⌘ *Author's Choice*

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 284, NO. 23, pp. 15668–15675, June 5, 2009

Motogenic Sites in Human Fibronectin Are Masked by Long Range Interactions[S]

Received for publication, February 5, 2009, and in revised form, April 3, 2009 Published, JBC Papers in Press, April 14, 2009, DOI 10.1074/jbc.M109.003673

Ioannis Vakonakis[‡1], David Staunton[‡2], Ian R. Ellis[§3], Peter Sarkies[‡], Aleksandra Flanagan[‡2], Ana M. Schor[§3], Seth L. Schor[§3], and Iain D. Campbell[‡2]

From the [‡]Department of Biochemistry, University of Oxford, Oxford OX1 3QU and the [§]Unit of Cell and Molecular Biology, University of Dundee, Dundee DD1 4HR, Scotland, United Kingdom

Fibronectin (FN) is a large extracellular matrix glycoprotein important for development and wound healing in vertebrates. Recent work has focused on the ability of FN fragments and embryonic or tumorigenic splicing variants to stimulate fibroblast migration into collagen gels. This activity has been localized to specific sites and is not exhibited by full-length FN. Here we show that an N-terminal FN fragment, spanning the migration stimulation sites and including the first three type III FN domains, also lacks this activity. A screen for interdomain interactions by solution-state NMR spectroscopy revealed specific contacts between the Fn N terminus and two of the type III domains. A single amino acid substitution, R222A, disrupts the strongest interaction, between domains $^{4-5}$FnI and ^{3}FnIII, and restores motogenic activity to the FN N-terminal fragment. Anastellin, which promotes fibril formation, destabilizes ^{3}FnIII and disrupts the observed $^{4-5}$FnI-^{3}FnIII interaction. We discuss these findings in the context of the control of cellular activity through exposure of masked sites.

Fibronectin (FN),[4] a large multidomain glycoprotein found in all vertebrates, plays a vital role in cell adhesion, tissue development, and wound healing (1). It exists in soluble form in plasma and tissue fluids but is also present in fibrillar networks as part of the extracellular matrix. The structures of many FN domains of all three types, FnI, FnII, and FnIII, are known, for example (2–4). Although interactions between domains that are close in primary sequence have been demonstrated (3, 5), studies of multidomain fragments generally assume a beads-on-string model (2). There is, however, much evidence for the presence of long range order in soluble FN as a number of functional sites, termed cryptic, are not active in the native molecule, until exposed through conformational change. These include self-association sites (5–8), sites of protein interactions (9), and sites that control cellular activity (10, 11). Low resolution studies of the FN dimer suggest a compact conformation under physiological conditions (12–14); however, attempts to define large scale structure in FN by small angle scattering or electric birefringence (15–17) have yielded contradictory results. Interpretation of domain stability changes in terms of interaction sites (18) has also not been straightforward (2), possibly because of domain stabilization through nearest-neighbor effects (19, 20).

A FN splicing variant produced in fetal and cancer patient fibroblasts, termed migration stimulation factor (MSF), stimulates migration of adult skin fibroblasts into type I collagen gels (10, 21) and breast carcinoma cells using the Boyden chamber (22). MSF comprises FN domains ^{1}FnI to ^{9}FnI, a truncated ^{1}FnIII, and a small C-terminal extension; a recombinant FN fragment corresponding to ^{1}FnI-^{9}FnI (Fn70kDa) displays the same activity (10). An overview of FN domain structure and nomenclature is presented in Fig. 1*a*. Further experiments sublocalized full motogenic activity to the gelatin binding domain of FN (GBD, domains ^{6}FnI-^{9}FnI) (23) and partial activity to a shorter fragment spanning domains $^{7-9}$FnI (24). Two IGD tripeptides of domains ^{7}FnI and ^{9}FnI were shown to be essential through residue substitutions and reconstitution of partial motogenic activity in synthetic peptides (10, 24, 25); however, similar IGD tripeptides outside the GBD, on domains ^{3}FnI and ^{5}FnI, appear to have little effect (10, 23). Full-length adult FN does not affect cell migration in similar assays (10, 23); thus motogenic activity sites are presumed to be masked in the conformation adopted by soluble FN, although they could be exposed by molecular rearrangement.

Here we show that a recombinant fragment, closely matching a truncated form of FN identified in zebrafish (26), as well as amphibians, birds, and mammals (27), does not stimulate cell migration. This fragment is similar to MSF but includes the first three FnIII domains ($^{1-3}$FnIII), suggesting that these domains are responsible for a conformational transition that masks the activity sites in this construct and probably in full-length FN. To identify the mechanism behind this transition, we performed structural studies by solution NMR spectroscopy and identified a specific long range interaction between domains $^{4-5}$FnI and ^{3}FnIII as essential for this masking effect. Interestingly, this interaction does not involve direct contacts with the GBD but possibly represses motogenic activity through chain compaction, evident in analytical size exclusion assays. Intramolecular interactions thus provide a mechanism by which confor-

⌘ *Author's Choice*—Final version full access.

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Figs. 1–8.

[1] Recipient of support from the Marie Curie Fellowships program, the Royal Society, and Trinity College, Oxford, UK. To whom correspondence should be addressed: Dept. of Biochemistry, South Parks Rd., Oxford OX1 3QU, UK. Tel.: 44-1865-285332; Fax: 44-1865-275253; E-mail: ioannis.vakonakis@bioch.ox.ac.uk.

[2] Supported by the Wellcome Trust.

[3] Supported by the Breast Cancer Campaign.

[4] The abbreviations used are: FN, fibronectin; MSF, migration stimulation factor; FnI/II/III, FN type I/II/III domains; Fn30kDa, fibronectin domains ^{1}FnI-^{5}FnI; Fn70kDa, fibronectin domains ^{1}FnI-^{9}FnI; Fn100kDa, fibronectin domains ^{1}FnI-^{3}FnIII; GBD, fibronectin gelatin binding domain; HSQC, heteronuclear single quantum coherence.

mational rearrangement induced, for example, by tension or splicing variation can result in cellular activity differences.

EXPERIMENTAL PROCEDURES

Protein Expression and Purification—Expression and purification of type I domain pairs of FN ($^{2-3}$FnI, FN residues 93–182; $^{4-5}$FnI, residues 183–275; and $^{8-9}$FnI, residues 516–608) were described previously (24, 28). ^{6}FnI$^{1-2}$FnII^{7}FnI (residues 305–515) and GBD (residues 305–608) were produced as described (29). Expression and purification of ^{1}FnIII and ^{2}FnIII have been described previously (5). Domain ^{3}FnIII (residues 810–900) and $^{2-3}$FnIII were produced in a similar manner. In all cases, protein purity was evaluated by SDS-PAGE and NMR as being in excess of 95%.

$^{2-5}$FnI was integrated in *Pichia pastoris* in a manner analogous to that described previously (30). Expression in media buffered at pH 3.5 at 25 °C produced small amounts of protein and a number of proteolytic degradation products. $^{2-5}$FnI was concentrated from the media by retention in a cation exchange column, and protein-containing fractions were pooled and purified by size exclusion chromatography. SDS-PAGE under reducing and nonreducing conditions produced the expected protein band essentially free of degradation by-products (supplemental Fig. 6). ^{1}H-^{15}N HSQC spectra of this protein showed substantial heterogeneity and significant presence of an unfolded population of molecules (supplemental Fig. 6). The different species were separated using a shallow NaCl gradient in a high resolution anion exchange column (MonoQ, GE Healthcare) at pH 10.6 (supplemental Fig. 6). Approximately 0.2 mg of correctly folded purified protein could be obtained from 0.5 liters of *P. pastoris* culture under high density fermentation conditions.

Recombinant MSF was produced in *Escherichia coli*, as described previously (10), from a transcript that included the complete N-terminal FN sequence to the middle of ^{1}FnIII (residue 647) and a unique C-terminal amino acid tail. FN fragments consisting of residues 48–608 (Fn70kDa) or 48–900 (Fn100kDa, wild-type or with residue substitutions) were cloned into vector pHLsec resulting in final proteins with three vector-derived residues at the N terminus (ETG) and nine, including a His tag, at the C terminus (GTKHHHHHH). These proteins were produced in human cell line HEK293T by transient expression as described previously (31). Proteins were expressed for 4 days and then purified directly from the media by metal affinity chromatography followed by size exclusion in phosphate-buffered saline. A typical yield of purified Fn100kDa was 20 mg per liter of conditioned media. Protein purity was evaluated by SDS-PAGE as 85–90%. The proteolytic Fn30kDa fragment was purchased from Sigma.

Analytical Ultracentrifugation Equilibrium—The molecular mass of Fn100kDa was determined by sedimentation equilibrium analysis using an Optima XL-A analytical ultracentrifuge (Beckman). Fn100kDa samples (5 and 11 μM) in phosphate-buffered saline were centrifuged in double sector 12-mm centerpieces at 10,000, 12,000, and 16,000 rpm at 20 °C. Protein sedimentation was monitored at 280 nm. Although expected to be glycosylated at three separate sites, Fn100kDa gave a molecular mass of ~95 kDa suggesting little glycosylation of the final protein.

Analytical Size Exclusion Chromatography—Small (0.1 ml) samples of FN fragments were passed through an analytical Superdex 200 size exclusion column (GE Healthcare) equilibrated in phosphate-buffered saline. An Äkta FPLC system with 0.5-cm path length (GE Healthcare) was used for data recording of UV absorption at 280 nm.

Collagen Gel Migration Assays—Type I collagen was extracted from rat tail tendons and used to make 2-ml collagen gels in 35-mm plastic tissue culture dishes as described previously (32). Collagen gels were overlaid with 1 ml of either serum-free minimum Eagle's medium or serum-free minimum Eagle's medium containing four times the final concentration of recombinant proteins. Confluent stock cultures of fibroblasts were trypsinized, pelleted by centrifugation, and resuspended in growth medium containing 4% donor calf serum at 2×10^5 cells/ml, and 1-ml aliquots were added to the overlaid gels for a final concentration of 1% serum in both control and recombinant proteins containing cultures. The assay cultures were incubated for 4 days, and the percentage of fibroblasts found within the three-dimensional gel matrix was then ascertained by microscopic observation of 10 randomly selected fields in each of two duplicate cultures, as described previously (10, 24). Each experiment was repeated a minimum of two times.

NMR Spectroscopy and Data Analysis—All experiments were performed at 30 °C using home-built spectrometers with 11.7–22.3 tesla field strengths in a 20 mM sodium phosphate, pH 6.0, 20 mM NaCl, 2 mM EDTA, 0.1 mM 2,2-dimethyl-2-silapentanesulfonic acid, 0.02% NaN_3, and 5% v/v D_2O sample buffer unless otherwise noted. Sequential chemical shift assignments were performed using standard triple-resonance experiments. Analysis of spectral perturbations upon protein interactions and determination of equilibrium parameters were performed as described (30).

Notes—The chemical shift assignments of $^{2-3}$FnI, $^{4-5}$FnI, and ^{3}FnIII have been deposited in the BioMagResBank under accession numbers 15756, 15758, and 15759, respectively. The amino acid sequences and numbering schemes used here correspond to the human FN UniProt accession number P02751. The $^{2-5}$FnI model was constructed from the $^{2-3}$FnI and $^{4-5}$FnI crystal structures (4) assuming that the $^{3-4}$FnI interface is similar to that of other FnI domain pairs. The $^{2-3}$FnIII model was constructed by threading the amino acid sequence through the $^{9-10}$FnIII structure (Protein Data Bank code 2MFN) using the Phyre web service (33) and then substituting ^{2}FnIII with the high resolution NMR structure of that domain excluding the flexible N terminus (5).

RESULTS

N-terminal FN Fragment Does Not Affect Fibroblast Migration—A previously identified alternatively spliced mRNA variant of FN (26, 27) encodes the FN N terminus up to the complete domain ^{3}FnIII and a highly variable C-terminal amino acid tail. We produced a recombinant version of this fragment excluding the tail (Fig. 1*a*) using transient expression in HEK293T cell line; we refer to this fragment as Fn100kDa,

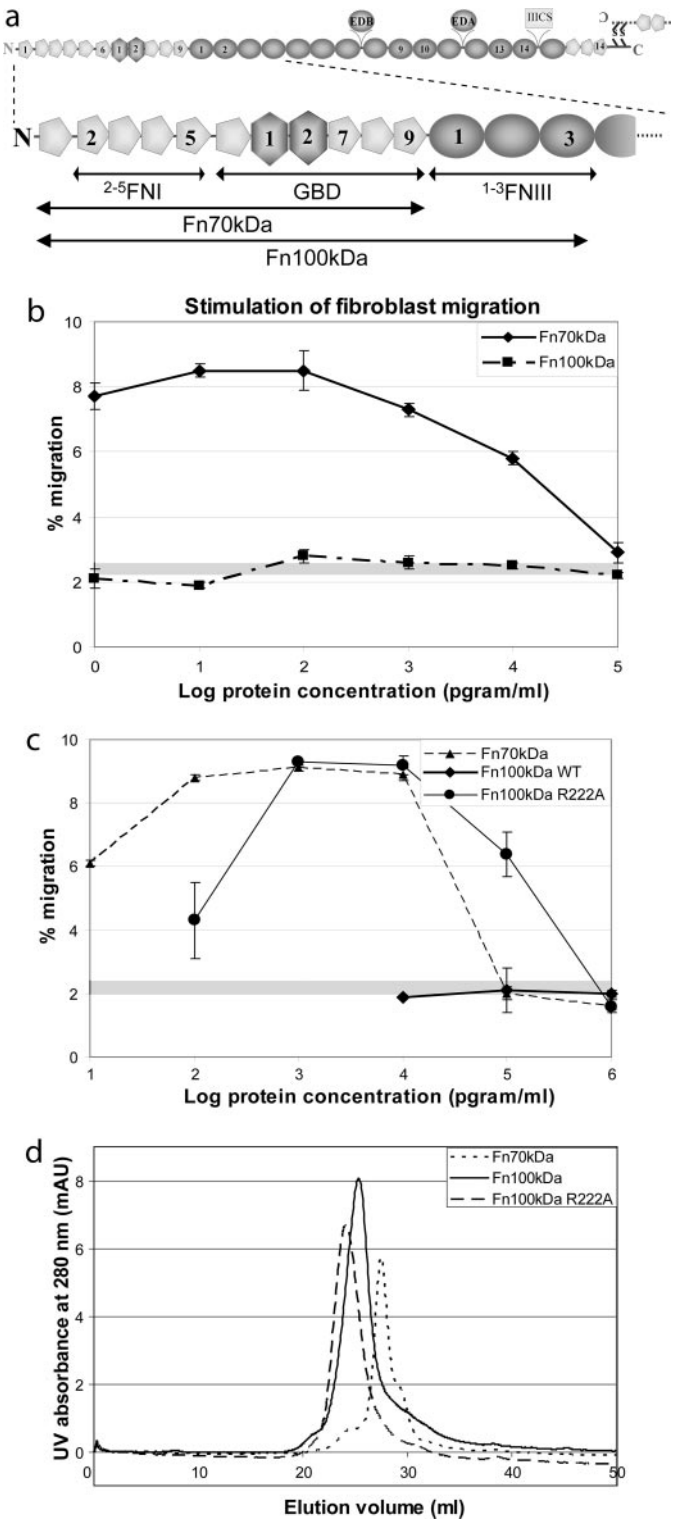

FIGURE 1. **Motogenic activity of FN fragments.** *a,* schematic representation of the FN domain structure (*top*) and enlargement of the FN N terminus (*bottom*). Type I domains are shown as *pentagons*; type II domains as *hexagons*; and type III domains as *ovals*. *b,* comparison of motogenic activity *versus* protein concentration of wild-type Fn70kDa and Fn100kDa fragments. *Error bars* are derived from duplicate experiments, and a *gray band* denotes migration activity of media without additives. *c,* similar comparisons for mutant Fn100kDa fragments. *d,* analytical size exclusion chromatography of large FN fragments. The trace of UV absorbance at 280 nm *versus* elution volume shown here indicates a larger hydrodynamic radius for Fn100kDa R222A compared with the wild type, consistent with our model (Fig. 6*a*).

TABLE 1
Dissociation constants from inter-domain interactions (K_d, mM)

FN fragments	^{1}FnIII	^{2}FnIII	^{3}FnIII
$^{1-2}$FnI	No interaction detected[a]	–[b]	–[b]
$^{2-3}$FnI		0.71 ± 0.02	0.82 ± 0.07
$^{4-5}$FnI		1.2 ± 0.2	0.146 ± 0.009
^{6}FnI$^{1-2}$FnII^{7}FnI		No interaction detected[a]	
$^{8-9}$FnI			

[a] Based on the extent of perturbations observed in other titrations, and the stoichiometric ratios used, we estimate a lower limit for these interactions as in excess of 7.5–10 mM.
[b] Small perturbations were detected on ^{2}FnI resonances but not on ^{1}FnI; hence we consider these interactions in the context of $^{2-3}$FnI.

based on its apparent mobility in denaturing gels. Fn100kDa is highly soluble, and analytical ultracentrifugation experiments showed that it is monomeric at ~11 μM (1.0 mg/ml) concentration (supplemental Fig. 1). This construct includes domains ^{7}FnI and ^{9}FnI that harbor the two IGD tripeptides necessary for stimulation of cell migration; however, assays using adult skin fibroblasts showed no effect on cell motility (Fig. 1*a*), a result similar that obtained with full-length FN (10, 23). Inclusion of the C-terminal amino acid tail in this construct did not alter this result (data not shown). In contrast, a recombinant fragment that lacks the first three FnIII domains (Fn70kDa) displays full activity in the same assays, in agreement with previous results (10). Hence, we hypothesized that domains $^{1-3}$FnIII are responsible for a structural rearrangement in Fn100kDa that masks the sites of motogenic activity. A likely cause of this rearrangement would be long range interactions between the FnIII domains and the remainder of Fn100kDa. We therefore sought to identify any such interactions using NMR.

Specific Interactions between Domains in Fn100kDa—We performed assays for interdomain interactions using solution NMR spectroscopy by monitoring chemical shift perturbations in ^{1}H-^{15}N HSQC spectra during titrations of FN fragments. Domains ^{1}FnIII, ^{2}FnIII, and ^{3}FnIII were tested against fragments $^{1-2}$FnI, $^{2-3}$FnI, $^{4-5}$FnI, ^{6}FnI$^{1-2}$FnII^{7}FnI, and $^{8-9}$FnI in all pairwise combinations at 30 °C and 20 mM NaCl, 20 mM sodium phosphate pH 6.0 buffer; the choice of buffer pH reflects a compromise between solubility of FnI domains (high at lower pH), and solubility and stability of FnIII domains (high at physiological pH). Under these conditions FnIII domains remained folded and soluble to at least 1 mM concentration, whereas FnI domains are less soluble (~0.2 mM for $^{4-5}$FnI).

As shown in Table 1, a total of four specific interactions were detected between $^{2-3}$FnI or $^{4-5}$FnI and ^{2}FnIII or ^{3}FnIII (Fig. 2); the interaction between $^{4-5}$FnI and ^{3}FnIII has an equilibrium dissociation constant (K_d) of 146 ± 9 μM under these conditions, whereas others are weaker (K_d of 700–1200 μM). FN conformation is sensitive to ionic strength (11, 34), and thus we tested for the persistence of these four interactions under physiological conditions (150 mM NaCl). Chemical shift perturbations were still observed, although we were unable to estimate dissociation constants accurately in some cases because of low fractional saturation in the respective titrations. The $^{4-5}$FnI-^{3}FnIII titration yielded a K_d of 800 ± 46 μM (data not shown); the 5-fold reduction in affinity is consistent with the increase in ionic strength, assuming a linear anticorrelation between these phenomena (35). Combined analysis of ^{1}H and ^{15}N perturbations shows good correlation between low and physiological

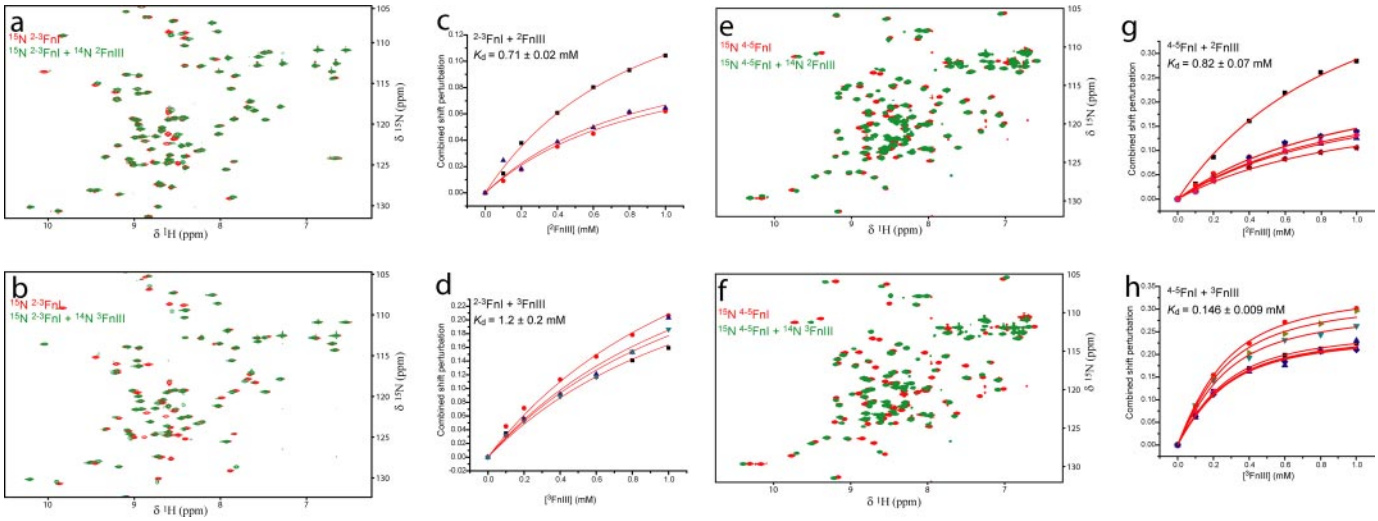

FIGURE 2. **NMR spectra of interdomain interactions.** *a* and *b*, ^{1}H-^{15}N HSQC spectra of $^{2-3}$FnI alone (*red*) or with excess ^{2}FnIII or ^{3}FnIII (*green*), respectively. *c* and *d*, fits of chemical shift perturbations against protein concentration to derive K_d values for the $^{2-3}$FnI-^{2}FnIII and $^{2-3}$FnI-^{3}FnIII titrations, respectively. *e–h*, HSQC spectra and titration fits for the $^{4-5}$FnI-^{2}FnIII and $^{4-5}$FnI-^{3}FnIII titrations.

NaCl concentrations (supplemental Fig. 2), suggesting that the domains interact in a similar fashion over this range of ionic strength.

Previously, we reported an interaction K_d under physiological ionic strength conditions of ~85 μM between wild-type $^{1-2}$FnIII and $^{1-5}$FnI (Fn30kDa), determined by surface plasmon resonance (5). This is substantially tighter than the interactions reported here between isolated FnIII domains and FnI pairs under similar conditions. However, the nature of the two techniques used, especially the difference of limited diffusion in surface plasmon resonance *versus* three-dimensional diffusion in NMR, makes comparisons across techniques difficult. The $^{4-5}$FnI-^{3}FnIII interaction did not prove amenable to surface plasmon resonance analysis (supplemental Fig. 3); thus, we tested the $^{1-2}$FnIII-Fn30kDa interaction by NMR (supplemental Fig. 4). As shown, only few and very small perturbations are detected in the NMR spectra under physiological conditions, compared with similar spectra of $^{4-5}$FnI-^{3}FnIII. Similarly, a recent study of wild-type $^{1-2}$FnIII and Fn70kDa did not show substantial interactions between these components using fluorescence (36). Hence, we infer that $^{4-5}$FnI-^{3}FnIII is the strongest among these interactions of wild-type proteins.

Mapping the chemical shift perturbations on the amino acid sequence (supplemental Fig. 5) and protein structure highlights the surfaces involved in the $^{2-5}$FnI and $^{2-3}$FnIII contacts (Fig. 3). ^{2}FnIII perturbs both FnI fragments along the domain-domain (^{2}FnI/^{3}FnI, ^{4}FnI/^{5}FnI) interfaces, which probably reflects a small change in the preferred orientation within the domain pairs; similar changes in FnI pairs have been noted earlier as a result of changing local contacts (28). In addition, ^{5}FnI is perturbed along the β-strand A/E interface and ^{2}FnI along strand E. The pattern of perturbations by ^{3}FnIII is similar to that of ^{2}FnIII; however, the perturbation on ^{2}FnI is more pronounced, and an additional interaction surface is present on ^{4}FnI along the β-strand D/E loop. It is likely that this additional interaction surface accounts for the higher affinity of ^{3}FnIII for $^{4-5}$FnI compared with ^{2}FnIII (146 ± 9 μM *versus* 820 ± 70 μM K_d). Neither ^{2}FnIII nor ^{3}FnIII significantly affect domain ^{3}FnI. Large scale mapping of the ^{2}FnIII and ^{3}FnIII interactions to a $^{2-5}$FnI model shows that the vast majority of perturbed areas localize on a single side of the model (Fig. 3, *a* and *b*).

Similar interface mapping of the $^{2-3}$FnI and $^{4-5}$FnI titrations on FnIII domains (Fig. 3, *c* and *d*) shows strong perturbations along ^{3}FnIII β-strand D and the C-terminal end of strand E for both interactions, although this effect is more highly pronounced in the $^{4-5}$FnI titration. In contrast, perturbations on ^{2}FnIII differ for the two interactions as $^{2-3}$FnI primarily affects the C terminus of β-strand G and the strand C/D loop (the part of the molecule proximal to ^{3}FnIII), whereas $^{4-5}$FnI affects the N terminus of strand G and the strand F/G loop (the distal part of the molecule compared with ^{3}FnIII). Inspection of the modeled $^{2-3}$FnIII indicates that, in both cases, the perturbed interfaces of ^{2}FnIII and ^{3}FnIII orient in opposite directions (Fig. 3, *c* and *d*); however, chains of FnIII domains are known to be prone to domain reorientation in solution (37).

The flexibility inherent in these multiple module fragments suggested that the observed interactions between adjacent domains could act cooperatively to enhance the binding affinity. $^{2-3}$FnIII can be easily produced, but we were initially unable to express $^{2-5}$FnI. Optimization of expression and purification procedures yielded only small amounts of this fragment (supplemental Fig. 6); these were, however, sufficient to test the effects of increasing fragment size. ^{15}N-enriched $^{2-3}$FnI at low concentrations (20 μM) was titrated with 50 μM unenriched $^{2-3}$FnIII (supplemental Fig. 7); titrations of substantially higher amounts of $^{2-3}$FnIII produced broadened spectra with little signal (data not shown). Chemical shift perturbations were detected for both $^{2-3}$FnI and $^{4-5}$FnI residues, and comparisons of the chemical shift changes to those observed for smaller fragments indicated a fractional saturation of ~25%. At the protein concentrations used, this fraction corresponds to an estimated K_d of 125 μM, equivalent to that of the $^{4-5}$FnI-^{3}FnIII titration. Indeed, titrations of $^{2-5}$FnI with ^{3}FnIII alone produced similar spectra to those with $^{2-3}$FnIII (supplemental Fig. 7), consistent with no cooperative ^{2}FnIII binding. We interpret the $^{2-3}$FnI perturbations in the latter titration as occurring from a second-

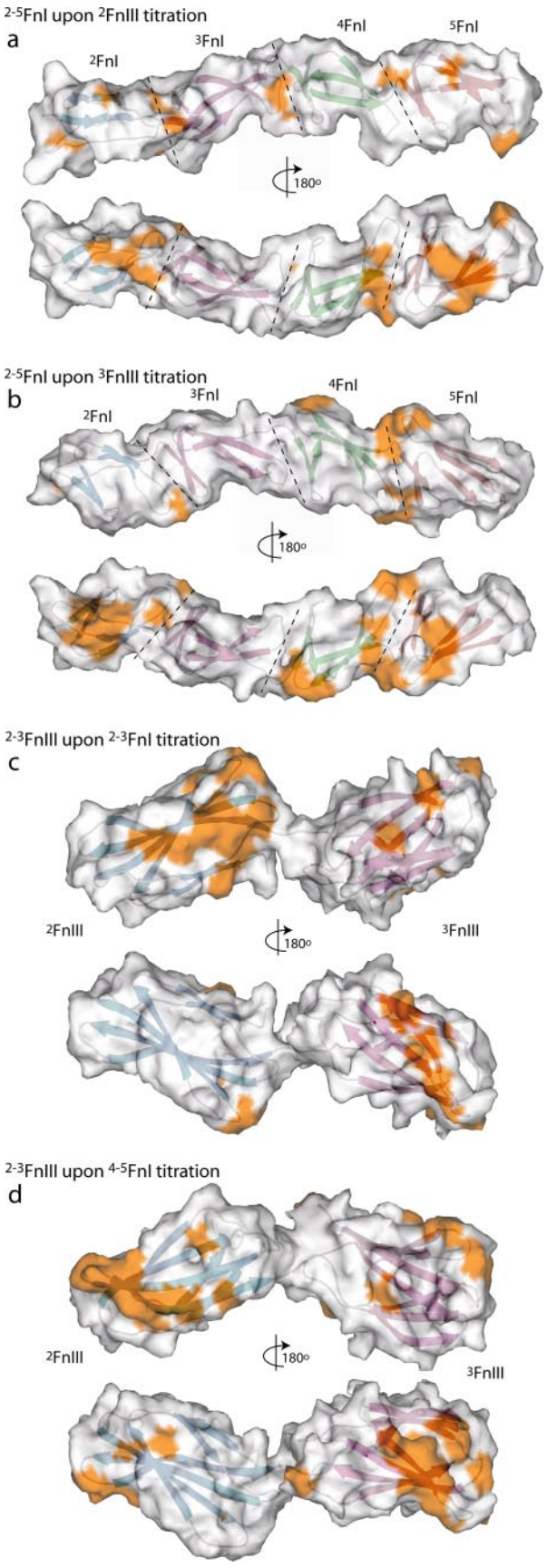

FIGURE 3. **Interaction surfaces from chemical shift perturbations.** *a* and *b*, molecular surface area representations of putative $^{2-5}$FnI models in two opposite orientations. The surfaces of residues with chemical shift perturbations in the ^{2}FnIII (*a*) or ^{3}FnIII (*b*) titrations greater than 1 S.D. from the average are colored in *gold*. The structures of domains ^{2}FnI (*blue*), ^{3}FnI (*purple*), ^{4}FnI (*green*), and ^{5}FnI (*red*) are shown, and the domain-domain interfaces are indicated by *dashed lines*. *c* and *d*, similar representations of a putative $^{2-3}$FnIII model upon $^{2-3}$FnI (*c*) or $^{4-5}$FnI (*d*) titration. ^{2}FnIII is shown in *blue* and ^{3}FnIII in *purple*.

ary binding event. ^{3}FnIII would preferentially interact with the $^{4-5}$FnI pair in $^{2-5}$FnI, leading to an increase in the effective local ^{3}FnIII concentration and a higher probability of a subsequent $^{2-3}$FnI-^{3}FnIII interaction.

Anastellin Disrupts the $^{4-5}$FnI-^{3}FnIII Interaction through ^{3}FnIII Destabilization—Anastellin, a truncated form of ^{1}FnIII, causes multimerization of plasma FN to a matrix-like form, superfibronectin (38). Structurally, anastellin resembles canonical type III folds (39) but displays properties different to those of ^{1}FnIII, including direct binding to ^{3}FnIII (5, 40). This interaction results in ^{3}FnIII destabilization, which is important for superfibronectin formation (40). We monitored ^{3}FnIII destabilization in the ^{1}H-^{15}N HSQC spectra of this domain during titrations of unenriched anastellin at pH 7.0 (Fig. 4, *a–c*). Large spectral changes indicative of complete domain unfolding were observed at anastellin concentrations only slightly higher than equimolar, consistent with past reports (40). In contrast, similar titrations of anastellin with ^{2}FnIII did not result in molecular interactions or ^{2}FnIII destabilization (data not shown). To test the effect of anastellin on the $^{4-5}$FnI-^{3}FnIII interaction, we performed titrations of both components in ^{15}N-enriched $^{4-5}$FnI (Fig. 4, *d* and *e*). The chemical shift perturbations caused by ^{3}FnIII become smaller upon addition of anastellin, and resonance peaks tend to revert to those of uncomplexed $^{4-5}$FnI. Together, these data indicate that the $^{4-5}$FnI-^{3}FnIII interaction is disrupted *in vitro* by anastellin, probably because of ^{3}FnIII destabilization; this disruption is potentially the initiation event for superfibronectin formation (38).

Disruption of the $^{4-5}$FnI-^{3}FnIII Interaction Restores Motogenic Activity to Fn100kDa—The $^{4-5}$FnI-^{3}FnIII interaction is the strongest observed; thus, we attempted to disrupt it through amino acid substitutions to investigate its possible role in stabilizing long range order in Fn100kDa. Our results show that this interaction includes a large electrostatic component with substantially reduced affinity at higher ionic strength. The electrostatic and interaction surface maps of these proteins (Fig. 5, *a* and *b*) show a number of charged residues located at, or near, perturbed regions. Some of these were targeted for substitutions, including Arg-222, Arg-228, and Arg-237 in $^{4-5}$FnI, and Glu-850, Glu-855, and Asp-869 in ^{3}FnIII. Titrations with proteins bearing these substitutions resulted in a range of affinities (Fig. 5*f*); the largest effect, of over 20-fold compared with the wild-type interaction, was observed for the R222A substitution (Fig. 5, *c–e*).

Mutagenesis of Fn100kDa yielded no effect on stimulation of fibroblast migration for the R228A and R228A/E850A substitutions compared with the wild-type protein. In contrast, R222A and R222A/E850A in Fn100kDa produced motogenic activities with characteristic bell-shaped curves in migratory response *versus* protein concentration profiles (Fig. 1*c* and supplemental Fig. 8). Compared with Fn70kDa, these variants require ~10 times more protein for maximal activity, which

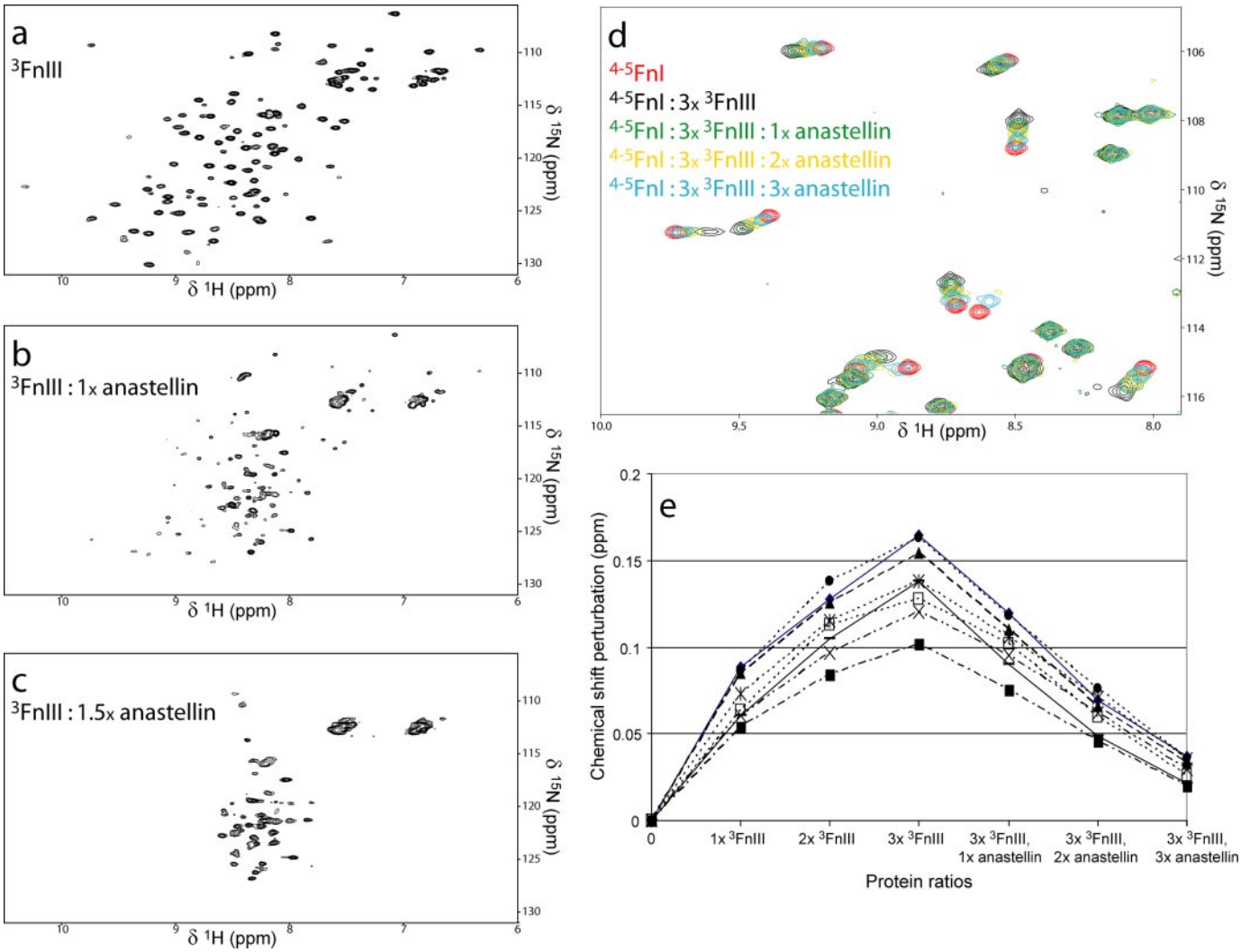

FIGURE 4. **Interaction of anastellin with ^{3}FnIII.** *a–c*, ^{1}H-^{15}N HSQC spectra of ^{3}FnIII alone (*a*) or in the presence of equimolar (1×, *b*) or higher (1.5×, *c*) amounts of anastellin. *d*, detail from an overlay of HSQC spectra of $^{4-5}$FnI alone, in the presence of 3-fold excess ^{3}FnIII or ^{3}FnIII, and stoichiometric ratios of anastellin. *e*, extent of chemical shift perturbations on select $^{4-5}$FnI resonances upon ^{3}FnIII or ^{3}FnIII and anastellin titrations denoted as protein ratios with respect to $^{4-5}$FnI.

indicates that reversal of the conformational change responsible for occlusion of the stimulation activity is only partial.

Exposure of motogenic activity sites in our Fn100kDa variants could be the result of change in structural state between a "closed" and "open" conformation (Fig. 6*a*), which may involve large differences in chain compaction. Such differences could be detectable in assays that report on the hydrodynamic properties of the macromolecule. Thus, we performed analytical size exclusion chromatography assays on Fn70kDa, Fn100kDa, and Fn100kDa R222A. As shown in Fig. 1*d*, wild-type Fn100kDa is retained longer in the size exclusion column compared with the R222A variant, indicating a more compact conformation, consistent with our hypothesis (Fig. 6*a*).

DISCUSSION

Long range FN structure and the nature of conformational transitions uncovering functionally cryptic sites have been the subject of much research and speculation. However, the large size of the molecule, its potential heterogeneity when purified from natural source, its sensitivity to changing environment, and difficulties with specific labeling have hindered description of the soluble plasma FN state. The system described here couples a sensitive biological assay that can detect functional changes, with FN variants that are relatively simple to produce and characterize. Furthermore, wild-type Fn100kDa displays activity properties similar to intact FN, which make it a good model system and an exciting new resource in the field of FN and matrix biology. Applications of fluorescent resonance energy transfer experiments (12, 13) in single molecule systems and coupled with computation are of particular interest.

Fn100kDa includes domains that can confer motogenic activity, but it does not stimulate fibroblast migration. We have shown here that the conformational transition responsible for this masking effect depends on the presence of a specific interaction between domains ^{3}FnIII and $^{4-5}$FnI; this interaction is expected to occur within a single Fn100kDa molecule as this protein remains monomeric at concentrations much higher than those used in biological assays. It is likely that a similar intramolecular interaction is responsible for the lack of motogenic activity in full-length FN; however, interactions across the FN dimer cannot be excluded. Although we cannot construct a reliable structural model of Fn100kDa at this time,

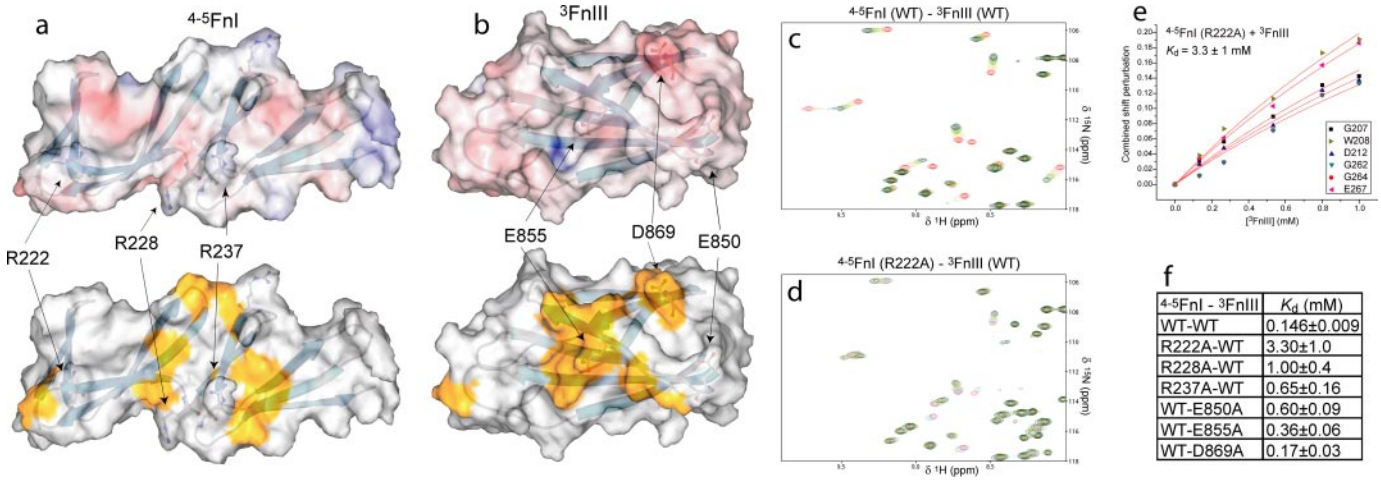

FIGURE 5. **Disruption of the $^{4-5}$FnI-^{3}FnIII interaction.** Molecular surface area representation of $^{4-5}$FnI (*a*) and ^{3}FnIII (*b*) colored by electrostatic potential (*top*) or showing residues strongly perturbed in the interaction in *gold* (*bottom*). The residues substituted are indicated. Details of overlays of the NMR spectra from the wild-type titration (*c*) and the $^{4-5}$FnI (R222A)-^{3}FnIII titration (*d*) are shown. The fit of chemical shift change against protein concentration for this substitution is shown in *e* and all dissociation constants measured for substitutions in *f*.

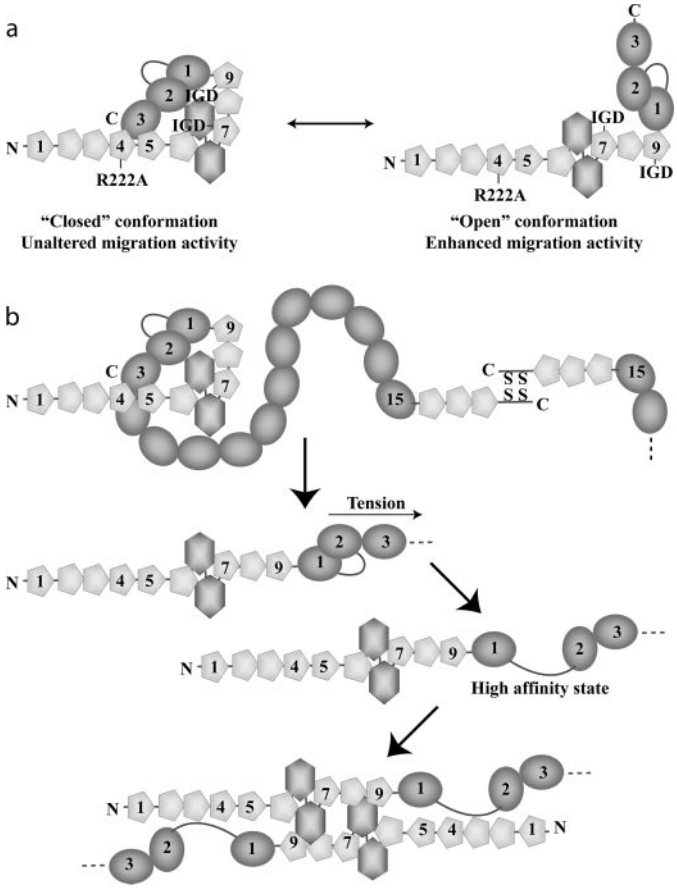

FIGURE 6. **Model of FN conformational transitions and fibrillogenesis.** *a,* Fn100kDa possibly adopts a bent-back compact conformation (closed state, *left*) stabilized by intramolecular interactions, including $^{4-5}$FnI-^{3}FnIII; a similar conformation is likely present in full-length FN. The motogenic activity sites on ^{7}FnI and ^{9}FnI are occluded while in this state, which is destabilized by the R222A substitution toward an alternative, open state (*right*). *b,* in FN, tension or fibril-promoting agents disrupt the $^{4-5}$FnI-^{3}FnIII interaction, although further tension yields a high affinity state prone to oligomerization (5).

primarily because of flexible regions present in the molecule and the unknown conformation of intact GBD, our results suggest that the termini of Fn100kDa are in proximity creating an overall "bent-back" compact conformation we term the closed state (Fig. 6*a*). This state is likely to be stabilized by additional interactions acting in tandem; these interactions can be very weak, but if they are intramolecular the effective local concentration will be high enough to allow them to form readily. The importance of these interactions, particularly between $^{4-5}$FnI and ^{3}FnIII, is illustrated by anastellin-mediated destabilization of ^{3}FnIII and disruption of the interaction with $^{4-5}$FnI, the likely trigger for superfibronectin formation (38). The interactions of ^{2}FnIII described here are less direct than those of ^{3}FnIII in the FN fragments tested, but they could be important in the context of matrix formation where interaction heterogeneity and diversity may contribute to fibril growth.

In previous studies we showed an interaction between domains ^{1}FnIII and ^{2}FnIII (5). Disruption of this interaction by amino acid substitutions (a construct we called KADA) led to a high affinity (nanomolar) interaction between $^{1-2}$FnIII KADA and the FN N terminus (Fn30kDa, $^{1-5}$FnI); a similar result was recently shown by fluorescence (36). Technical limitations prevent direct characterization of the high affinity complex by NMR, as Fn30kDa-$^{1-2}$FnIII KADA complexes tend to aggregate and precipitate (data not shown) under physiological conditions. However, a strong interaction is clearly present in the KADA mutant, and we believe that it will likely be dominant in conditions that favor fibrillogenesis, such as cell-generated tension or in the presence of fibril-promoting agents (38). Combining these earlier results with our present study allows us to create a more complete model of the initial fibrillogenesis steps (Fig. 6*b*). Soluble adult FN is likely to be held in compact form by inter-domain interactions, including those between $^{4-5}$FnI and ^{3}FnIII. Tension, or fibril-promoting agents such as anastellin, will disrupt this initial conformation. Further application of force, or thermal fluctuations, could disrupt the ^{1}FnIII-^{2}FnIII complex creating a strong interaction with the FN N terminus. FN will then create oligomers that, in concert with the FN dimeric state, produce protofibrils and eventually the FN matrix. Interestingly, "handles" in soluble FN used by cells to initiate FN matrix formation have been mapped on Fn70kDa

and ^{1}FnIII (41, 42). It is possible that these handles permit unraveling of the FN100kDa compact state by cells *in vivo*.

The Fn100kDa closed conformation occludes motogenic activity sites on GBD as shown in our functional assays (Fig. 1). It is possible that this state also interferes with other identified GBD-related activities, such as collagen binding (29, 43, 44), attachment of pathogenic bacteria (45, 46), or transglutaminase-mediated cross-linking (47, 48). In contrast, removal of $^{1-3}$FnIII in MSF or Fn70kDa alters the conformation to expose sites of biological activity (Fig. 1 and Fig. 6*a*). MSF has been proposed to be a by-product of matrix metalloprotease cleavage of FN (49), possibly increasing the effect of metalloprotease activity on cell migration (50). However, MSF (10) and the Fn100kDa construct characterized here (26, 27) are also found as splicing variants of the FN gene. Therefore, migration stimulation and possibly other FN-mediated activities can be regulated at the level of mRNA processing. We have shown here that sites for these activities are always present in FN, but specific association interactions regulate their display; this mechanism renders FN activities sensitive to splicing variation by addition or removal of protein domains. We postulate that alternative splicing as well as mechanosensing could be functionally conserved methods for regulation of FN properties, as well as control of other multipotent biomolecules.

Acknowledgments—We thank Nick Soffe and Dr. Jonathan Boyd for assistance with the NMR instrumentation, Dr. Michèle C. Erat for critical reading of the manuscript, and Dr. Radu Aricescu for help with the HEK293T transient expression system. Funding for the Oxford Instruments 22.3 tesla (950 MHz ^{1}H frequency) superconducting magnet was provided by the Wellcome Trust.

REFERENCES

1. Vakonakis, I., and Campbell, I. D. (2007) *Curr. Opin. Cell Biol.* **19,** 578–583
2. Leahy, D. J., Aukhil, I., and Erickson, H. P. (1996) *Cell* **84,** 155–164
3. Pickford, A. R., Smith, S. P., Staunton, D., Boyd, J., and Campbell, I. D. (2001) *EMBO J.* **20,** 1519–1529
4. Bingham, R. J., Rudiño-Piñera, E., Meenan, N. A., Schwarz-Linek, U., Turkenburg, J. P., Höök, M., Garman, E. F., and Potts, J. R. (2008) *Proc. Natl. Acad. Sci. U. S. A.* **105,** 12254–12258
5. Vakonakis, I., Staunton, D., Rooney, L. M., and Campbell, I. D. (2007) *EMBO J.* **26,** 2575–2583
6. Morla, A., and Ruoslahti, E. (1992) *J. Cell Biol.* **118,** 421–429
7. Hocking, D. C., Sottile, J., and McKeown-Longo, P. J. (1994) *J. Biol. Chem.* **269,** 19183–19187
8. Ingham, K. C., Brew, S. A., Huff, S., and Litvinovich, S. V. (1997) *J. Biol. Chem.* **272,** 1718–1724
9. Ingham, K. C., Brew, S. A., and Erickson, H. P. (2004) *J. Biol. Chem.* **279,** 28132–28135
10. Schor, S. L., Ellis, I. R., Jones, S. J., Baillie, R., Seneviratne, K., Clausen, J., Motegi, K., Vojtesek, B., Kankova, K., Furrie, E., Sales, M. J., Schor, A. M., and Kay, R. A. (2003) *Cancer Res.* **63,** 8827–8836
11. Ugarova, T. P., Zamarron, C., Veklich, Y., Bowditch, R. D., Ginsberg, M. H., Weisel, J. W., and Plow, E. F. (1995) *Biochemistry* **34,** 4457–4466
12. Lai, C. S., Wolff, C. E., Novello, D., Griffone, L., Cuniberti, C., Molina, F., and Rocco, M. (1993) *J. Mol. Biol.* **230,** 625–640
13. Wolff, C., and Lai, C. S. (1988) *Biochemistry* **27,** 3483–3487
14. Nelea, V., Nakano, Y., and Kaartinen, M. T. (2008) *Protein J.* **27,** 223–233
15. Sjöberg, B., Pap, S., Osterlund, E., Osterlund, K., Vuento, M., and Kjems, J. (1987) *Arch. Biochem. Biophys.* **255,** 347–353
16. Vuillard, L., Roux, B., and Miller, A. (1990) *Eur. J. Biochem.* **191,** 333–336
17. Pelta, J., Berry, H., Fadda, G. C., Pauthe, E., and Lairez, D. (2000) *Biochemistry* **39,** 5146–5154
18. Litvinovich, S. V., and Ingham, K. C. (1995) *J. Mol. Biol.* **248,** 611–626
19. Spitzfaden, C., Grant, R. P., Mardon, H. J., and Campbell, I. D. (1997) *J. Mol. Biol.* **265,** 565–579
20. Altroff, H., Schlinkert, R., van der Walle, C. F., Bernini, A., Campbell, I. D., Werner, J. M., and Mardon, H. J. (2004) *J. Biol. Chem.* **279,** 55995–56003
21. Schor, S. L., Schor, A. M., Grey, A. M., and Rushton, G. (1988) *J. Cell Sci.* **90,** 391–399
22. Houard, X., Germain, S., Gervais, M., Michaud, A., van den Brûle, F., Foidart, J. M., Noël, A., Monnot, C., and Corvol, P. (2005) *Int. J. Cancer* **116,** 378–384
23. Schor, S. L., Ellis, I., Dolman, C., Banyard, J., Humphries, M. J., Mosher, D. F., Grey, A. M., Mould, A. P., Sottile, J., and Schor, A. M. (1996) *J. Cell Sci.* **109,** 2581–2590
24. Millard, C. J., Ellis, I. R., Pickford, A. R., Schor, A. M., Schor, S. L., and Campbell, I. D. (2007) *J. Biol. Chem.* **282,** 35530–35535
25. Schor, S. L., Ellis, I., Banyard, J., and Schor, A. M. (1999) *J. Cell Sci.* **112,** 3879–3888
26. Zhao, Q., Liu, X., and Collodi, P. (2001) *Exp. Cell Res.* **268,** 211–219
27. Liu, X., Zhao, Q., and Collodi, P. (2003) *Matrix Biol.* **22,** 393–396
28. Rudiño-Piñera, E., Ravelli, R. B., Sheldrick, G. M., Nanao, M. H., Korostelev, V. V., Werner, J. M., Schwarz-Linek, U., Potts, J. R., and Garman, E. F. (2007) *J. Mol. Biol.* **368,** 833–844
29. Erat, M. C., Slatter, D. A., Lowe, E. D., Millard, C. J., Farndale, R. W., Campbell, I. D., and Vakonakis, I. (2009) *Proc. Natl. Acad. Sci. U. S. A.* **106,** 4195–4200
30. Vakonakis, I., Langenhan, T., Prömel, S., Russ, A., and Campbell, I. D. (2008) *Structure* **16,** 944–953
31. Aricescu, A. R., Lu, W., and Jones, E. Y. (2006) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **62,** 1243–1250
32. Schor, S. L. (1980) *J. Cell Sci.* **41,** 159–175
33. Bennett-Lovsey, R. M., Herbert, A. D., Sternberg, M. J., and Kelley, L. A. (2008) *Proteins* **70,** 611–625
34. Alexander, S. S., Jr., Colonna, G., and Edelhoch, H. (1979) *J. Biol. Chem.* **254,** 1501–1505
35. Schreiber, G. (2002) *Curr. Opin. Struct. Biol.* **12,** 41–47
36. Karuri, N. W., Lin, Z., Rye, H. S., and Schwarzbauer, J. E. (2009) *J. Biol. Chem.* **284,** 3445–3452
37. Copié, V., Tomita, Y., Akiyama, S. K., Aota, S., Yamada, K. M., Venable, R. M., Pastor, R. W., Krueger, S., and Torchia, D. A. (1998) *J. Mol. Biol.* **277,** 663–682
38. Morla, A., Zhang, Z., and Ruoslahti, E. (1994) *Nature* **367,** 193–196
39. Briknarová, K., Akerman, M. E., Hoyt, D. W., Ruoslahti, E., and Ely, K. R. (2003) *J. Mol. Biol.* **332,** 205–215
40. Ohashi, T., and Erickson, H. P. (2005) *J. Biol. Chem.* **280,** 39143–39151
41. Tomasini-Johansson, B. R., Annis, D. S., and Mosher, D. F. (2006) *Matrix Biol.* **25,** 282–293
42. Xu, J., Bae, E., Zhang, Q., Annis, D. S., Erickson, H. P., and Mosher, D. F. (2009) *PLoS ONE* **4,** e4113
43. Ingham, K. C., Brew, S. A., and Isaacs, B. S. (1988) *J. Biol. Chem.* **263,** 4624–4628
44. Ingham, K. C., Brew, S. A., and Migliorini, M. (2002) *Arch. Biochem. Biophys.* **407,** 217–223
45. Ozeri, V., Tovi, A., Burstein, I., Natanson-Yaron, S., Caparon, M. G., Yamada, K. M., Akiyama, S. K., Vlodavsky, I., and Hanski, E. (1996) *EMBO J.* **15,** 989–998
46. Talay, S. R., Zock, A., Rohde, M., Molinari, G., Oggioni, M., Pozzi, G., Guzman, C. A., and Chhatwal, G. S. (2000) *Cell. Microbiol.* **2,** 521–535
47. Akimov, S. S., and Belkin, A. M. (2001) *J. Cell Sci.* **114,** 2989–3000
48. Radek, J. T., Jeong, J. M., Murthy, S. N., Ingham, K. C., and Lorand, L. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90,** 3152–3156
49. Mott, J. D., and Werb, Z. (2004) *Curr. Opin. Cell Biol.* **16,** 558–564
50. Zaman, M. H., Matsudaira, P., and Lauffenburger, D. A. (2007) *Ann. Biomed. Eng.* **35,** 91–100