Forced unfolding of the fibronectin type III module reveals a tensile molecular recognition switch

ANDRÉ KRAMMER*, HUI LU†, BARRY ISRALEWITZ‡, KLAUS SCHULTEN†‡, AND VIOLA VOGEL*‡§

Departments of *Physics and 3Bioengineering, University of Washington, Seattle, WA 98195; and 5Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL 61801

Communicated by Jiri Jonas, University of Illinois at Urbana-Champaign, Urbana, IL, December 14, 1998 (received for review September 25, 1998)

ABSTRACT The 10th type III module of fibronectin possesses a β-sandwich structure consisting of seven β-strands (A–G) that are arranged in two antiparallel sheets. It mediates cell adhesion to surfaces via its integrin binding motif, Argβ8, Glyβ9, and Aspβ10 (RGD), which is placed at the apex of the loop connecting β-strands F and G. Steered molecular dynamics simulations in which tension is applied to the protein’s terminal ends reveal that the β-strand G is the first to break away from the module on forced unfolding whereas the remaining fold maintains its structural integrity. The separation of strand G from the remaining fold results in a gradual shortening of the distance between the apex of the RGD-containing loop and the module surface, which potentially reduces the loop’s accessibility to surface-bound integrins. The shortening is followed by a straightening of the RGD-loop from a tight β-turn into a linear conformation, which suggests a further decrease of affinity and selectivity to integrins. The RGD-loop therefore is located strategically to undergo strong conformational changes in the early stretching stages of the module and thus constitutes a mechanosensitive control of ligand recognition.

Cellular response to mechanical force has emerged to be a critical regulator for sensory functions as well as for cell signal transduction, proliferation, and gene expression (ref. 1 and references therein). The mechanical tension generated within the contractile microfilaments of the cytoskeleton pull inward on the cell membrane. These inward-directed forces are counterbalanced by cell adhesion to anchoring sites within the extracellular environment. Transmembrane integrins thereby act as structural links to couple the intracellular microfilaments to extracellular ligands—for example, to RGD-containing proteins of the extracellular matrix (ECM). Current research efforts seek to characterize the “inside-out” signaling pathway by which the adhesiveness of integrins to external ligands is regulated by intracellular events affecting the cytoplasmic part of integrins. Simultaneously, “outside-in” signaling pathways exist that are triggered by binding of external ligands to membrane-bound integrins, leading to the assembly of integrins and cytoplasmic proteins into large focal adhesion complexes. Models describing both pathways implicitly assume that the RGD-containing proteins, either surface-bound or integrated into the ECM, serve as static anchoring sites against which cells can build up tensile strength. As a result, the dynamic regulation of cell adhesiveness is attributed solely to activation processes involving integrins. Simulations of forced unfolding of the 10th type III module of fibronectin (FnIII10) that are reported below now suggest that FnIII10 itself acts as a mechanosensitive switch. Accessibility of its RGD-loop to integrins is reduced when the FnIII10 module is stretched beyond an extension threshold value. This may explain how cells retract from RGD-containing matrices, a process that is key to understanding cell motility. The speed of cell migration, for example, is limited by the retraction rate of the rear of the cell from adhesive substrates, which requires disruption of all bonds between the cell membrane and the ECM (2).

The functional role of the RGD sequence to bind to integrins was first identified for fibronectin (3). Fibronectin is composed of three repeating structural motifs, of which one is the FnIII module. The three modules form a linear sequence of multiple tandem copies connected by short linker peptides as shown in Fig. 1a. The secondary structure of the FnIII10 module, which is the only fibronectin module to possess the RGD motif, consists of two β-sheets containing the antiparallel β-strands ABE and DCFG, respectively, which fold up to form a β-sandwich (Fig. 1b). The RGD sequence is located in the loop connecting the β-strands F and G. Fibronectin is a glycoprotein of 450–500 kDa with multiple recognition sites, thereby constituting a major building block of the extracellular matrix. In addition to cell recognition, essential for fibronectin’s function as an ECM protein is its ability to self-assemble into fibrils and to exhibit binding sites to other ECM proteins, including collagen, fibrin, heparin, and fibrinogen (4). The fibronectin type III module also is encountered in many other proteins, including tenasin, as well as the neural cell adhesion protein neuroglian, the muscle protein titin, and in a variety of different cytokine receptors (5, 6).

Forced unfolding of single proteins containing modular repeats recently has been studied experimentally for titin (7–10) and tenasin (11) by using atomic force microscopy (AFM) and optical tweezers. In all cases, pulling on the entire protein resulted in the consecutive rupture of the tertiary structure of individual modules. Each peak in AFM force-extension curves could be correlated to the rupture of a single protein module. Our previous studies of the forced unfolding of the 27th Ig module belonging to the I-band of titin (I27) by steered molecular dynamics (SMD) have reproduced the occurrence of a single peak in a computed force–extension graph (12) in agreement with experimental results. SMD simulations furthermore have been applied successfully to analyze ligand-receptor unbinding reactions (13–17). SMD simulations possess the advantage of providing an atomic scale picture of the unraveling process and can reveal how the folding scaffold of a protein responds to external forces. The simulations of the stretching and unfolding of the FnIII10 module illustrate how the accessibility of a recognition site can be modulated by external forces, opening the door to future

Abbreviations: FnIII, fibronectin type III; I27, 27th Ig module of titin’s I-band; SMD, steered molecular dynamics; ECM, extracellular matrix; AFM, atomic force microscopy.

†To whom reprint requests should be addressed. K.S., Beckman Institute for Advanced Science and Technology, University of Illinois, 405 North Mathews Avenue, Urbana, IL 61801, e-mail: kschulte@ks.uiuc.edu. Or: V.V., Center for Nanotechnology and Department of Bioengineering, Box 352125, University of Washington, Seattle, WA 98195, e-mail: vogel@bioeng.washington.edu.
studies of the role of tensile forces in regulating biological activity.

METHODS

Molecular dynamics simulations were carried out by using the programs XPLOR (18) and NAMD (19) with the CHARMM (20) force field. The x-ray crystallographic structure of FnIII10 was obtained from the tetramer FnIII7–10 (Protein Data Bank code 1fnf) (21). All atoms, including hydrogens, are described explicitly. The simulations were performed with a time step of 1 fs, TIP3P water parameters (22), a uniform dielectric constant of 1, and a cut-off of Coulomb forces with a switching function starting at a distance of 10 Å and reaching zero at 13 Å. FnIII10 was solvated by a sphere of water covering the module by at least five layers of water molecules before the entire system gradually was heated up to 300 K and equilibrated at a temperature of 300 K. A detailed description of the equilibration procedure is given in ref. 12. Equilibrating for 50 ps before SMD simulations yielded a stable solvated structure of FnIII10 with temperature fluctuations of 5 K and an averaged rms coordinate deviation of 1.5 Å in reference to the crystal structure. In SMD simulations, a time-dependent harmonic restraint directed along a specified coordinate then was imposed on the simulated molecular system, resulting in a nonequilibrium simulation. In the simulations, the N-terminal Cα atom (Leu1) of FnIII10 was constrained to a fixed point whereas the C-terminal Cα atom (Thr94) was attached to one

Fig. 1. (a) Modular structure of the fibronectin family. Fibronectin consists of two monomeric strands linked by two disulfide bridges (S—S). Each monomer contains three types of modules, types I, II, and III, and/or variable modules V. Modules that are either inserted or missing in various spliced forms of fibronectin are shown as filled circles. The cell adhesion site and other recognition sites are indicated. (b–e) Forced unfolding sequence of FnIII10 for various extensions of the module. A different gray scale is used to highlight the RGD motif and strand G. Hydrogen bond pairs that are broken in clusters during the unfolding burst of the module are displayed by dotted (AB bond pairs) and dashed (FG bond pairs) lines. The ball-and-stick representation indicates the position of the RGD cell binding region along the FG loop of FnIII10. Water molecules that were included in the SMD simulation are omitted for clarity. The equilibrated structure of FnIII10 is shown in b with its fixed terminal end (×) and the applied external pulling force (gray arrow). Partial distortion of the module is evident in c with the hydrogen bonds between strand F and G being stretched. Finally, in d, strand G clearly is separated from β-sheet FCD with the hydrogen bonds between F and G broken apart. Only strands C, D, and E remain in e to form a structural unit whereas the other strands form a line along the pulling direction. Further pulling on the module leads to total disintegration of the module with all strands aligned. This figure was created with MOLSCRIPT (37).
end of a harmonic spring with a force constant $k$ of $10 k_B T / \AA^2$, limiting the spatial fluctuation of its displacement to $\delta z = \sqrt{(k_B T / k)} = 0.32 \AA$. The other end of the spring was moved at a nearly constant pulling speed $v$ of 0.5 $\AA$/ps, which was achieved by changing its position every 100 fs by a distance of 0.05 $\AA$. The direction of the velocity vector $v$ was chosen to point from the constrained C$_a$ atom to the moving C$_a$ atom. Thus, the two termini of the protein module were subjected to an external force $F$ of the form:

$$F = k(vt - z(t)),$$

where $z(t)$ is the displacement vector at time $t$ of the C-terminal C$_a$ atom. The unfolding events are independent of the pulling direction. SMD simulations on 127 (12) yielded similar results either when fixing one of the terminal C$_a$ atoms and pulling on the other one or when pulling on both termini with half of the pulling speed. The extension of the module at any time $t$ during the simulation is defined by the distance $z(t)$. The simulation mimics an AFM experiment in which the cantilever represented by the harmonic spring is moving at constant speed with its tip attached to one end of the protein module and the other bound to a substrate.

**RESULTS**

**Forced Unfolding of the FnIII$_{10}$ Module.** FnIII$_{10}$ was stretched from its initially compact and folded structure to its fully elongated configuration. A sequence of forced unfolding events is shown in Fig. 1 b–e. Details of the forced unfolding pathway can be correlated to characteristics of the force–extension curve. As seen in Fig. 2a, a force threshold of 1,808 pN must be overcome to rupture the tertiary structure of FnIII$_{10}$, occurring at an extension of 24.6 Å. Two minor features were observed in close vicinity at an extension of 17.3 and 36.2 Å, with force peaks of 1,556 and 1,572 pN, respectively. The pulling force then steadily decreased until it leveled out to an average of $\approx$500 pN while the ruptured module was unraveling further. The module was fully elongated at an extension of 315 Å, at which time the force again rapidly increased and the simulation was halted. Simulations also were carried out for FnIII$_{9}$ as well as at a pulling speed of 1.0 Å/ps, providing similar results to the ones discussed below.

The observation in SMD simulations that a force threshold must be overcome to unravel the module agrees well with experimental force–extension curves from the FnIII modules of tenascin (11). The computed force peak of 1,800 pN, however, is nearly an order of magnitude higher than the 200 pN force measured by AFM. This discrepancy results from differences in pulling speeds and spring constants: a pulling speed of 0.5 Å/ps had to be chosen for the SMD simulation because of the limitation in simulation time whereas the pulling speed of typical AFM experiments is several orders of magnitude smaller (23). However, reconstructing the potential of mean force along the SMD pulling coordinate allows one to estimate the force needed in AFM experiments. This method has been used successfully in the case of a one-dimensional potential, relating forces at different pulling distances to the position of the hydrogen and oxygen atoms. Hydrogen bond pairs with residue and participant backbone atoms between $\beta$-strand F and G are (line pattern and color are solid and black unless otherwise noted in parenthesis) Tyr$_{68}$–Tyr$_{17}$, Ile$_{70}$–Ile$_{72}$ (dashed), Val$_{17}$–Ile$_{19}$ (dot-dashed), Ile$_{68}$–Val$_{72}$ (long dashed), Ile$_{21}$–Ile$_{11}$ (dashed, gray), and Tyr$_{68}$–Tyr$_{88}$ (gray). Hydrogen bond pairs between $\beta$-strand A and B are Ser$_{12}$–Glu$_{40}$ (gray), Glu$_{40}$–Ser$_{12}$, Leu$_{59}$–Val$_{72}$ (dashed), Ala$_{12}$–Leu$_{11}$ (dot-dashed), and Ser$_{12}$–Thr$_{11}$ (dashed, gray). Of backbone hydrogen bonds was correlated with the force peak and the first minor feature seen in the force–extension diagram (Fig. 2a). Coinciding with strand G sliding in the direction of the applied force, the 12-aa loop FG that contains the RGD sequence was dragged along (Fig. 1d) whereas strand F stayed closely bound to strand C. Distinct from this first rupture event that broke the backbone hydrogen bonds between strands G and F collectively, a second set of five backbone hydrogen bonds connecting strands A and B was distorted starting at an extension of 36 Å, as seen in Fig. 2d and e. These hydrogen bonds did not, however, rupture collectively. Three broke apart at 40 Å followed by the rupture of the other two at $\approx$55 Å. At this point, strand A separated from strand B. After this last rupture, the module gradually unfolded, and the remaining strands, B through F, unraveled one by one. The hydrogen bond pairs between strands A and B, as well as F and G, were the only ones observed to break.
apart in clusters whereas all other inter-β-sheet backbone hydrogen bonds later broke individually during the extension of FnIII10. This behavior is expected to be independent of the amino acid sequence of those strands as long as mutations do not perturb the module’s tertiary structure.

The force required for further extension of the module was now mainly attributable to friction with the surrounding water molecules and surface tension of the water droplet used in the simulation. Finally, at full extension, all of the β-strands were aligned linearly (Fig. 1e). Further stretching beyond this point required changing the length and angle of covalent bonds along the module backbone, which resulted in the drastic increase in pulling force (Fig. 2a).

**RGD-Loop Conformations Along the Unfolding Pathway.**
Dramatic conformational changes of the RGD-loop were observed in the initial stages of the forced unfolding pathway of the module (Fig. 1b–d). The loop initially was located at the apex of a hairpin β-turn connecting strands G and F. As illustrated in Fig. 1b as well as schematically in Fig. 3, this hairpin was bent at the Cα atoms of residues Thr76 and Ala83. On extension of the module past the rupture between strands G and F, this bend was straightened, as can be seen in the steep increase in the angle defined by the Cα atoms of residues Asp80, Ala83, and Ser84 at 20 Å (Fig. 3a). Subsequently, the apex of the loop was pulled closer to the surface of the module, i.e., to the plane defined by β-strands C and F. It was shortened from 12 to 8 Å whereas the module was extended from 25 to 35 Å (Fig. 3b). After this transformation, the RGD-loop opened up, which is indicated by a gradual change in the distance between the Cα atoms of residues Arg78 and Pro82 starting at 30 Å (Fig. 3c). Residue Asp80 of the RGD sequence initially defined the apex of the RGD-loop at extensions ~31.4 Å. As the loop was dragged behind the G-strand, the apex of the loop moved closer to the F-strand. This resulted in a transition of the (φ,
Simulations illustrate how the RGD-loop loses its tight hairpin-like configuration if FnIII\textsubscript{10} is stretched beyond the rupture of strands G and F. The SMD simulations, combined with experimental evidence, suggest that the accessibility and affinity of the RGD sequence for membrane-bound integrins can be altered if tension is applied to the terminal ends of FnIII\textsubscript{10}. The FnIII\textsubscript{10} module thereby acts as a tensile molecular recognition switch that is activated under applied external force. On mechanical activation, the module’s affinity to membrane-bound integrins is reduced.

Reversible unfolding after activation, however, is crucial to a tensile molecular recognition switch that has undergone a function-driven evolutionary process. The FnIII\textsubscript{10} module is known to refold on a time scale of milliseconds after complete denaturization, which is considerably faster than the refolding of other FnIII modules (34, 35). Rapid refolding to the initial conformation is facilitated if only one β-strand needs to be pulled out to activate the switch while the remaining module experiences only minor structural perturbations. This brings up the question regarding how the β-sandwich motif of FnIII\textsubscript{10} functionally differs from other β-sandwich motifs and why the RGD sequence is spliced predominately into FnIII\textsubscript{10} rather than into other β-sandwich modules. SMD simulations forcing the unfolding of the C2 module of synaptotagmin I, a β-sandwich motif with the N- and C-termini in close proximity, and of cytochrome C6, an all α-helical protein, displayed no force peak during unfolding (36). This indicates that no structural element of these modules is protected from being readily unfolded by external force. These types of modules would not be suitable to include the RGD motif as a tensile molecular recognition switch because the loop would be disturbed too easily by applied forces.

A closer comparison of the two unrelated yet similar β-sandwich motifs, titin I27 and fibronectin FnIII\textsubscript{10}, which both show a single force peak early in the unfolding pathway, experimentally as well as in SMD simulations, provides even further insight. The seven β-strands of I27 also fold into a β-sandwich in which the upper and lower β-sheets are formed by β-strands ABED and A’G’F’, respectively. Nevertheless, the backbone topology of I27 is slightly different from the one of FnIII\textsubscript{10} (compare Fig. 1 c and Fig. 5). For instance, β-strands A and A’ of I27, which are equivalent to β-strand A of FnIII\textsubscript{10}, belong to different β-sheets. Backbone hydrogen bonds link β-strands A and B, as well as A’ and G in I27, whereas FnIII\textsubscript{10} lacks hydrogen bonds between β-strands A and G. Previous SMD simulations of the forced unfolding of I27 under the same conditions as the ones described here have demonstrated that the I27 structure unraveled by sequentially breaking the hydrogen bonds between β-strands AB, A’G’, and GF (12). Although, for FnIII\textsubscript{10}, the C-terminal β-strand G was pulled out first, in the case of I27, the N-terminal β-strands A and A’ were released first on applying tension to the module’s termini. The major advantage of functionalizing the GF-loop of FnIII\textsubscript{10}
thus could be that the G-strand forms backbone hydrogen bonds to only one neighboring strand (F) whereas the G-strand of I27 forms backbone hydrogen bonds to two neighboring strands (A′ and F). On activation of the recognition switch, the structural perturbation of the FnIII10 module therefore was minimized. This illustrates how the study of forced unfolding pathways may provide important clues for an understanding of the structure/function relations of this module.

In summary, SMD simulations suggest that FnIII10 operates as a tensile molecular recognition switch if the module is extended beyond a threshold value to rupture the hydrogen bonds between strands F and G. The existence of a force threshold ensures the structural stability for the RGD motif to function properly unless the tensile force exceeds this threshold. The FnIII10 module, thereby modulating cell binding. FnIII10 thus constitutes a critical link in transmitting mechanical stress from various extracellular matrix proteins. Tensile stress from the unfolding of FnIII10. The FnIII10 module within fibronectin, regulated if a tensile force is applied that initiates partial structural perturbation of the FnIII10 module therefore was straightened out early in the forced unfolding process. The conformation of the RGD-loop thus is altered while introducing only minor structural disturbances to the remaining module. Identification of the FnIII10 module as a tensile molecular recognition switch by SMD simulations has important implications in cell biology and biotechnology. Experiments now can be designed to test whether cell binding to FnIII10 can be regulated if a tensile force is applied that initiates partial unfolding of FnIII10. The FnIII10 module within fibronectin, for example, is tightly linked to other repeats that recognize various extracellular matrix proteins. Tensile stress from the extracellular matrix potentially can be transmitted to the FnIII10 module, thereby modulating cell binding. FnIII10 thus constitutes a critical link in transmitting mechanical stress from the extracellular matrix to the cytoskeleton and vice versa by functioning not only as an anchor but also as a mechanosensitive regulator.

We thank G. Banexy and Dr. T. P. Lybrand for fruitful discussions. This work was supported by grants from the National Institute of Health (Grants GM 40863 and PHS 5 P41 RR05969), by the National Science Foundation (Grants BIR 94-23827 EQ, NSF/GCAG BIR 93-18159, and MCA93S028), by the Roy J. Carver Charitable Trust, and the Whitaker Foundation.