration at (16 ± 12) mN/m. During the experiments, the approach and retract velocity of the cantilever was maintained at 4 μ m/s. The cantilever was moved towards cell surfaces with an applied force of 0.5 nN and a dwell time of 100 ms. The approach/retract cycle was repeated at least 200 times per cell within 2 h with at least 10 cells measured per setup. The obtained force-distance curves were analyzed focusing on the work of de-adhesion, the peak force during de-adhesion, and the number, height, and position of distinct steps in the force-distance curves using the data analysis software described in sections 7.1 and 7.2.

5.2. Results

The force-distance retraction curves obtained by the AFM measurements showed steps typically

seen in this kind of experiments^{55, 117} (fig. 41). From these curves, we counted the average number of adhesive events and the number of steps per adhesive force curve (fig. 42). For cells expressing TMD-GpA, nearly 8 force steps per adhesive event were observed, compared to approximately 3 force steps for TMD- $\alpha_v\beta_3$ and TMD-GpA-I. Upon blockade of $\alpha_v\beta_3$ /VN-interactions by use of the mAb directed to $\alpha_v\beta_3$ # 23C6, both the number of force steps and the number of adhesive events were reduced. However, we still observed frequent unspecific and weak adhesion events. We further evaluated the peak force (fig. 43), the work needed to retract the cantilever, which equals the area under the force-distance curves (fig. 44), as well as height and position of the force steps (fig. 45 and 46). The median of the peak force increased from 91.2 pN for unspecific adhesion measured after blocking by the $\alpha_v\beta_3$ -directed mAb over 102.3 pN for TMD- $\alpha_v\beta_3$, 126.8 pN for TMD-GpA-I up to 188.8 pN for TMD-GpA (fig. 43). The work required for detaching the cantilever from the cells increased from 52.9 aJ (blocking by $\alpha_v\beta_3$ -directed mAb) over 125.8 aJ (TMD- $\alpha_v\beta_3$), 188.5 aJ (TMD-GpA-I) up to 287.1 aJ (TMD-GpA) (fig. 44). The low forces and low work after treatment with the $\alpha_v\beta_3$ -directed mAb demonstrated the importance of this integrin for strong cell adhesion and also the $\alpha_v\beta_3$ -specificity of the AFM measurements. We also observed a significant increase in the step height from 21.5 pN for TMD- $\alpha_v\beta_3$ to 27.0 pN for TMD-GpA and TMD-GpA-I, respectively (fig. 45). The strengthened adhesion correlated with a decrease in the median step position for TMD-GpA-I in comparison to TMD- $\alpha_v\beta_3$ from 0.9 to 0.5 μm , while TMD-GpA expression exerted no effect on the step position with a median position of 1.0 μm (fig. 46).

5.3. Discussion

The measurements documented increased adhesion of cells expressing TMD-GpA-I when compared to TMD- $\alpha_v\beta_3$. These data disclosed a robust increase in adhesive strength of cells displaying TMD-GpA when compared to TMD-GpA-I. Furthermore, TMD-GpA-expressing cells formed bonds more rapidly than TMD-GpA-I. When probed under external forces, TMD-GpA shows an increased force needed for braking single tethers compared to TMD- $\alpha_v\beta_3$, indicating that TMD-GpA/ligand bonds are better suited to withstand external forces (fig. 43). Since the ligand binding site itself is unaltered by the TMD mutation, the observed effect argues for an allosteric effect. The observed combination of an increased step height with a decreased step position of TMD-GpA-I (fig. 45 and 46) indicates its constitutive linkage to the cytoskeleton⁵⁵. This causes a stiffer integrin environment, which increases the rate at which the integrin-ligand bond experiences the force

exerted by the AFM cantilever.

The findings of this study are in apparent contrast to the decreased affinity observed by use of a similar construct for integrin $\alpha_v\beta_3$ ¹¹⁸. However, in that study, the affinity of the integrin to a soluble ligand was measured, while we determined cell adhesion to an immobilized ligand. As already noted in the seminal work by Bell¹¹⁹, the kinetics of bond formation are drastically affected when confining reaction partners to surfaces or membranes. In general, the bond formation can be separated into two steps: Initially, the reaction partners diffuse towards each other to form an encounter complex (diffusion step). The encounter complex then evolves into the final complex, the so-called reaction step. In solution, the diffusion step is in general faster than the reaction step. In membranes, however, the diffusion step is slowed down by several orders of magnitude. The complex formation becomes diffusion-controlled. Differences in lateral diffusion will therefore have a strong impact on the rate of complex formation. Since talin is required to couple integrins to the actin cytoskeleton, inhibition of talin binding increases the lateral diffusion of the integrin within membranes. Therefore, the increased adhesion rate by TMD-GpA expression (fig. 42) compared to TMD-GpA-I and TMD- $\alpha_v\beta_3$ can be fully explained by its increased diffusion in the membrane. The fast diffusion overcompensates for the low basal affinity caused by TMD tethering. The increased step height - at least in case of TMD-GpA expression - indicates a decrease of the off-rate under force. Also this is in apparent contrast to the previous affinity measurements of Springer and coworkers¹¹⁸. However, in the latter study, integrin-ligand bonds were tested in a force-free environment¹¹⁸, while in all adhesion assays of the present study, the integrin-ligand bond is exposed to forces. In impedance measurements, cells exert forces generated by molecular motors such as myosin II to the environment through integrins. During AFM measurements, integrin-ligand bonds are probed under external forces. This implies that TMD-GpA puts the integrin into a primed state, which is easily activated by a force applied to the integrin-ligand bond, even without TMD separation (fig. 47).

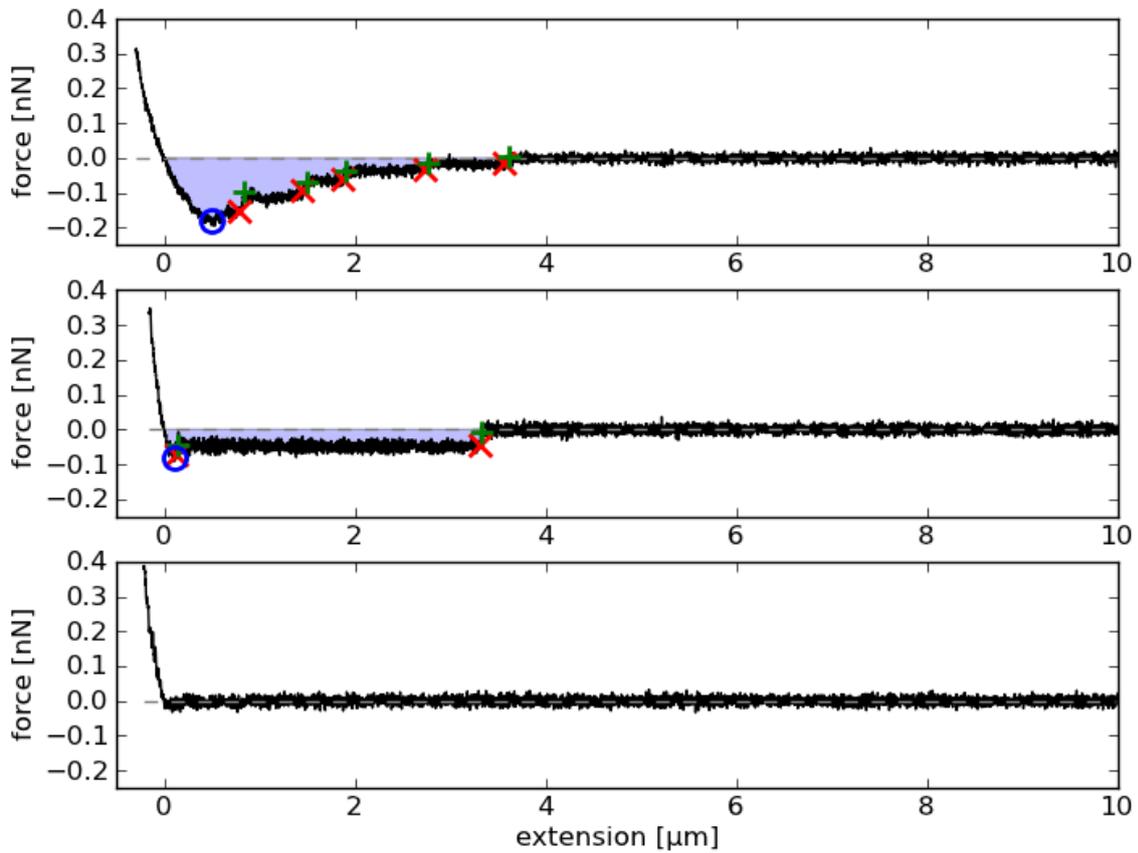


Fig. 41: Typical AFM force-distance curves recorded to evaluate cell adhesive strength. The first two curves depict adhesion with five and two force steps, respectively. The last one does not show adhesion. The peak force (blue circles), the work of cantilever retraction (light blue areas under the curves), the positions, and the heights of the steps (red and green markers) were evaluated.

average number of steps and adhesion rates

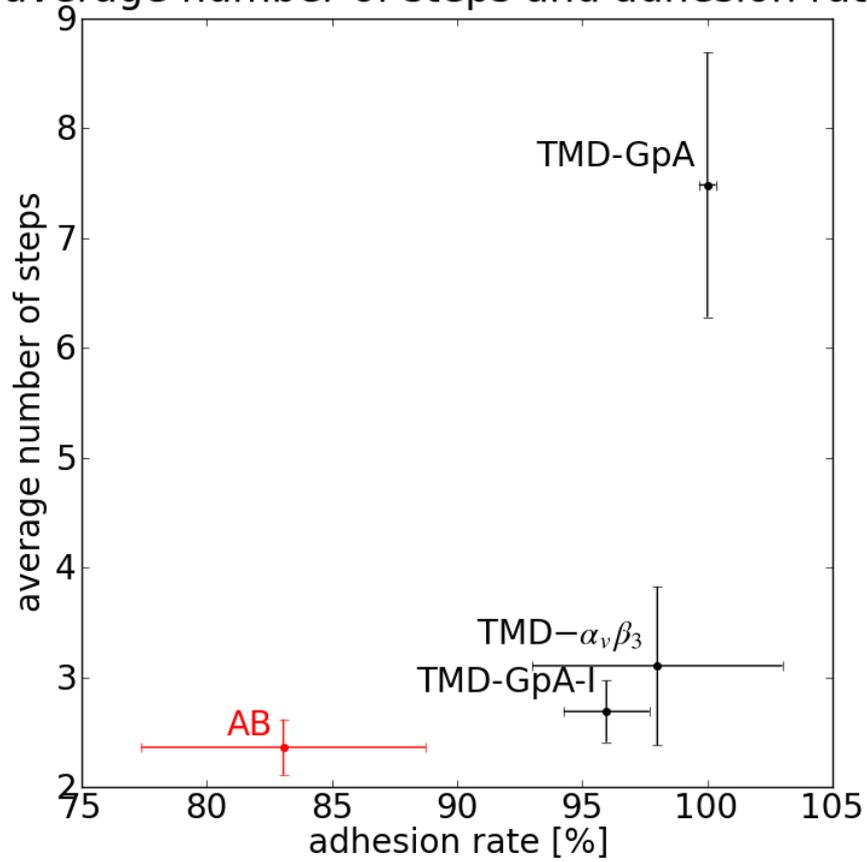


Fig. 42: Average adhesion rate number of steps per adhesive force-distance curve for different $\alpha_v\beta_3$ TMD constructs

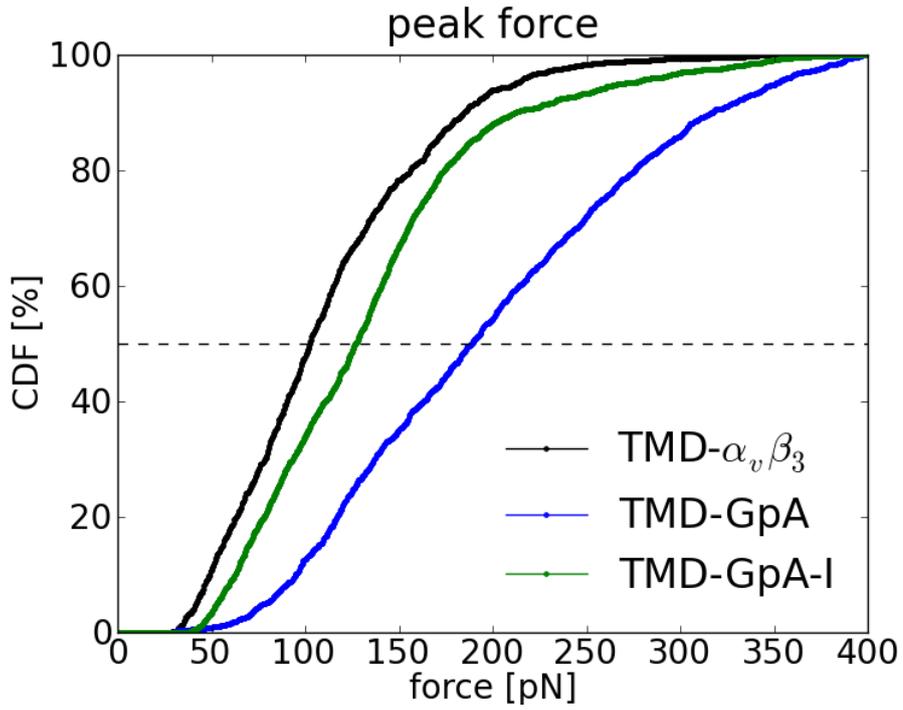


Fig. 43: Cumulative distribution functions of the peak force (blue circles in fig. 41)

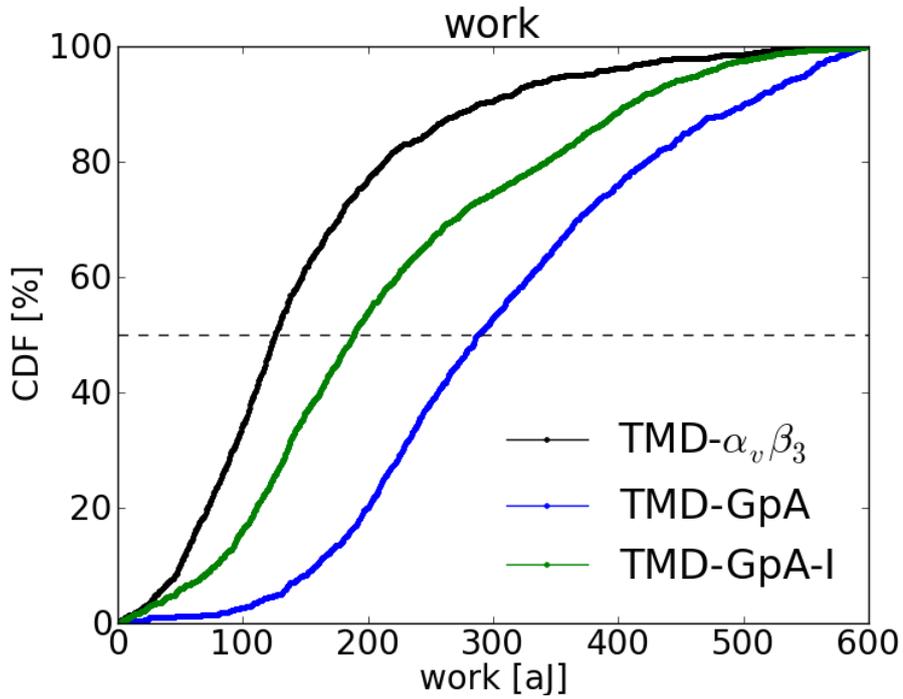


Fig. 44: Cumulative distribution functions of the work of cantilever retraction (light blue area in fig. 41)

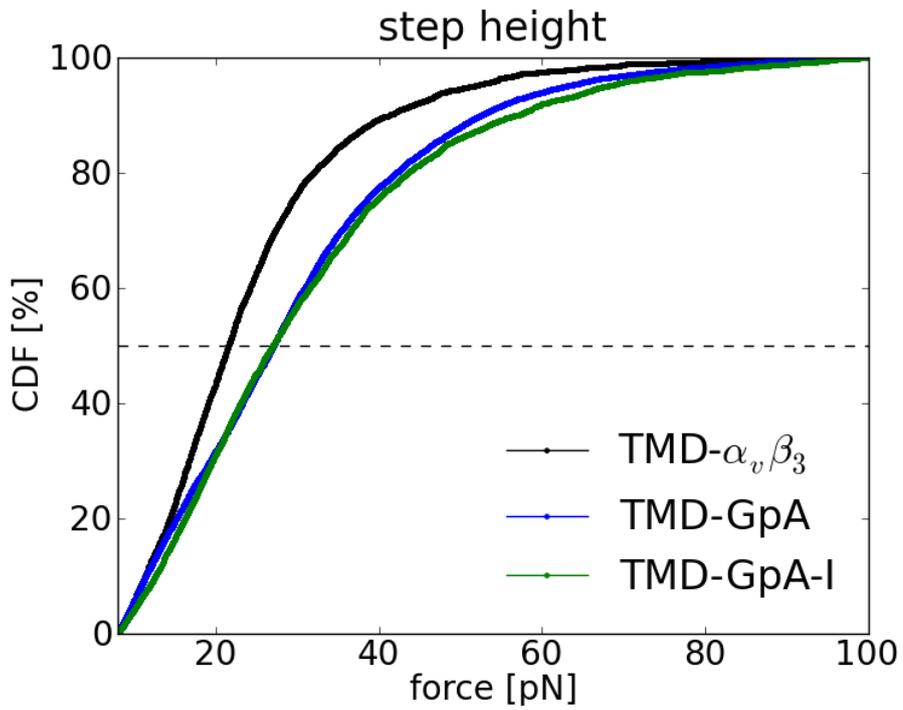


Fig. 45: Cumulative distribution functions of the step heights

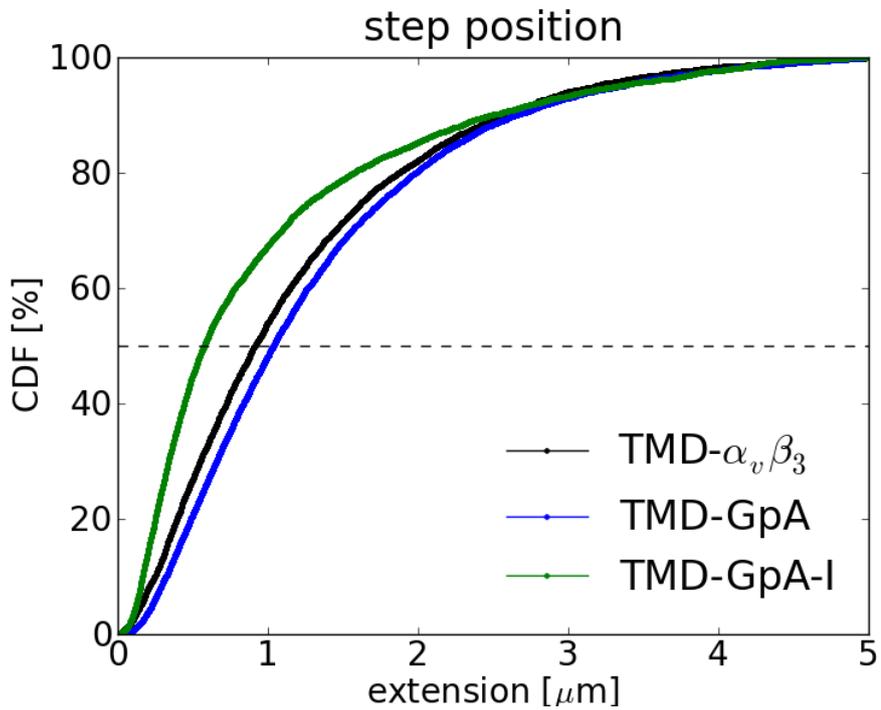


Fig. 46: Cumulative distribution functions of the step positions

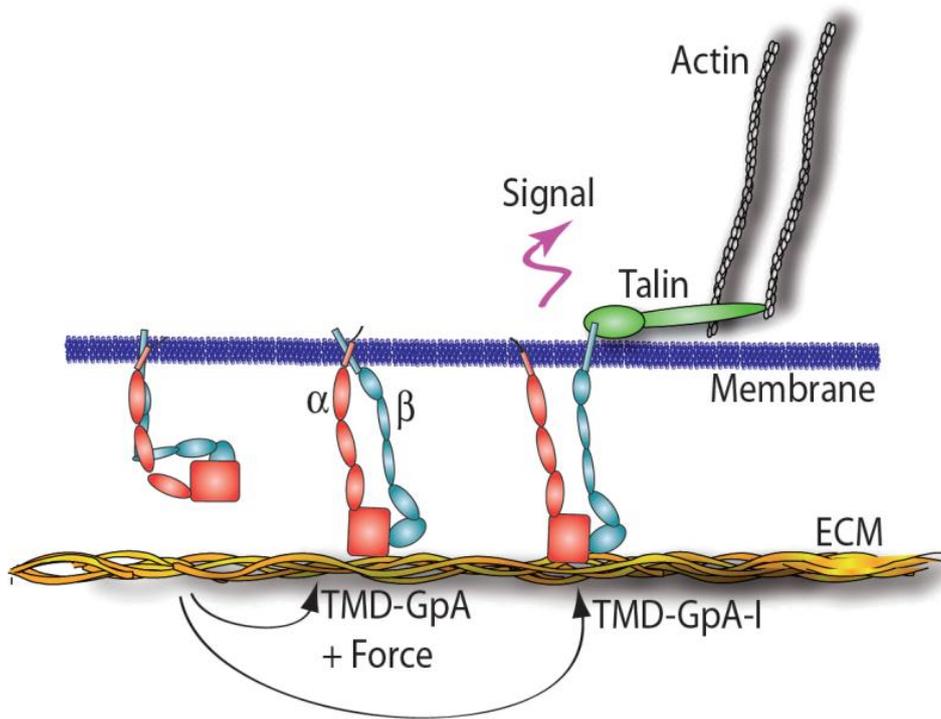


Fig. 47: Effect of the TMD substitutions (modified from Müller et al.¹⁰⁶). From left to right: The TMD-GpA substitution puts the integrin in a state that is easily activated by force, but inhibits full activation by helix-helix interactions of the TMD. The TMD-GpA-I substitution causes a constitutively active integrin, which is linked to intracellular adaptor proteins, such as talin. The two model constructs therefore imitate different stages of cell signaling mediated by the $\alpha_v\beta_3$ TMD.

6. Conclusions

Two novel post-processing techniques have been developed to evaluate single-molecule force spectroscopy data with low signal-to-noise ratios:

ReNoiR, a recursive wavelet-based algorithm, aims at reducing the noise in order to make hidden features of the data visible without inducing disturbing artifacts. In contrast to kernel- or Fourier-based filters, the method does hardly blur or distort the signal, but preserves sudden transitions. As a consequence of increased redundancy of the wavelet expansion and a flexible way of thresholding, it reduces noise more effectively than the other tested techniques, which was demonstrated by various simulated and measured data. The method allowed for the first time to observe the tether onset force of human T lymphocytes. Although the automated parameter estimation is suited for colored noise as well and effectuates a significant improvement in comparison to existing optimization methods for SWT, the underlying empiric findings need a theoretical back-up. For simplicity, only the case $R = 1$ was considered here. By computing $R > 1$ recursions, the quality of the recovered signal might be further enhanced. Since noise is a fundamental problem in experimental sciences and many technologies, ReNoiR is not limited to force spectroscopy data. It may be deployed in completely different applications as well.

In a second approach, hidden rupture events were identified directly in unprocessed noisy force spectra by means of the MSF algorithm. It was shown that an exclusion principle holds for the detection of steps: To be detected, they must be sufficiently wide if their SNR is small and their SNR must be sufficiently high if they are narrow. Within these limits, MSF can be configured to perform a long-range search for low steps or a locally confined search for narrow steps. Thereby, it generally obtains better detection rates than the χ^2 method while needing less computation time. Further, it does not require the user to specify the number of steps to be detected. Instead, a detection sensitivity can be chosen. In contrast to the Matlab implementation of Kersemakers et al., it is able to detect single steps, decaying parts do not result in false-positive detections, and the calculated heights are correct, even if the sections between the steps are not constant. The increased height resolution provides the possibility to detect discrete states in biological data, which are limited by low SNRs. Using the MSF method, unbinding forces below 10 pN could be resolved in atomic force spectroscopy data obtained from living cells.

Out of these primitives, data analysis software was built that allows for an automatic and objective

evaluation of large data sets. With its help, it could be shown that the chemokine SDF-1 α leads to a strengthening of individual bonds between VLA-4 and its corresponding ligand VCAM-1. Further studies should be aimed at identifying residues that inhibit SDF-1 α signaling.

By analyzing the effects of integrin $\alpha_v\beta_3$ mutants on cellular adhesion, a three-state model of integrin activation could be made plausible. Future experiments should focus on two questions: a) is a GpA-like intermediate state physiologically relevant, and b) do the differences in the integrin TMD variants reflect different requirements in signaling plasticity.

7. Appendix

7.1. Curvalyser reference manual

7.1.1. Basic configuration

Before first use, the format and location of the input files must be specified. To this end, open the configuration file *config-curvalyser* in a text editor. If the data is saved in text files, set the variable `file_format` to `'text'` and `columns` to the column numbers of the extension and force records, respectively (note that counting starts with 0). The columns can either be separated by any whitespace (default) or by the string given in `column_delimiter`. This example assumes that the data is contained in the comma-separated columns 2 and 3:

```
file_format = 'text'  
columns = (1,2)  
column_delimiter = ','
```

In case of Curvalyser (**.crv*) and old JPK files (**.out*), a single line is sufficient, as the correct columns are determined automatically:

```
file_format = 'crv'
```

```
file_format = 'jpk-old'
```

New JPK files (**.jpk-force*) can contain multiple “segments”. Therefore, `JPK_segments` must be defined to select the number of the trace and of the retrace segment, respectively:

```
file_format = 'jpk'  
JPK_segments = (0,1)
```

Extension and force can optionally be scaled to a decent order of magnitude:

```
multiplier_x = 1e6  
multiplier_y = 1e12
```

Negative values may be needed if the extension does not increase with the distance from the surface or if indentation does not correspond to a positive sign of the force. All program output is based on the units of the values in the input files, which are multiplied by these factors.

A separate configuration file should be created for each set of force spectra (“experiment”) and stored in the directory *config* (e.g. *001*, *002* and so on). This allows distinguishing the corresponding input and output files. Common configuration parameters can still be defined in *config-curvalyser* and included in each experiment-specific configuration file (such as *config/001*):

```
execfile('config-curvalyser')
file_pattern = 'data/001/*.txt'
```

Here, only a different set of input files is selected by the variable `file_pattern`, which points to the absolute or relative directory containing the force spectra to be analyzed.

By default, the name of the configuration file (e.g. `001`) is used as output directory below `base_output_dir` (here: `output/001`). This can be overridden by the parameter `output_dir`:

```
output_dir = 'output2/001'
```

Further parameters are listed in section 7.1.9 (e.g. to select a certain range of input files or to define the limits of experimental settings to be checked).

7.1.2. Program execution

The Curvalyser is run using the configuration file `config/001` by typing

```
> python Curvalyser.py 001
```

By default, configuration files are searched in the folder `config`. If no configuration file is given, all files in this directory will be processed. Multiple configuration files can be selected individually

```
> python Curvalyser.py 001 002 003 config2/010 config2/020
```

or by file masks:

```
> python Curvalyser.py 00* config2/0?0
```

The following command line options control the tasks performed by the program:

Option	Example	Description
-h		show a help message
-l	-l output/log.txt	set log file
-c	- c"exp_id=='001'"	select experiments by a Python expression
-e	-e WT	select only experiments belonging to the given experiment type
-O		overwrite existing output files
-f	-f data/*.txt	same as configuration parameter <code>file_pattern</code>
-r	-r 100:200:10	same as configuration parameter <code>file_range</code>
-o	-o output	same as configuration parameter <code>base_output_dir</code>
-v, -vv		be verbose (-v) or very verbose (-vv)
-d	-d 4.5	set the primary de-noising parameter

-i	-i 5	set the relative indicator threshold
-p		create force plots (-pf), indicator plots (-pi) or both (-pfi)
-s		show plots in a graphical user interface
-V		display program version

A very flexible way to pick out experiments according to some rule is provided by the option “-c”. It takes a Python expression as parameter returning either true or false, depending on whether the file is to be included or not. The condition can be attached to the consecutive configuration file number (`config_file_no`), the file name (`config_file`), the experiment ID (`exp_id`) or any configuration parameter (stored in the dictionary `config`). In this example, only experiments whose ID start with the prefix “WT” are chosen:

```
> python Curvalyser.py -c"exp_id[:2]=='WT'"
```

To select experiments of a certain type (specified by the configuration parameter `experiment_type`), it is more convenient to use the option “-e”.

7.1.3. Baseline correction

Force curves can be baseline-corrected to compensate instrumental drift and to allow for correct calibration of the contact point as well as of the zero-force level. Depending on the configuration parameter `fit_baseline`, either a linear (1) or a quadratic fit (2) is subtracted from the retrace curve before any further evaluation is done. Fitting is performed in a smooth part at the end of the retrace curve where the distance from the surface is maximal and only background noise is present. This range must be free of steps, tip-surface interactions or other sample-specific effects and sufficiently long to obtain an accurate fit. Therefore, the length of the retrace extension should not be too short. A baseline fit is calculated iteratively every `baseline_step_len` data points, starting from the end of the curve. It stops automatically if the residual sum of squares (RSS) locally exceeds the expected value by the factor `baseline_max_rel_local_RSS`. The end point of the fit interval is reverted to the last local minimum of the RSS if `baseline_return_to_local_min` is set to 1. In case the length of the fit falls below `baseline_fit_min_width`, the curve is excluded from further analysis. If the baseline fit stops too early, a higher threshold `baseline_max_rel_local_RSS` or a lower step interval `baseline_step_len` should be chosen (and vice versa).

7.1.4. Noise reduction

Noise reduction is crucial for reliable step detection, contact point calibration and for accurate determination of some characteristics of the force curves, such as the peak force. Different methods can be selected: ReNoiR (see section 2.1.2), Gaussian smoothing⁴⁴ and the Savitzky-Golay filter⁴³. The first is preferable if the force spectra contain sharp features (such as spikes or steep, narrow steps) and the second is ideal for curves with more or less constant plateaus, which may be separated by smooth transitions.

The filter strength of ReNoiR (`denoising_method = 'renoir'`) depends on the parameters T_0 (`denoising_param`) and T_1 (`denoising_param2`). See section 2.1.2 for details about the algorithm. The number of recursions (`denoising_recursions`) and levels (`denoising_levels`) usually do not need to be changed. To minimize border distortion caused by the wavelet filter, the signal can be extended by point reflection at both ends. `denoising_padding_lt` and `denoising_padding_rt` are the numbers of padded data points on the left and right, respectively. The resulting number of samples should be a power of two.

If Gaussian smoothing (`denoising_method = 'gauss'`) is performed, `denoising_param` represents the standard deviation of the smoothing kernel⁴⁴.

In case of Savitzky-Golay filtering (`denoising_method = 'savgol'`), the same parameter determines the half window size and `savgol_order` the order of the polynomial kernel⁴³.

`denoising_param` and `denoising_param2` can be user-defined callback functions. They must take three parameters: `noise_level` (the estimated standard deviation of the noise), `rows` (the number of samples), and `meta`. The latter is a dictionary containing metadata, such as the headers of JPK force spectra (`*.out`, `*.jpk-force`) or the additional information comprised by Curvalyser files (`*.crv`).

Example:

```
denoising_param = lambda noise_level, rows, meta: noise_level * 3
```

7.1.5. Contact point calibration

The extension where the indentation force becomes zero for the first time during retraction is referred to as “contact point”. Its determination is necessary to calibrate the zero point of the extension, e.g. of detected steps (see section 7.1.6) or of the peak force. Although most

subsequent calculations are influenced by the calibration, it is only performed if `find_contact_pos` is `1`. Otherwise, the extension at the reversal point between trace and retrace curve defines the zero point.

If no contact point was found in the retrace curve, if it was found further away from the reversal point than stipulated by `max_contact_pos` or if the maximum indentation force is lower than `min_contact_force`, the trace curve is used instead. Should this also fail, then the calibration and all parameters depending on it will be omitted.

7.1.6. Step detection

A moving step fit is deployed to detect upward steps in the retrace curve (see section 3.1.2 for a description of the algorithm). Briefly, it produces an indicator that is correlated to the step heights. Local maxima correspond to possible step positions. The half window size `MSF_window` should be set to `1` if the steps are narrow or sharp (i.e. showing steep flanks) or to a higher value if the steps are smooth (try e.g. `10` or `100`).

An absolute threshold `indicator_threshold` and a relative one (`indicator_relative_threshold`) that is multiplied by the noise level of the indicator, define the detection sensitivity (the greater of the two values counts). False-positives can be suppressed by smoothing the indicator with a Gaussian kernel with standard deviation `MSF_sigma`, by defining the minimum step height (`step_min_height`), the minimum/maximum step width (`step_min_width/step_max_width`), or the minimum average amplitude of the indicator (`step_min_slope`).

The distance between lower and upper corner of detected steps can be narrowed by increasing `step_confinement_lt` and `step_confinement_rt` (and vice versa). Both parameters can vary between `0` and `1`.

Step heights are determined by the force difference between the corners. The force levels at the left/right of the steps are determined by linear fits over `step_fit_min_len` to `step_fit_max_len` data points to the left/right starting `step_fit_clearance_lt/step_fit_clearance_rt` points left/right of the lower/upper corner. Within the given limits, the fit interval is optimized automatically to yield the best fit in the sense of minimum root mean square error. In any case, the interval is terminated at the nearest neighboring step.

To exclude the beginning or end of the retrace curve, `indicator_margin_lt` or `indicator_margin_rt` can be set to a positive value.

7.1.7. Fitting procedures

The slopes on the left and right of the detected steps and on the left and right of the contact point are fitted if `step_fit_width`, `contact_pos_lt_fit_width` and `contact_pos_rt_fit_width`, respectively, are set to a positive value. The indentation part of the trace curve (from the contact point to the left) can be fitted with the user-defined callback function `trace_fit_function`, which is initialized with the parameters defined in the Python sequence `trace_fit_init_params`. Analogously, the section of the retrace curve between the contact point and the first step can be fitted using `retrace_fit_function` and `retrace_fit_init_params`.

Plotting

To determine if the force curves and indicators used for step detection are plotted into files of the format `plot_format`, set `plot_force_curves` and `plot_indicators` to either `0` or `1`. The plots will be written to the subdirectory `plots`. Additionally, they are displayed in a graphical user interface if `show_plots` is `1`. The horizontal and vertical data range is specified by `plot_xmin`, `plot_xmax`, `plot_ymin`, and `plot_ymax`. The resolution of the output files can be changed by `plot_size_force_curves` and `plot_size_indicator_curves`. `fit_drawing_width` defines the extension of the lines drawn into the plots to depict the fitted slopes on the left and right corner of the steps.

7.1.8. Output files

All extracted information about the force curves is saved into two files in the output directory (see section 7.1.1): `curves.txt` contains all parameters specific to the whole curve and `steps.txt` all data related to the steps. The first line of the files is a header shortly describing the meaning of the columns. Basically, the units found in the input files are used, but may be scaled by multiplication factors (`multiplier_x` and `multiplier_y`).

Plots are written to the subdirectory `plots`. The file names contain the names of the corresponding input files. “(failed)” is appended if a curve could not be analyzed (e.g. because the baseline correction failed).

7.1.9. Overview of important configuration parameters

Parameter	Values / example	Description
Input and output		
<code>file_pattern</code>	<code>'data/*.txt'</code>	file mask for force curves to be analysed
<code>file_format</code>	<code>'txt'</code>	text format (also set <code>columns!</code>)
	<code>'jpk'</code>	<u>new</u> JPK format (<code>*.jpk-force</code>)
	<code>'jpk-old'</code>	<u>old</u> JPK format (<code>*.out</code>)
	<code>'crv'</code>	Curvalyser format; default
<code>file_range</code>	<code>'100:200:10'</code>	select a range of files (format: 'start:stop:step,start:stop:step,...'; counting starts with 1; negative value: count from the end)
<code>columns</code>	<code>(1,2)</code>	specify the columns containing extension and force data, respectively (counting starts with 0!)
<code>column_delimiter</code>	<code>','</code>	string used to separate values; default: any whitespace
<code>JPK_segments</code>	<code>(0,2)</code>	segment numbers of the trace and retrace data, respectively
<code>multiplier_x</code>	<code>1e6</code>	multiplier for rescaling the extension data
<code>multiplier_y</code>	<code>1e12</code>	multiplier for rescaling the force data
<code>output_dir</code>	<code>'output2/001'</code>	output directory (automatically determined by <code>base_output_dir</code> and <code>exp_id</code> if omitted)
<code>exp_id</code>	<code>'123'</code>	ID used to distinguish multiple sets of force spectra (experiments); default: name of the configuration file
<code>base_output_dir</code>	<code>'output'</code>	base output directory; only used if <code>output_dir</code> is omitted
<code>nominal_values</code>	<code>{'sensitivity': (50e-9,10e-9)}</code>	Python dictionary; first field: nominal value; second field: maximum deviation
<code>assert_nominal_values</code>	<code>0</code>	warn only
	<code>1</code>	exclude force curves conflicting with the limits defined in <code>nominal_values</code> ; default
<code>unit_x</code>	<code>'um'</code>	unit of the extension data (only used for plots)
<code>unit_y</code>	<code>'pN'</code>	unit of the force data (only used for plots)
Baseline correction		
<code>fit_baseline</code>	<code>0</code>	do not correct baseline
	<code>1</code>	subtract a linear baseline; default
	<code>2</code>	subtract a quadratic baseline
<code>baseline_max_rel_local_RSS</code>	<code>2.5</code>	relative threshold for the termination of

		the baseline fit; default: 2.0
baseline_fit_min_width	4.0	minimum baseline fit length for a curve to be further analysed (in units of the extension data)
baseline_step_len	100	step length for the iterative baseline fit (default: automatic)
baseline_return_to_local_min	0	the baseline fit is performed up to the point, where it is terminated
	1	the baseline fit is performed up to the last local minimum of the RSS after termination; default
De-noising		
denoising_method	'renoir'	use ReNoiR for noise reduction
	'gauss'	use Gaussian smoothing; default
	'savgol'	use the Savitzky-Golay filter
denoising_param	4.1	primary parameter for the noise reduction filter or user-defined Python function
denoising_param2	20.8	secondary parameter for the noise reduction filter or user-defined Python function
denoising_recurions	1	number of recursions (ReNoiR only); default: 1
denoising_wavelet	'haar'	name of the wavelet used; default: 'haar'
denoising_levels	0	number of levels (ReNoiR only); default: 0 (automatic)
savgol_order	5	order of the polynomial kernel (Savitzky-Golay filter only)
Contact point calibration		
find_contact_pos	0	do not detect the contact point
	1	detect the contact point; default
max_contact_pos	2.0	maximum distance of the contact point from the beginning of the curve; default: do not check
min_contact_force	0	minimum force at the contact point; default: 0
Step detection		
MSF_sigma	2.5	Gaussian smoothing of the curve
MSF_window	100	width of the moving fit window (number of samples); default: 1
indicator_margin_lt	100	left margin of the indicator; default: 0
indicator_margin_rt	100	right margin of the indicator; default: 0
indicator_threshold	40	step detection sensitivity (absolute value)
indicator_relative_threshold	10	step detection sensitivity (relative to

		noise level after de-noising); default: 5
max_steps	20	maximum number of detected steps
step_confinement_lt	0.9	sensitivity for locating the lower corner of steps; default: 0
step_confinement_rt	0.9	sensitivity for locating the upper corner of steps; default: 0
step_fit_min_len	3	minimum size of the window for step fits
step_fit_max_len	2000	maximum size of the window for step fits; default: no limit
step_fit_clearance_lt	10	additional gap between step and left fit window; default: 0
step_fit_clearance_rt	10	additional gap between step and right fit window; default: 0
step_min_height	15	minimum step height; default: do not check
step_min_width	1	minimum step width; default: do not check
step_max_width	100	maximum step width; default: do not check
step_min_slope	250	minimum average indicator amplitude; default: do not check
Fitting		
indentation_fit	0	fit indentation slopes at contact point; default
	> 0	fit slopes at specified indentation force
indentation_curve	0	use retrace curve; default
	1	use trace curve (never denoised)
contact_pos_lt_fit_width	0.02	width of the indentation fit to the left; default: do not fit
contact_pos_rt_fit_width	0.01	width of the indentation fit to the right; default: do not fit
indentation_fit_avg_window	3	half size of the averaging window for finding a specified indentation force; default: 0
step_fit_width	0.01	fit length for step slopes; default: do not fit
trace_fit_function		fit function for trace curve from beginning to contact point
trace_fit_init_params		initial fit parameters for trace fit
retrace_fit_function		fit function for retrace curve from contact point to first step (if number of steps is 1)
retrace_fit_init_params		initial fit parameters for retrace fit
single_step_fit_function		fit function for retrace curve from

		contact point to first and only step (number of steps must be 1)
<code>single_step_init_params</code>		initial fit parameters for first tether fit
Plotting		
<code>plot_force_curves</code>	0	do not plot force curves
	1	plot force curves; default
<code>plot_indicators</code>	0	do not plot indicators
	1	plot indicators; default
<code>show_plots</code>	0	do not show plots; default
	1	show plots in a graphical user interface
<code>plot_format</code>	'pdf'	file format for plots; default: 'png'
<code>plot_xmin</code>	-1	minimum extension; default: auto-scale
<code>plot_xmax</code>	20	maximum extension; default: auto-scale
<code>plot_ymin</code>	-500	minimum force; default: auto-scale
<code>plot_ymax</code>	500	maximum force; default: auto-scale
<code>plot_size_force_curves</code>	(640,480)	plot size (pixels)
<code>plot_size_indicator_curves</code>	(640,480)	plot size (pixels)
<code>fit_drawing_width</code>	0.01	width of the step slope fits; default: do not draw
Miscellaneous		
<code>verbose</code>	2	level of verbosity (0-2); default: 0
<code>spring_constant</code>	0.1	spring constant used for force calibration if not specified in input file
<code>length_correction</code>	0	do not convert extension to distance; default
	1	convert extension to distance
<code>experiment_type</code>	'WT'	string used to categorise experiments by a user-defined experiment type

7.2. Paramalyser reference manual

7.2.1. Configuration

The Paramalyser uses the same configuration files as the Curvalyser (see section 7.1.1) and an additional one, which defines the settings needed for statistical evaluation. It is specified by the command line option “-C” (default: *config-paramalyser*). The following configuration parameters are evaluated:

Parameter	Values / example	Description
Input and output		
<code>fixed_input_dir</code>	<code>'output2'</code>	input directory; default: Curvalyser output directory
<code>fixed_output_dir</code>	<code>'output2/stats'</code>	output directory; default: <code>statistics_dir</code> in Curvalyser output directory
<code>statistics_dir</code>	<code>'stats'</code>	relative output directory; default: <i>statistics</i>
<code>multiplier_x</code>	<code>1e6</code>	multiplier for rescaling the extension data
<code>multiplier_y</code>	<code>1e12</code>	multiplier for rescaling the force data
<code>experiment_types</code>	<code>['WT', 'A2', 'B1']</code>	selection and order of experiment types to be analysed
<code>analyse_all</code>	<code>0</code>	analysed only experiment types specified by <code>experiment_types</code> in the given order
	<code>1</code>	analyse all experiment types; the order specified by <code>experiment_types</code> is still regarded
<code>ignore_ids</code>	<code>['001', '002']</code>	experiment IDs to ignore
<code>curves_range_start</code>	<code>100</code>	first curve to include (counting starts with 1; negative value: count from the end); can be overwritten by <code>paramalyser_curves_range_start</code> in Curvalyser configuration file
<code>curves_range_stop</code>	<code>200</code>	last curve to include (counting starts with 1; negative value: count from the end); can be overwritten by <code>paramalyser_curves_range_stop</code> in Curvalyser configuration file
<code>curves_range_step</code>	<code>10</code>	step size (negative value: go backwards); default: auto +1/-1; can be overwritten by <code>paramalyser_curves_range_step</code> in Curvalyser configuration file
<code>autosplit_size</code>	<code>10</code>	split curves into chunks of the given size
<code>autosplit_shift</code>	<code>10</code>	shift between successive auto-split chunks
Tasks		

tasks	['CDF', 'CDF_cum']	list of tasks to perform (see section 7.2.3); default: all
plot_params	['force', 'work']	list of parameters to analyse within each task (see section 7.2.4); default: all
plot_steps	['first', 'last']	list of subset plots to create (<code>all</code> : all curves, <code>single</code> : only curves with exactly one step, <code>first</code> : only first step, <code>last</code> : only last step); default: all
Filters		
step_filter_limits	{'height': (10, None), 'work': (0, 100)}	Python dictionary of criteria to filter steps (tuple of minimum and maximum for each parameter to be checked); all criteria must be met
curve_filter_limits	{'noise_sigma': (None, 10), 'steps_ctr': (1, 1)}	Python dictionary of criteria to filter curves <u>and</u> steps (tuple of minimum and maximum for each parameter to be checked); all criteria must be met; applied after <code>step_filter_limits</code>
Plotting		
plot_ranges	{'work': (0, 100)}	Python dictionary of default plot ranges (tuple of minimum and maximum for each parameter)
histogram_ranges		like <code>plot_ranges</code> , but for histograms
boxplot_ranges		like <code>plot_ranges</code> , but for boxplots
scatterplot_ranges		like <code>plot_ranges</code> , but for scatterplots
plot_size	(800, 600)	default plot size (tuple of width and height)
plot_size_histogram		like <code>plot_size</code> , but for histograms
plot_size_CDF		like <code>plot_size</code> , but for CDFs
plot_size_scatter		like <code>plot_size</code> , but for scatterplots
plot_size_boxplot		like <code>plot_size</code> , but for boxplots
plot_size_errorbar		like <code>plot_size</code> , but for errorbar plots
plot_size_errorbar_cum		like <code>plot_size</code> , but for cumulated errorbar plots
plot_size_errorbar_corr		like <code>plot_size</code> , but for correlated errorbar plots
histogram_bins	100	number of histogram bins; default: 50
histogram_ylim	(0, 1)	vertical range of histograms (tuple of minimum and maximum)
histogram_norm_curves	0	do not normalise histograms over curve-specific parameters; default
	1	normalise to one
	2	normalise to average number of steps
	3	normalise to adhesion rate
histogram_norm_steps	0	do not normalise histograms over step-specific parameters; default

	1	normalise to one
	2	normalise to average number of steps
	3	normalise to adhesion rate
CDF_histogram_bins	100	number of histogram bins for CDF plots (default: sum of counts)
experiment_type_labels	{'WT': 'wild-type'}	Python dictionary defining custom labels for experiment types
experiment_type_colors	{'WT': 'blue'}	Python dictionary defining custom colours for experiment types
Miscellaneous		
cumulation_mode	0	cumulate data for errorbar plots globally
	1	cumulate and calculate averages/medians/modals by experiment type; default
steps_ctr_averaging	0	consider all curves to calculate the average number of steps
	1	consider only adhesive curves; default
output_format	'text'	save data files in standard text format; default
	'igor'	save data files in an Igor-compatible text format
custom_script	'custom.py'	custom script to be executed
curvalyser_config		Python dictionary of Curvalyser configuration parameters (overwrites other settings)

7.2.2. Program execution

The Paramalyser is invoked the same way as the Curvalyser (see section 7.1.2). Example:

```
> python Paramalyser.py -e WT
```

It understands these command line options:

Option	Example	Description
-h		show a help message
-C	-C pconfig.txt	set the configuration file to be used; default: <i>config-paramalyser</i>
-c	-c "exp_id=='001'"	select experiments by a Python expression
-e	-e WT,A2,B1	selection and order of experiment types (same as configuration parameter <code>experiment_types</code>)
-a		analyse all experiment types (same as configuration parameter <code>analyse_all</code>); default: only specified types
-i	-i output2	set the input directory (same as configuration parameter

		<code>fixed_input_dir</code>); default: Curvalyser output directory
<code>-o</code>	<code>-o output2/stats</code>	set the output directory (same as configuration parameter <code>fixed_output_dir</code>); default: <code>statistics_dir</code> in Curvalyser output directory
<code>-t</code>	<code>-t CDF, CDF_cum</code>	same as configuration parameter <code>tasks</code>
<code>-p</code>	<code>-p force, work</code>	same as configuration parameter <code>plot_params</code>
<code>-s</code>	<code>-s first, last</code>	same as configuration parameter <code>plot_steps</code>
<code>-V</code>		display program version

7.2.3. Tasks

Data can either be analyzed per-experiment or cumulated over all experiments of the same type.

The following tasks can be performed (specified by `tasks`, see section 7.2.1):

Task	Description	Output directory
<code>data</code>	save parameters to text files	<code>data</code>
<code>data_cum</code>	save cumulated parameters to text files	<code>cumulated data</code>
<code>histos</code>	create histograms	<code>histograms/*</code>
<code>histos_cum</code>	create histograms of cumulated data	<code>histograms</code>
<code>CDF</code>	plot cumulative distribution function (CDF)	<code>CDFs</code>
<code>CDF_cum</code>	plot cumulative distribution function (CDF) of cumulated data	<code>cumulated CDFs</code>
<code>CDF_data_cum</code>	save the 50% values of the cumulated CDF plots to text files	<code>cumulated data</code>
<code>scatter</code>	create scatter plots	<code>scatter plots/*</code>
<code>scatter_cum</code>	create cumulated scatter plots	<code>scatter plots</code>
<code>boxplots</code>	create boxplots	<code>boxplots</code>
<code>boxplots_cum</code>	create cumulated boxplots	<code>cumulated boxplots</code>
<code>averages</code>	plot average values	<code>averages</code>
<code>averages_cum</code>	plot cumulated average values	<code>cumulated averages</code>
<code>averages_data</code>	save averages to text files	<code>data</code>
<code>averages_data_cum</code>	save cumulated averages to text files	<code>cumulated data</code>
<code>averages_boxplots</code>	create boxplots of average values	<code>boxplots of averages</code>
<code>medians</code>	plot medians	<code>medians</code>
<code>medians_cum</code>	plot cumulated medians	<code>cumulated medians</code>
<code>medians_data</code>	save medians to text files	<code>data</code>
<code>medians_data_cum</code>	save cumulated medians to text files	<code>cumulated data</code>
<code>medians_boxplots</code>	create boxplots of medians	<code>boxplots of medians</code>
<code>modals</code>	plot modals	<code>modals</code>

<code>modals_cum</code>	plot cumulated modals	<i>cumulated modals</i>
<code>modals_data</code>	save modals to text files	<i>data</i>
<code>modals_data_cum</code>	save cumulated modals to text files	<i>cumulated data</i>
<code>modal_boxplots</code>	create boxplots of modals	<i>boxplots of modals</i>
<code>other</code>	plot adhesion rates and average number of steps	.
<code>tests</code>	calculate Mann-Whitney and Kruskal-Wallis tests	.
<code>custom</code>	execute the custom script defined by <code>custom_script</code>	

7.2.4. Parameters

The following parameters can be analyzed (specified by `plot_params`, see section 7.2.1):

Parameter	Description
Curve-specific	
<code>steps_ctr</code>	number of steps
<code>peak_pos</code>	sample number of the peak force (global minimum/maximum force)
<code>peak_extension</code>	extension of the peak force
<code>peak_force</code>	peak force
<code>indent_force</code>	indentation force
<code>work</code>	work (area between the baseline and the retrace curve)
<code>bl_itcpt</code>	interception of the baseline
<code>bl_slope</code>	slope of the baseline
<code>bl_crvtr</code>	curvature of the baseline (in case of a quadratic fit)
<code>bl_fit_len</code>	length of the baseline fit (number of samples)
<code>bl_avg_RSS</code>	average residual sum of squares of the baseline fit
<code>contact_pos</code>	sample number of the contact point
<code>indent_slope_l</code>	left indentation slope of the retrace curve
<code>indent_slope_r</code>	right indentation slope of the retrace curve
<code>indent_slope_r_normed</code>	like <code>indent_slope_r</code> , but divided by the number of steps
<code>trace_fit1</code>	parameters obtained by the custom fit function
<code>trace_fit2</code>	<code>trace_fit_function</code>
<code>trace_fit3</code>	
<code>retrace_fit1</code>	parameters obtained by the custom fit function
<code>retrace_fit2</code>	<code>retrace_fit_function</code>
<code>retrace_fit3</code>	
<code>noise_sigma</code>	estimated standard deviation of the noise
<code>denoising_param</code>	value of the actually used primary de-noising parameter
<code>denoising_param2</code>	value of the actually used secondary de-noising parameter
<code>indicator_thld</code>	value of the actually used indicator threshold
Step-specific	
<code>lt_pos</code>	sample number of the lower corner of the step

lmax_pos	sample number of the local maximum position
rt_pos	sample number of the upper corner of the step
extension	extension of the lower corner of the step
force	force at the lower corner of the step
height	relative height of the step (difference of the forces at the upper and lower corner)
avg_slope	average value of the indicator
max_slope	maximum value of the indicator
plateau_slope_l	fitted slope on the left of the step
plateau_slope_r	fitted slope on the right of the step
stiffness	$\text{height} / \text{extension}$

References

1. Binnig G, Quate CF, & Gerber C (1986) Atomic Force Microscope. *Phys. Rev. Lett.* 56(9):930-933.
2. Ashkin A (1970) Acceleration and Trapping of Particles by Radiation Pressure. *Phys. Rev. Lett.* 24(4):156–159.
3. Ashkin A, Dziedzic JM, Bjorkholm JE, & Chu S (1986) Observation of a single-beam gradient force optical trap for dielectric particles. *Optics Letters* 11(5):288-290.
4. Smith S, Finzi L, & Bustamante C (1992) Direct mechanical measurements of the elasticity of single DNA molecules by using magnetic beads. *Science* 258:1122–1126.
5. Schmitz J (2008) Cell Adhesion mediated by the Integrin VLA-4. (Ludwig Maximilian University Munich).
6. Holschneider M, Kronland-Martinet R, Morlet J, & Tchamitchian P (1989) A real-time algorithm for signal analysis with the help of the wavelet transform. *Wavelets, Time-Frequency Methods and Phase Space*, eds J. M. Combes AG & Tchamitchian P (Springer), pp 286-297.
7. Dutilleul P (1989) An implementation of the algorithm à trous to compute the wavelet transform. *Wavelets: Time-Frequency Methods and Phase Space, Inverse Problems and Theoretical Imaging*, (Springer), p 298.
8. Durand S & Froment J (2001) Artifact free signal denoising with wavelets. *IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp 3685-3688.
9. Xiong J-P, *et al.* (2001) Crystal Structure of the Extracellular Segment of Integrin $\alpha V\beta 3$. *Science* 294(5541):339-345.
10. Humphries MJ (2000) Integrin structure. *Biochemical Society Transactions* 28:311-339.
11. Hynes RO (2002) Integrins: bidirectional, allosteric signaling machines. *Cell* 110:673-687.
12. Hemler ME (1990) VLA Proteins in the Integrin Family: Structures, Functions, and Their Role on Leukocytes. *Annual Review of Immunology* 8(1):365-400.
13. Sixt M, Bauer M, Lämmermann T, & Fässler R (2006) Beta1 integrins: zip codes and signaling relay for blood cells. *Curr Opin Cell Biol* 18(5):482-490.
14. Brakebusch C, *et al.* (2000) Skin and hair follicle integrity is crucially dependent on beta 1 integrin expression on keratinocytes. *EMBO J* 19(15):3990-4003.
15. Williams MA & Solomkin JS (1999) Integrin-mediated signaling in human neutrophil

- functioning. *J Leukoc Biol* 65(6):725-736.
16. Kim C, Ye F, & Ginsberg MH (2011) Regulation of Integrin Activation. *Annual Review of Cell and Developmental Biology* 27(1):321-345.
 17. Harburger DS & Calderwood DA (2009) Integrin signalling at a glance. *Journal of cell science* 122:159-163.
 18. Kinashi T (2005) Intracellular signalling controlling integrin activation in lymphocytes. *Nat Rev Immunol* 5:546–559.
 19. Shimaoka M, *et al.* (2006) AL-57, a ligand-mimetic antibody to integrin LFA-1, reveals chemokine-induced affinity up-regulation in lymphocytes. *PNAS* 103:13991-13996.
 20. Hyduk SJ & Cybulsky MI (2009) Role of $\alpha 4\beta 1$ Integrins in Chemokine-Induced Monocyte Arrest under Conditions of Shear Stress. *Microcirculation* 16(1):17-30.
 21. Ye F, Kim C, & Ginsberg MH (2011) Molecular mechanism of inside-out integrin regulation. *Journal of Thrombosis and Haemostasis* 9:20-25.
 22. Laudanna C & Bolomini-Vittori M (2009) Integrin activation in the immune system. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine* 1(1):116-127.
 23. Kinashi T (2012) Overview of integrin signaling in the immune system. (Translated from eng) *Methods Mol Biol* 757(1940-6029 (Electronic)):261-278 (in eng).
 24. Plow EF & Ma Y-Q (2007) Inside-out, outside-in: what's the difference? *Blood* 109(8):3128-3129.
 25. Tadokoro S, *et al.* (2003) Talin binding to integrin beta tails: a final common step in integrin activation. *Science* 302:103-106.
 26. Hyduk SJ, *et al.* (2007) Phospholipase C, calcium, and calmodulin are critical for $\alpha 4\beta 1$ integrin affinity up-regulation and monocyte arrest triggered by chemoattractants. *Blood* 109(1):176-184.
 27. Ley K, Laudanna C, Cybulsky MI, & Nourshargh S (2007) Getting to the site of inflammation: the leukocyte adhesion cascade updated. *Nat Rev Immunol* 7(9):678-689.
 28. Kobilka BK (2007) G protein coupled receptor structure and activation. *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1768(4):794-807.
 29. Calderwood DA, *et al.* (2002) The Phosphotyrosine Binding-like Domain of Talin Activates Integrins. *Journal of Biological Chemistry* 277(24):21749-21758.
 30. Schmitz J & Gottschalk K-E (2008) Mechanical regulation of cell adhesion. *Soft Matter* 4(7):1373-1387.

31. Peled A, *et al.* (2000) The chemokine SDF-1 activates the integrins LFA-1, VLA-4, and VLA-5 on immature human CD34+ cells: role in transendothelial/stromal migration and engraftment of NOD/SCID mice. *Blood* 95(11):3289-3296.
32. Lapidot T, Dar A, & Kollet O (2005) How do stem cells find their way home? *Blood* 106:1901-1910.
33. DiVietro JA, Brown DC, Sklar LA, Larson RS, & Lawrence MB (2007) Immobilized Stromal Cell-Derived Factor-1 α Triggers Rapid VLA-4 Affinity Increases to Stabilize Lymphocyte Tethers on VCAM-1 and Subsequently Initiate Firm Adhesion. *The Journal of Immunology* 178(6):3903-3911.
34. Woolf E, *et al.* (2007) Lymph node chemokines promote sustained T lymphocyte motility without triggering stable integrin adhesiveness in the absence of shear forces. *Nat Immunol* 8:1076-1085.
35. Alon R, *et al.* (2005) $\alpha 4\beta 1$ -dependent adhesion strengthening under mechanical strain is regulated by paxillin association with the $\alpha 4$ -cytoplasmic domain. *The Journal of Cell Biology* 171(6):1073-1084.
36. Laudanna C & Alon R (2006) Right on the spot. Chemokine triggering of integrin-mediated arrest of rolling leukocytes. (Translated from eng) *Thromb Haemost* 95(1):5-11 (in eng).
37. Campbell ID & Humphries MJ (2011) Integrin Structure, Activation, and Interactions. *Cold Spring Harbor Perspectives in Biology* 3(3).
38. Springer TA (1997) Folding of the N-terminal, ligand-binding region of integrin α -subunits into a β -propeller domain. *Proceedings of the National Academy of Sciences* 94(1):65-72.
39. Benoit M, Gabriel D, Gerisch G, & Gaub HE (2000) Discrete interactions in cell adhesion measured by single-molecule force spectroscopy. *Nat Cell Biol* 2(6):313-317.
40. Agrawal GP (2010) *Fiber-Optic Communication Systems* (Wiley) 4 Ed.
41. Cassen P, Guillot T, & Quirrenbach A (2006) *Extrasolar Planets* (Springer).
42. Kenney JF & Keeping ES (1962) *Mathematics of Statistics* (Van Nostrand) 3 Ed.
43. Savitzky A & Golay MJE (1964) Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry* 36(8):1627-1639.
44. Gonzalez RC & Woods RE (1992) *Digital Image Processing* (Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA) 2nd Ed.
45. Rief M, Gautel M, Oesterhelt F, Fernandez JM, & Gaub HE (1997) Reversible Unfolding of Individual Titin Immunoglobulin Domains by AFM. *Science* 276(5315):1109-1112.

46. Lang M, Guo H, Odegard JE, Burrus CS, & Wells RO (1995) Nonlinear processing of a shift invariant DWT for noise reduction. in *SPIE Symp. OE/Aerospace Sensing and Dual Use Photonics* (Orlando, FL/USA).
47. Donoho DL & Johnstone JM (1994) Ideal spatial adaptation by wavelet shrinkage. *Biometrika* 81(3):425-455.
48. Haar A (1910) Zur Theorie der orthogonalen Funktionensysteme. *Mathematische Annalen* 69:331-371.
49. Fowler JE (2005) The redundant discrete wavelet transform and additive noise. *IEEE Signal Processing Letters* 12:629-632.
50. Daubechies I (1992) *Ten lectures on wavelets* (Society for Industrial and Applied Mathematics (SIAM)).
51. Mallat S (1999) *A Wavelet Tour of Signal Processing, Second Edition (Wavelet Analysis & Its Applications)* (Academic Press).
52. Donoho DL & Johnstone IM (1995) Adapting to Unknown Smoothness via Wavelet Shrinkage. *Journal of the American Statistical Association* 90(432):1200-1224.
53. Tikkanen PE (1999) Nonlinear wavelet and wavelet packet denoising of electrocardiogram signal. *Biological Cybernetics* 80(4):259-267.
54. Rosenthal-Allieri MA, Ticchioni M, Breitmayer JP, Shimizu Y, & Bernard A (2005) Influence of β 1 Integrin Intracytoplasmic Domains in the Regulation of VLA-4-Mediated Adhesion of Human T Cells to VCAM-1 under Flow Conditions. *The Journal of Immunology* 175(2):1214-1223.
55. Schmitz J, Benoit M, & Gottschalk K-E (2008) The Viscoelasticity of Membrane Tethers and Its Importance for Cell Adhesion. *Biophysical Journal* 95(3):1448-1459.
56. Gonçalves MR & Marti O (2010) Influence of the roughness of metal templates on surface enhanced Raman scattering. in *Proceedings of SPIE*.
57. Gebhardt JCM, Bornschlöggl T, & Rief M (2010) Full distance-resolved folding energy landscape of one single protein molecule. *Proceedings of the National Academy of Sciences* 107(5):2013-2018.
58. Koster G, Cacciuto A, Derényi I, Frenkel D, & Dogterom M (2005) Force Barriers for Membrane Tube Formation. *Physical Review Letters* 94(6):068101.
59. Li Z, *et al.* (2002) Membrane Tether Formation from Outer Hair Cells with Optical Tweezers. *Biophysical Journal* 82(3):1386-1395.

60. Klaue D & Seidel R (2009) Torsional Stiffness of Single Superparamagnetic Microspheres in an External Magnetic Field. *Physical Review Letters* 102(2):028302.
61. Hinterdorfer P, Baumgartner W, Gruber HJ, Schilcher K, & Schindler H (1996) Detection and localization of individual antibody-antigen recognition events by atomic force microscopy. *Proceedings of the National Academy of Sciences* 93(8):3477-3481.
62. Hinterdorfer P & Dufrene YF (2006) Detection and localization of single molecular recognition events using atomic force microscopy. *Nature Methods* 3(5):347-355.
63. Bornschlöggl T & Rief M (2011) Single-molecule protein unfolding and refolding using atomic force microscopy. *Methods Mol Biol* 783:233-250.
64. Abbondanzieri EA, Greenleaf WJ, Shaevitz JW, Landick R, & Block SM (2005) Direct observation of base-pair stepping by RNA polymerase. *Nature* 438:460–465.
65. Mallik R, Carter BC, Lex SA, King SJ, & Gross SP (2004) Cytoplasmic dynein functions as a gear in response to load. *Nature* 427:649–652.
66. Svoboda K, Schmidt CF, Schnapp BJ, & Block SM (1993) Direct observation of kinesin stepping by optical trapping interferometry. *Nature* 365:721–727.
67. Neuman KC & Nagy A (2008) Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy. *Nat Meth* 5(6):491-505.
68. Kerssemakers JWJ, *et al.* (2006) Assembly dynamics of microtubules at molecular resolution. *Nature* 442(7103):709-712.
69. Carter BC, Vershinin M, & Gross SP (2008) A Comparison of Step-Detection Methods: How Well Can You Do? *Biophys J* 94(1):306-319.
70. Knight AE, Veigel C, Chambers C, & Molloy JE (2001) Analysis of single-molecule mechanical recordings: application to acto-myosin interactions. *Progress in Biophysics and Molecular Biology* 77(1):45-72.
71. Evans E & Ritchie K (1997) Dynamic strength of molecular adhesion bonds. *Biophys Journal* 72(4):1541-1555.
72. Fritz J, Katopodis AG, Kolbinger F, & Anselmetti D (1998) *Force-mediated kinetics of single P-selectin/ligand complexes observed by atomic force microscopy* pp 12283-12288.
73. Box G & Muller M (1958) A Note on the Generation of Random Normal Deviates. *The Annals of Mathematical Statistics* 29(2):610-611.
74. Matsumoto M & Nishimura T (1998) Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator. *ACM Trans. Model. Comput.*

Simul. 8(1):3-30.

75. Brunnbauer M, *et al.* (2010) Regulation of a heterodimeric kinesin-2 through an unprocessive motor domain that is turned processive by its partner. *Proceedings of the National Academy of Sciences*.
76. Shamri R, *et al.* (2005) Lymphocyte arrest requires instantaneous induction of an extended LFA-1 conformation mediated by endothelium-bound chemokines. *Nat. Immunol* 6:497–506.
77. Grabovsky V, *et al.* (2000) Subsecond Induction of α 4 Integrin Clustering by Immobilized Chemokines Stimulates Leukocyte Tethering and Rolling on Endothelial Vascular Cell Adhesion Molecule 1 under Flow Conditions. *The Journal of Experimental Medicine* 192(4):495-506.
78. Manevich E, Grabovsky V, Feigelson SW, & Alon R (2007) Talin 1 and paxillin facilitate distinct steps in rapid VLA-4-mediated adhesion strengthening to vascular cell adhesion molecule 1. *J Biol Chem* 282:25338-25348.
79. Romzek NC, *et al.* (1998) Use of a beta1 integrin-deficient human T cell to identify beta1 integrin cytoplasmic domain sequences critical for integrin function. *Mol Biol Cell* 9(10):2715-2727.
80. Lossner D, Abou-Ajram C, Bengel A, & Reuning U (2008) Integrin alpha(v) beta(3) mediates upregulation of epidermal growth-factor receptor expression and activity in human ovarian cancer cells. *The international journal of biochemistry & cell biology* 40:2746-2761.
81. Beck V, *et al.* (2005) ADAM15 decreases integrin alphavbeta3/vitronectin-mediated ovarian cancer cell adhesion and motility in an RGD-dependent fashion. *The international journal of biochemistry & cell biology* 37:590-603.
82. Bleul CC, Fuhlbrigge RC, Casasnovas JM, Aiuti A, & Springer TA (1996) A highly efficacious lymphocyte chemoattractant, stromal cell-derived factor 1 (SDF-1). (Translated from eng) *J Exp Med* 184(3):1101-1109 (in eng).
83. Sanz-Rodríguez F, Hidalgo A, & Teixidó J (2001) Chemokine stromal cell-derived factor-1 α modulates VLA-4 integrin-mediated multiple myeloma cell adhesion to CS-1/fibronectin and VCAM-1. *Blood* 97(2):346-351.
84. Alon R & Feigelson SW (2009) Chemokine Signaling to Lymphocyte Integrins Under Shear Flow. *Microcirculation* 16(1):3-16.
85. Alon R & Dustin ML (2007) Force as a Facilitator of Integrin Conformational Changes during

- Leukocyte Arrest on Blood Vessels and Antigen-Presenting Cells. *Immunity* 26(1):17-27.
86. Alon R & Shulman Z (2011) Chemokine triggered integrin activation and actin remodeling events guiding lymphocyte migration across vascular barriers. *Experimental Cell Research* 317(5):632-641.
 87. Grimmett G & Stirzaker D (1992) *Probability and Random Processes, 2nd ed.* (Oxford University Press, Oxford, England) 2nd Ed.
 88. Chigaev A, *et al.* (2009) Real-time Analysis of Conformation-sensitive Antibody Binding Provides New Insights into Integrin Conformational Regulation. *Journal of Biological Chemistry* 284(21):14337-14346.
 89. Puklin-Faucher E, Gao M, Schulten K, & Vogel V (2006) How the headpiece hinge angle is opened: New insights into the dynamics of integrin activation. *J Cell Biol* 175(2):349-360.
 90. Puklin-Faucher E & Vogel V (2009) Integrin activation dynamics between the RGD-binding site and the headpiece hinge. *J Biol Chem* 284(52):36557-36568.
 91. Icard-Arcizet D, Cardoso O, Richert A, & Hénon S (2008) Cell Stiffening in Response to External Stress is Correlated to Actin Recruitment. *Biophysical Journal* 94(7):2906-2913.
 92. Greene GW, Anderson TH, Zeng H, Zappone B, & Israelachvili JN (2009) Force amplification response of actin filaments under confined compression. *Proceedings of the National Academy of Sciences* 106(2):445-449.
 93. Voermans C, Anthony EC, Mul E, van der Schoot E, & Hordijk P (2001) SDF-1-induced actin polymerization and migration in human hematopoietic progenitor cells. *Experimental Hematology* 29(12):1456-1464.
 94. Pardo-Cabañas M, *et al.* (2004) Integrin $\alpha 4\beta 1$ involvement in stromal cell-derived factor-1 α -promoted myeloma cell transendothelial migration and adhesion: role of cAMP and the actin cytoskeleton in adhesion. *Experimental Cell Research* 294(2):571-580.
 95. Moyer RA, Wendt MK, Johanesen PA, Turner JR, & Dwinell MB (2007) Rho activation regulates CXCL12 chemokine stimulated actin rearrangement and restitution in model intestinal epithelia. *Lab Invest* 87(8):807-817.
 96. Murata K, *et al.* (2012) Stromal Cell-Derived Factor 1 Regulates the Actin Organization of Chondrocytes and Chondrocyte Hypertrophy. *PLoS ONE* 7(5):e37163.
 97. Rullo J, *et al.* (2012) Actin polymerization stabilizes $\alpha 4\beta 1$ integrin anchors that mediate monocyte adhesion. *The Journal of Cell Biology* 197(1):115-129.
 98. Moser M, Legate KR, Zent R, & Fässler R (2009) The Tail of Integrins, Talin, and Kindlins.

Science 324(5929):895-899.

99. Franco-Cea A, *et al.* (2010) Distinct developmental roles for direct and indirect talin-mediated linkage to actin. *Developmental Biology* 345(1):64-77.
100. Tanentzapf G, Martin-Bermudo MD, Hicks MS, & Brown NH (2006) Multiple factors contribute to integrin-talin interactions in vivo. *Journal of cell science* 119(8):1632-1644.
101. Calderwood DA, Shattil SJ, & Ginsberg MH (2000) Integrins and Actin Filaments: Reciprocal Regulation of Cell Adhesion and Signaling. *Journal of Biological Chemistry* 275(30):22607-22610.
102. Nieves B, *et al.* (2010) The NPIY motif in the integrin β 1 tail dictates the requirement for talin-1 in outside-in signaling. *Journal of cell science* 123(8):1216-1226.
103. Hyduk SJ, *et al.* (2011) Talin-1 and Kindlin-3 Regulate α 4 β 1 Integrin-Mediated Adhesion Stabilization, but Not G Protein-Coupled Receptor-Induced Affinity Upregulation. *The Journal of Immunology* 187(8):4360-4368.
104. Gingras AR, *et al.* (2009) Structural Determinants of Integrin Binding to the Talin Rod. *Journal of Biological Chemistry* 284(13):8866-8876.
105. Ellis SJ, Pines M, Fairchild MJ, & Tanentzapf G (2011) In vivo functional analysis reveals specific roles for the integrin-binding sites of talin. *Journal of cell science* 124(11):1844-1856.
106. Müller MA, *et al.* (2012) Integrin-mediated adhesion and signaling are decoupled and controlled by their transmembrane domains (manuscript in preparation).
107. Gottschalk KE, Adams PD, Brunger AT, & Kessler H (2002) Transmembrane signal transduction of the α (IIb) β (3) integrin. *Protein Science* 11(7):1800-1812.
108. Hoefling M, Kessler H, & Gottschalk KE (2009) The Transmembrane Structure of Integrin α (IIb) β (3): Significance for Signal Transduction. *Angewandte Chemie* 48:6590-6593.
109. Lau TL, Kim C, Ginsberg MH, & Ulmer TS (2009) The structure of the integrin α (IIb) β (3) transmembrane complex explains integrin transmembrane signalling. *The EMBO journal* 28:1351-1361.
110. Hantgan RR, Paumi C, Rocco M, & Weisel JW (1999) Effects of ligand-mimetic peptides Arg-Gly-Asp-X (X = Phe, Trp, Ser) on α (IIb) β (3) integrin conformation and oligomerization. *Biochemistry* 38:14461-14474.
111. Xiao T, Takagi J, Collier BS, Wang JH, & Springer TA (2004) Structural basis for allostery in integrins and binding to fibrinogen-mimetic therapeutics. *Nature* 432:59-67.

112. Gottschalk KE (2005) A coiled-coil structure of the alpha (IIb) beta(3) integrin transmembrane and cytoplasmic domains in its resting state. *Structure* 13:703-712.
113. Senes A, Engel DE, & DeGrado WF (2004) Folding of helical membrane proteins: the role of polar, GxxxG-like and proline motifs. *Current opinion in structural biology* 14:465-479.
114. Lemmon MA, Flanagan JM, Treutlein HR, Zhang J, & Engelman DM (1992) Sequence specificity in the dimerization of transmembrane alpha-helices. *Biochemistry* 31:12719-12725.
115. Vinogradova O, *et al.* (2004) Membrane-mediated structural transitions at the cytoplasmic face during integrin activation. *PNAS* 101:4094-4099.
116. Hapke S, *et al.* (2001) Integrin alpha(v)beta(3)/vitronectin interaction affects expression of the urokinase system in human ovarian cancer cells. *The Journal of biological chemistry* 276:26340-26348.
117. Helenius J, Heisenberg C-P, Gaub HE, & Müller DJ (2008) Single-cell force spectroscopy. *J Cell Sci* 121(11):1785-1791.
118. Zhu J, *et al.* (2009) The structure of a receptor with two associating transmembrane domains on the cell surface: integrin alpha(IIb) beta(3). *Molecular cell* 34:234-249.
119. Bell GI (1978) Models for the specific adhesion of cells to cells. *Science* 200:618-627.

Publications

J. Opfer, K.-E. Gottschalk

A wavelet-based recursive noise reduction algorithm

submitted to PNAS

J. Opfer, K.-E. Gottschalk

Identifying discrete states of a biological system using a novel step detection algorithm

submitted to PLoS ONE

J. Opfer, J. Schmitz, E. Manewitch, M. Benoit, R. Alon, and K.-E. Gottschalk

SDF-1 α alters the mechanical properties of the integrin $\alpha_4\beta_1$ environment

manuscript in preparation

M. A. Müller, J. Opfer, L. A. Volkhardt, L. Brunie, E.-K. Sinner, D. Boettiger, A. Bochen, H. Kessler, K.-E. Gottschalk, and U. Reuning

Integrin-mediated adhesion and signaling are decoupled and controlled by their trans-membrane domains

manuscript in preparation

E. Sariisik, D. Docheva, D. Padula, J. Opfer, M. Schieker, H. Clausen-Schaumann, and M. Benoit

Probing the Interaction Forces of Prostate Cancer Cells with Collagen I and Bone Marrow Derived Stem Cells on the Single Cell Level with AFM Based Force Spectroscopy

manuscript in preparation

Acknowledgment

I would like to thank everybody who contributed to this work or supported me. Prof. Dr. Kay-E. Gottschalk supervised the thesis, Prof. Dr. Hermann E. Gaub provided the possibility to work at his chair, Prof. Dr. Joachim Rädler refereed the thesis, and Dr. Martin Benoit co-advised the projects. I am pleased for many fruitful discussions with Dr. Martin Höfling, Dr. Jan Neumann, Stefan Stahl, Stefan Scheuer, Dr. Sandra Baumann, Lorenz Rognoni, and Dr. Julia Schmitz. Angelika Kardinal and Thomas Nikolaus helped me a lot in the lab. The people at the chair made me enjoy an excellent working climate. Further, I thank Dr. Martina Müller, Prof. Dr. Ute Reuning, Tobias Petzold, Raphael Ruppert, Dr. Markus Moser, and Prof. Dr. Reinhard Fässler for pleasant and productive cooperations, and Dr. Marie-Christine Blüm and Marilena Pinto for their ambitious efforts far beyond working-to-rule. The Elite Network of Bavaria (ENB) financially supported the thesis with a scholarship for the International Doctorate Program NanoBioTechnology (IDK-NBT) at the Center for Nanoscience (CeNS). Last but not least, I highly appreciate the constant support by my family.