Fibronectin at a glance

Roumen Pankov and Kenneth M. Yamada

Craniofacial Developmental Biology and Regeneration Branch, National Institute of Dental and Craniofacial Research, National Institutes of Health, Bethesda, MD 20892, USA

(e-mail: Roumen.Pankov@nih.gov and Kenneth.Yamada@nih.gov)

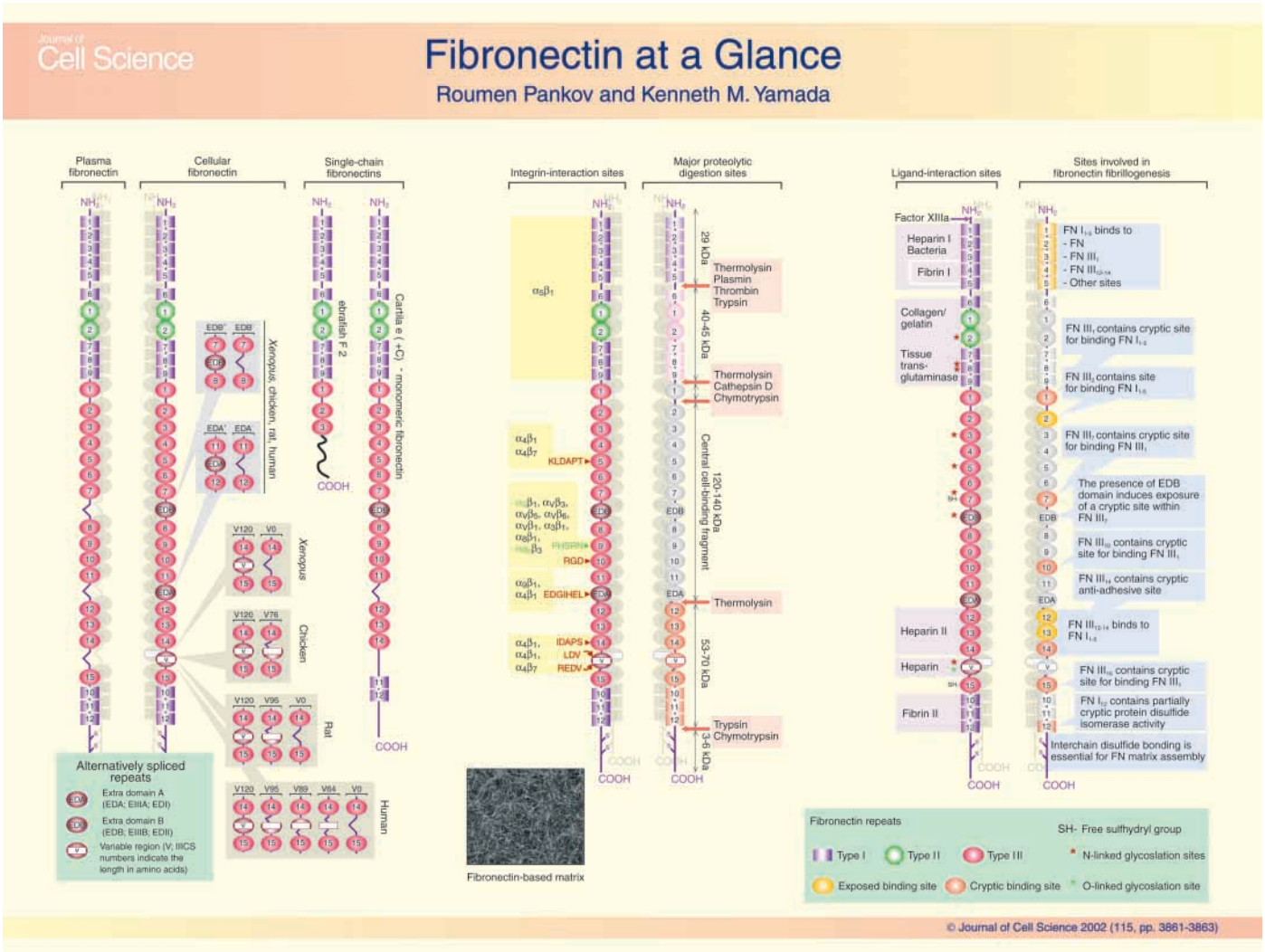
Journal of Cell Science 115, 3861-3863

doi:10.1242/jcs.00059

Fibronectin (FN) mediates a wide variety of cellular interactions with the extracellular matrix (ECM) and plays important roles in cell adhesion, migration, growth and differentiation (Mosher, 1989; Carsons, 1989; Hynes, 1990; Yamada and Clark, 1996). FN is widely expressed by multiple cell types and is critically important in vertebrate development, as demonstrated by the early embryonic lethality of mice with targeted inactivation of the *FN* gene (George et al., 1993). Although FN has been studied for more than two decades, this remarkably complex molecule is still the subject of exciting discoveries, such as finding new integrin- and heparin-binding sites (Mostafavi-Pour et al., 2001; Liao et al., 2002) or even a new form of the molecule (Zhao et al., 2001) that mediates a particular viral infection (Liu and Collodi, 2002).

FN usually exists as a dimer composed of two nearly identical ~250 kDa subunits linked covalently near their C-termini by a pair of disulfide bonds (see poster). Each monomer consists of three types of repeating units (termed FN repeats): type I (purple rectangles), type II (green octagons) and type III (red ovals). FN contains 12 type I repeats, two type II repeats and 15-17 type III repeats, which together account for approximately 90% of the FN sequence. Type I repeats are about 40 amino-acid residues in length and contain two disulfide bonds; type II repeats comprise a stretch of approximately 60 amino acids and two intrachain disulfide bonds; and type III repeats are about 90 residues long without any disulfide bonds. All three types of FN repeat are also found in other molecules, suggesting that FN evolved through exon shuffling (Patel et al., 1987).

Although FN molecules are the product of a single gene, the resulting protein can exist in multiple forms that arise from alternative splicing of a single pre-mRNA that can generate as many as 20 variants in human FN (for reviews, see



(See poster insert)

ffrench-Constant, 1995; Kosmehl et al., 1996) (left panel, Plasma and Cellular fibronectin). A major type of splicing occurs within the central set of type III repeats (left panel, FN III_7 to FN III_{15}). Exon usage or skipping leads to inclusion or exclusion of either of two type III repeats – EDB (also termed EIIIB or EDII and located between FN repeats III_7 and III_8) and EDA (also called EIIIA or EDI and located between FN repeats III_{11} and III_{12}). This 'yes or no' type of splicing of FN ED domains is found in many vertebrates, including *Xenopus*, chickens, rats and humans.

A third region of alternative splicing is localized to a non-homologous stretch called the V (variable in length) or IIICS (type III connecting segment) region. The structural variations in this region are more complex and species dependent (left panel, lower four gray boxes). In most species studied to date, except chicken, this region can be either partially or completely included or excluded; for example, in human FN, there can be five different V region variants.[1] A fourth type of splicing is found in cartilage, where the predominant form of FN [termed $(V+C)^-$] lacks the entire V region along with the FN III_{15} and FN I_{10} repeats. Interestingly, this FN isoform exists not only as a homodimer but also in an unusual monomeric configuration (left panel, Single-chain fibronectins). Recently, another single-chain FN (termed FN2) has been described in zebrafish, together with a FN1 form that is very similar to FNs identified in other vertebrates. The truncated FN2 form results from a fifth type of splicing in zebrafish (left panel, Single-chain fibronectins).

FN is an abundant soluble constituent of plasma (300 μg/ml) and other body fluids and also part of the insoluble extracellular matrix. On the basis of its solubility, FN can be subdivided into two forms – soluble plasma FN (pFN) and less-soluble cellular (cFN) FN. Plasma FN is synthesized predominantly in the liver by hepatocytes and shows a relatively simple splicing pattern. The alternatively spliced EDA and EDB domains are almost always absent from plasma FN, although both V0 and V+ are present (left panel, Plasma fibronectin). Cellular FN consists of a much larger and more heterogeneous group of FN isoforms that result from cell-type-specific and species-specific splicing patterns (left panel, Cellular fibronectin). Thus, alternative splicing of precursor mRNA from the single *FN* gene has the capacity to produce a large number of variants, generating FNs with different cell-adhesive, ligand-binding, and solubility properties that provide a mechanism for cells to precisely alter the composition of the ECM in a developmental and tissue-specific manner.

FN can be a ligand for a dozen members of the integrin receptor family (for a recent review, see Plow et al., 2000). Integrins are structurally and functionally related cell-surface heterodimeric receptors that link the ECM with the intracellular cytoskeleton. A large number of different integrins bind to FN, including the classic FN receptor $\alpha_5\beta_1$ (middle panel, Integrin interaction sites). Extensive analyses have narrowed down the regions involved in cell adhesion along the lengthy FN molecule to several minimal integrin-recognition sequences (middle panel, single amino-acid sequences in red). The best known of these – RGD – is located in FN repeat III_{10}. The recognition of this simple tripeptide sequence is complex and depends on flanking residues, its three-dimensional presentation and individual features of the integrin-binding pockets. For example, a second site in FN repeat III_9 (the 'synergy site' PHSRN, green) promotes specific $\alpha_5\beta_1$ integrin binding to FN, apparently via interactions with the α_5 subunit. However, binding of the FN receptor $\alpha_5\beta_1$ to FN is not restricted only to repeats III_9 and III_{10}. It can also interact with an N-terminal fragment containing repeats $I_{1\text{-}9}$ and $II_{1,2}$, which also promotes $\alpha_5\beta_1$-integrin-mediated cell adhesion. Interestingly, interaction with this N-terminal region can trigger integrin-mediated intracellular signals that are distinct from those generated in response to ligation with the RGD sequence.

A second set of FN sequences, which are bound by the $\alpha_4\beta_1$ integrin, has also received considerable attention. Two cell-recognition sequences (LDV and REDV) were originally identified in the alternatively spliced V region. Both of them are recognized by $\alpha_4\beta_1$ and $\alpha_4\beta_7$. Additional sites recognized by the $\alpha_4\beta_1$ integrin – IDAPS and KLDAPT – are also present in repeats III_{14} and III_5, respectively (the latter also binds to the $\alpha_4\beta_7$ integrin). Recently, binding of $\alpha_4\beta_1$ as well as $\alpha_9\beta_1$ to an EDGIHEL sequence located within the alternatively spliced EDA segment has been reported, suggesting a possible adhesive function for the increased EDA-containing FN species observed during wound healing (Liao et al., 2002).

Elucidation of the sites of integrin binding as well as other functionally important domains within the FN molecule was greatly facilitated by the early discovery that all FNs are cleaved only in specific regions when subjected to limited proteolytic digestion (reviewed by Mosher, 1989; Hynes, 1990). Even a protease capable of cleaving proteins at many sites (such as pronase) will initially cleave FN, and it will only do this at highly specific, probably non-folded, unprotected locations. A simplified scheme of the major proteolytic cleavage sites is shown in the middle panel (see Major proteolytic digestion sites). The binding activities of FN are often preserved after such proteolysis and can be identified within particular fragments.

FN has a remarkably wide variety of functional activities besides binding to cell surfaces through integrins. It binds to a number of biologically important molecules that include heparin, collagen/gelatin, and fibrin. These interactions are mediated by several distinct structural and functional domains, which have been defined by proteolytic fragmentation or recombinant DNA analyses (see Mosher, 1989; Hynes, 1990; Yamada and Clark, 1996; and the website

[1]In chicken, the whole 120 amino acid residues of the V region can be included or a 44 amino acid segment from the 5′ end can be excluded (creating V76), but the whole V region is never missing. A similar mechanism leads to exclusion of a 25 amino acid fragment in rat, generating V95 that can be detected together with V0 and V120 forms. Splicing of the V region is even more complicated in human where segments from both 5′ (25aa) and 3′ (31aa) ends can be omitted independently (creating V95 and V89 correspondingly) or together (V64) producing five different V regions.

http://www.gwumc.edu/biochem/ingham/fnpage.htm).

FN contains two major heparin-binding domains that interact with heparan sulfate proteoglycans (right panel, Ligand interaction sites). The strongest heparin-binding site is located in the C-terminal part (Heparin II) and a weaker binding domain is situated at the N-terminal end of the protein (Heparin I). The high-affinity heparin II domain can also bind to a widely distributed glycosaminoglycan, chondroitin sulfate, whereas the weaker heparin-binding domain contains a *Staphylococcus-aureus*-binding site that mediates FN interactions with bacteria. Recently, a novel glycosaminoglycan-binding site has been identified within the V region of FN (Mostafavi-Pour et al., 2001) (marked as 'Heparin' at the V domain). In at least some cell types, the heparin-binding domains of FN potentiate cell adhesion.

The collagen-binding domain includes repeats $I_{6\text{-}9}$ and $II_{1,2}$, and it binds far more effectively to denatured collagen (gelatin) than to native collagen. Thus, FN interactions with collagens in general may be due to its binding to unfolded regions of the collagen triple helix. It has been suggested that the physiological function of the collagen/gelatin-binding domain is related more to binding and clearance of denatured collagenous materials from blood and tissue than to mediating cell adhesion to collagen. Interestingly, however, a recent analysis of the physiological state of collagen indicates that the triple helix is likely to unfold locally at body temperature (Leikina et al., 2002), which suggests that this FN domain could be involved in interactions with native collagen in vivo.

FN also contains two major fibrin-binding sites (Fibrin I and Fibrin II). The major site is in the N-terminal domain and is formed by type I repeats 4 and 5. The interaction of FN with fibrin is thought to be important for cell adhesion or cell migration into fibrin clots. In both cases, cross-linking between FN and fibrin mediated by factor XIII transglutaminase is proposed to mediate the effect (the cross-linking site on the FN molecule is marked by factor XIIIa and an arrow). The interaction of FN with fibrin may also be involved in macrophage clearance of fibrin from circulation after trauma or in inflammation.

FNs are glycoproteins that contain 4-9% carbohydrate, depending on the cell source. Glycosylation sites that are either N-linked (red stars) or O-linked (green star) reside predominantly within type III repeats and the collagen-binding domain. The physiological role of the carbohydrates is not certain, although they appear to stabilize FN against hydrolysis and modulate its affinity to some substrates.

Although the plasma FN that circulates in blood is in a closed, reportedly non-active form, most of the FN activities in the body have been ascribed to the insoluble form of FN that exists as part of the extracellular matrix (see the immunofluorescence image obtained with anti-FN antibodies at the bottom of the middle panel labeled Fibronectin-based matrix). The creation and deposition of insoluble FN fibrils into the ECM is a tightly regulated, cell-mediated process termed FN fibrillogenesis or FN matrix assembly (for a review, see Geiger et al., 2001). A critical step in this process is self-association of FN into aggregates and fibrils, which is directed by multiple binding sites that have been identified along the molecule (right panel, Sites involved in fibronectin fibrillogenesis). Some of these self-interaction sites are exposed and available for binding (marked in yellow), while others are cryptic (marked in light brown) and become accessible only after conformational changes, for example, by cell-driven mechanical stretching of the FN molecule.

FN is one of the largest multi-domain proteins for which domain organization, molecular interactions, and key functions have been established in great detail. Exploration of the cell-type-specific splicing variants, glycosylation patterns and their relationship to health and disease will be further challenges in the study of this fascinating molecule.

References

Carsons, S. E. (1989). *Fibronectin in Health and Disease*. Florida: CRC Press, Inc.

ffrench-Constant, C. (1995). Alternative splicing of fibronectin – many different proteins but few different functions. *Exp. Cell Res.* **221**, 261-271.

Geiger, B., Bershadsky, A., Pankov, R. and Yamada, K. M. (2001). Transmembrane extracellular matrix-cytoskeleton crosstalk. *Nat. Rev. Mol. Cell. Biol.* **2**, 793-805.

George, E. L., Georges-Labouesse, E. N., Patel-King, R. S., Rayburn, H. and Hynes, R. O. (1993). Defects in mesoderm, neural tube and vascular development in mouse embryos lacking fibronectin. *Development* **119**, 1079-1091.

Hynes, R. O. (1990). *Fibronectins*. New York: Springer-Verlag.

Kosmehl, H., Berndt, A. and Katenkamp, D. (1996). Molecular variants of fibronectin and laminin: structure, physiological occurrence and histopathological aspects. *Virchows Arch.* **429**, 311-322.

Leikina, E., Mertts, M. V., Kuznetsova, N. and Leikin, S. (2002). Type I collagen is thermally unstable at body temperature. *Proc. Natl. Acad. Sci. USA* **99**, 1314-1318.

Liao, Y. F., Gotwals, P. J., Koteliansky, V. E., Sheppard, D. and Van De Water, L. (2002). The EIIIA segment of fibronectin is a ligand for integrins $\alpha_9\beta_1$ and $\alpha_4\beta_1$ providing a novel mechanism for regulating cell adhesion by alternative splicing. *J. Biol. Chem.* **277**, 14467-14474.

Liu, X. and Collodi, P. (2002). Novel form of fibronectin from zebrafish mediates infectious hematopoietic necrosis virus infection. *J. Virol.* **76**, 492-498.

Mosher, D. F. (1989). *Fibronectin*. San Diego: Academic Press, Inc.

Mostafavi-Pour, Z., Askari, J. A., Whittard, J. D. and Humphries, M. J. (2001). Identification of a novel heparin-binding site in the alternatively spliced IIICS region of fibronectin: roles of integrins and proteoglycans in cell adhesion to fibronectin splice variants. *Matrix Biol.* **20**, 63-73.

Patel, R. S., Odermatt, E., Schwarzbauer, J. E. and Hynes, R. O. (1987). Organization of the fibronectin gene provides evidence for exon shuffling during evolution. *EMBO J.* **6**, 2565-2572.

Plow, E. F., Haas, T. A., Zhang, L., Loftus, J. and Smith, J. W. (2000). Ligand binding to integrins. *J. Biol. Chem.* **275**, 21785-21788.

Yamada, K. M. and Clark, R. A. F. (1996). Provisional matrix. In *The Molecular and Cellular Biology of Wound Repair* (ed. R. A. F. Clark), pp. 51-93. New York: Plenum Press.

Zhao, Q., Liu, X. and Collodi, P. (2001). Identification and characterization of a novel fibronectin in zebrafish. *Exp. Cell Res.* **268**, 211-219.

Cell Science at a Glance on the Web

Electronic copies of the poster insert are available in the online version of this article at jcs.biologists.org. The JPEG images can be downloaded for printing or used as slides.