Crystal structure of a heparin- and integrin-binding segment of human fibronectin

Amit Sharma¹, Janet A. Askari²,
Martin J. Humphries², E. Yvonne Jones^{1,3}
and David I. Stuart^{1,3,4}

¹Laboratory of Molecular Biophysics, University of Oxford, Rex Richardson Building, South Parks Road, Oxford, OX1 3QU, ²Wellcome Trust Centre for Cell-Matrix Research, School of Biological Sciences, 2205 Stopford Building, University of Manchester, M13 9PT, ³Oxford Centre for Molecular Sciences, New Chemistry Laboratory, South Parks Road, Oxford, OX1 3QT, UK

⁴Corresponding author
e-mail: dave@biop.ox.ac.uk

The crystal structure of human fibronectin (FN) type III repeats 12–14 reveals the primary heparin-binding site, a clump of positively charged residues in FN13, and a putative minor site ~60 Å away in FN14. The IDAPS motif implicated in integrin $\alpha_4\beta_1$ binding is at the FN13–14 junction, rendering the critical Asp184 inaccessible to integrin. Asp184 clamps the BC loop of FN14, whose sequence (PRARI) is reminiscent of the synergy sequence (PHSRN) of FN9. Mutagenesis studies prompted by this observation reveal that both arginines of the PRARI sequence are important for $\alpha_4\beta_1$ binding to FN12–14. The PRARI motif may represent a new class of integrin-binding sites. The spatial organization of the binding sites suggests that heparin and integrin may bind in concert.

Keywords: fibronectin/heparin binding/integrin binding/protein structure/type III

Introduction

Fibronectin (FN) is a large glycoprotein found on cell surfaces, in the connective tissue matrix and in extracellular fluids. It participates in cell adhesion, spreading, migration, extracellular matrix formation, hemostasis and thrombosis (Hynes *et al.*, 1984; Ruoslahti, 1984, 1988; Mosher, 1989; Couchman *et al.*, 1990; Hynes, 1990; Schwarzbauer, 1991). FN binds to fibrin, collagens, gelatin, DNA, integrins, heparin and proteoglycans. It forms part of the extracellular matrix of many cell types, either as a fibrillar aggregate or as a secreted disulfide-linked dimer. In plasma, it is in dimeric form consisting of two 220 kDa monomers joined by a disulfide bond at the C-terminus. FN has a modular architecture: each monomer consists of homologous modules of 40–90 amino acids, classified as type I, II or III repeats (Petersen *et al.*, 1983; Ruoslahti, 1988). The type III repeats are the most common in FN (Figure 1A), and are found in many other proteins. Approximately 2% of all animal proteins contain this repeat (Bork and Doolittle, 1992, 1993); extracellular proteins, intracellular proteins, membrane spanning cytokine receptors, growth hormone receptors and adhesion

molecules all possess type III-like repeats within their architecture (Patthy, 1990; Engel, 1991; Hansen *et al.*, 1995). Previous structural analyses have shown that the type III repeats form a β -sandwich with four β -strands on one side and three on the other (Leahy *et al.*, 1992, 1996; Main *et al.*, 1992; Dickinson *et al.*, 1994). The core structure of the type III repeats is conserved, with differences in function attributed to key residues, often located in the loop regions. Due to their abundance, these repeats are involved in a plethora of binding interactions, and structural elements responsible for these interactions continue to be mapped. FN offers a variety of sites for interaction with other molecules rendering the relative arrangement of the individual repeats bearing these sites important for matrix assembly.

Integrins are a family of cell surface receptors that play a key role in many cell adhesion events (Hynes, 1992). The type III repeat 10 of FN (FN10) has been shown to bind integrins via an Arg-Gly-Asp (RGD) sequence (Pierschbacher and Ruoslahti, 1984). In the crystal structure of FN7–10, FN9 presents a synergy site (PHSRN) for additional interactions with integrins on the same face of FN as the RGD segment (Obara *et al.*, 1988; Leahy *et al.*, 1996). Aside from FN–integrin interaction, FN is involved in heparin binding, which primarily (though not solely, see Figure 1A) occurs via repeats FN12–14 (Benecky *et al.*, 1988; Ingham *et al.*, 1990; Barkalow and Schwarzbauer, 1991). Heparin is abundant in animal tissues (in the form of heparan sulfate proteoglycans), on cell surfaces and in the extracellular matrix. Heparin molecules are anionic polysaccharide chains consisting of alternating L-iduronic (Idu) and D-glucosamine sugars (GlcN) joined by an $\alpha(1\rightarrow4)$ glycosidic bond. Typically, a heparin disaccharide has three sulfate groups, on the 2-hydroxyl group of Idu and the 2-amino and 6-hydroxyl groups of GlcN. Heparin molecules adopt a right-handed helical shape and a typical heparin disaccharide is 8–8.7 Å in length (Mulloy *et al.*, 1993). Heparin-binding activity in FN has been proposed to act in concert with integrins to stimulate the formation of focal adhesions (Laterra *et al.*, 1983; Woods *et al.*, 1986; Haugen *et al.*, 1990). There are heparin- or glycosaminoglycan (GAG)-binding sites in numerous other proteins, for example antithrombin III (Mourey *et al.*, 1993), lipoprotein lipase (Hata *et al.*, 1993) and fibroblast growth factor (Faham *et al.*, 1996), all characterized by peptides or regions of protein surface with an overall positive charge.

A 30 kDa heparin-binding fragment of FN containing FN12–14 (also called Hep-2) has been isolated (Benecky *et al.*, 1988; Ingham *et al.*, 1990; Barkalow and Schwarzbauer, 1991) and shown to support heparin-dependent adhesion of melanoma and neuroblastoma cells (McCarthy *et al.*, 1988; Drake *et al.*, 1993). Fluorescence polarization studies suggest the presence of two heparin-

Fig. 1. (A) The modular architecture of human FN showing the type I, II and III repeats and the known binding sites for various ligands. Boxes in green and blue correspond to regions where structural information is available (Williams *et al.*, 1994; Leahy *et al.*, 1996; Sticht *et al.*, 1998).

(B) A structure-based sequence alignment of 17 type III human FN repeats. Secondary structure classification is based on all repeats in FN12-14; FNEDA and FNEDB denotes the extra repeats A and B. Residues in green are identical in all 17 repeats, in aquamarine are identical in at least 12 repeats, and in yellow identical in 8-11 repeats. Residues in red are implicated in integrin binding. The first nine residues of FN15 (shaded) are part of IIICS (see text). All molecular graphic figures are generated using BOBSRIPT (Esnouf, 1997) and RASTER3D (Merritt and Murphy, 1994), or GRASP (Nicholls *et al.*, 1991).

binding sites (Ingham *et al.*, 1990); the major heparin-binding site is reported to be in FN13, with possible contributions from FN14 (Ingham *et al.*, 1990; Barkalow and Schwarzbauer, 1991). Heparin binding to FN12-14 has two distinct dissociation constants of 10^{-8} and 10^{-7} M, and is largely based on electrostatic forces (Bentley *et al.*, 1985; Beneky *et al.*, 1988). In addition to the known heparin-binding capability of FN12-14, several reports have indicated that it may possess binding sites for integrins (Mould and Humphries, 1991; Mohri *et al.*, 1996a,b). Cell adhesion to this segment involves the

synergistic interaction of cell surface integrin $\alpha_4\beta_1$ and heparan sulfate proteoglycans (Mould *et al.*, 1994). Furthermore, FN12-14 is involved in FN fibrillogenesis (Bultmann *et al.*, 1998), and binds HIV-1 gp120/160, thereby reducing virus infectivity (Bozzini *et al.*, 1998). FN also exists in a splice variant form in which a domain (called IIICS or V) of up to 120 amino acids is inserted between FN14 and FN15 (Kornbliht *et al.*, 1985).

To determine the three-dimensional structure of the major heparin binding repeats of FN, we made a protein over-expression construct spanning repeats FN12-15.

Table I.**A. Protein crystal analysis**

Crystal space group	I222		
Cryo unit cell (native) (Å)	$a = 68.6$	$b = 86.3$	$c = 142.8$
Cryo unit cell (SeMet crystals) (Å)	$a = 68.2$	$b = 86.8$	$c = 141.6$
Crystallization buffer	0.1 M HEPES pH 7.5, 0.8 M Na, K Tartrate		
Cryo-cooling buffer	crystallization buffer with either 35% ethylene glycol (native) or 35% glycerol (SeMet)		

B. Diffraction data collection

	Native	SeMet- λ_1	SeMet- λ_2	SeMet- λ_3	SeMet- λ_4
Wavelength (Å)	0.87	0.9787	0.9788	0.9115	0.9796
Reflections	10 582	5479 ^a	5479 ^a	5479 ^a	5479 ^a
R_{merge} (%)	4.5 (24.2)	8.0 (27.7)	9.5 (39.8)	8.2 (30.5)	8.7 (36.4)
Resolution (Å)	15–2.8	15–3.5	15–3.5	15–3.5	15–3.5
Completeness (%)	99 (99)	100 (100)	100 (100)	100 (100)	100 (100)

C. Model refinement statistics

Resolution range (Å)	15–2.8
Total reflections	10424
R_{cryst} (%)	21.6
R_{free} (%)	29.5
Protein atoms	2074
Water molecules	241
R.m.s.d. bonds (Å)	0.014
R.m.s.d. angles (°)	1.8

In (C), $R_{\text{cryst}} = \Sigma[(F_{\text{obs}}) - (F_{\text{calc}})]/\Sigma(F_{\text{obs}})$, and R_{free} is same as R_{cryst} but comprises a test set (5% of total reflections which was not used in the model refinement).

In (B), $R_{\text{merge}} = \Sigma(I - \langle I \rangle)/\Sigma \langle I \rangle$, and values in parentheses correspond to the highest resolution shell.

^aKeeping Bijvoet pairs separate.

R.m.s.d., root mean square deviation.

Purified protein gave crystals readily, and the structure was determined using multiple anomalous dispersion (MAD) phasing techniques (Hendrickson, 1991). Although present in the crystals (as confirmed by both gel electrophoresis of protein crystals and by mass spectroscopy; data not shown), no ordered protein density corresponding to FN15 was visible, and therefore the atomic model presented spans FN12–14. The model contains 269 amino acids and for simplicity (since there is no consensus view between the database entries for FN sequences) we number the residues sequentially from the first amino acid in the construct (Ala1). This start point corresponds to residue Ala1721 in Swiss-Prot accession No. P02751. This provides an atomic resolution view of heparin-binding sites in FN12–14. Prompted by the structural observations, additional mutagenesis experiments were performed, identifying a new integrin-binding motif in FN14. The structure presented provides a platform for dissecting the atomic basis of interaction between FN12–14 and heparin and integrins.

Results and discussion

Structure determination

For protein over-production, a cDNA sequence spanning FN12–15 for human FN was introduced into a bacterial over-expression vector. The expressed protein was purified using conventional chromatographic methods (see Materials and methods). Protein crystals were grown by vapour diffusion; protein (4–6 mg/ml in 10 mM Tris pH 7.4, 150 mM NaCl) and mother liquor (0.8 M Na,

K Tartrate, 0.1 M HEPES pH 7.5) were mixed in equal ratios and equilibrated over the mother liquor solution at room temperature (RT). Crystals (space group I222, $a = 68.6$ Å, $b = 86.3$ Å, $c = 142.8$ Å, one molecule in the asymmetric unit, solvent content = 49%) grew within a week. Crystallization of FN12–14 and FN12–15 has been reported independently (Pechik *et al.*, 1993).

The wild-type protein crystals diffracted to 2.8 Å under cryo-cooled conditions in a synchrotron beam, whereas selenomethionyl (SeMet)-substituted protein crystals diffracted to 3.1 Å under similar conditions (Table I). For structure determination, SeMet was added during the growth of the bacterial cultures (Hendrickson *et al.*, 1990; Bottomley *et al.*, 1994) and crystals of SeMet-FN12–15 were grown under similar conditions as wild-type crystals. MAD data were collected at the ESRF BM14 facility, processed with DENZO (Otwinowski and Minor, 1993), and phases were calculated using SHARP (La Fortelle *et al.*, 1997). Histogram matching, solvent flattening and phase-extension techniques using programs (DM and SOLOMON) in the CCP4 package (CCP4, 1994) yielded an excellent electron-density map at 3.5 Å resolution, into which a model was built using the program 'O' (Jones and Kjeldgaard, 1991). A total of 269 ordered amino acid residues for FN12–14 were successfully phased using the anomalous scattering from two SeMet atoms. No electron density was found for the C-terminal repeat (FN15). Electron density was weak for residues 76–79, 131–132 and 145, and the model is imperfect in these regions. Refinement used X-PLOR (Brunger, 1992) and CNS (Brunger *et al.*, 1998), yielding a final R -factor of 21.6%

Fig. 2. Orthogonal views of FN12–14 structure. The structure is represented as a cartoon with the β -strands (defined by inspection) shown as arrows. The chain is coloured from blue to green to red as the polypeptide goes from the N- to the C-terminus. The repeats correspond to FN12, FN13 and FN14. Individual strands are labelled in one repeat in each case.

Table II. Modular architecture of FN7–10 and FN12–14

	Tilt (°)	Twist (°)	Buried area (Å ²)	Linker
FN7–8	50	105	610	AVPPP
FN8–9	48	157	650	GLDSP
FN9–10	23	–38	340	VSDVP
FN12–13	62	312	450	LENVSP
FN13–14	51	147	660	<u>ALDAP</u>

Tilt and twist angles are defined as (Bork *et al.*, 1996) and all angles are calculated based on the conserved TRP residues which are used to define the repeat axes (Bork *et al.*, 1996). The buried surface area is calculated using GRASP (Nicholls *et al.*, 1991). Linker sequences between FN repeats are shown with solvent inaccessible residues underlined.

($R_{\text{free}} = 29.5\%$) for all data from 15 to 2.8 Å (Table I). There were no outliers in the Ramachandran plot. The final FN12–14 model contained residues Pro3–Glu92 of FN12, Asn93–Thr181 of FN13 and Ala182–Thr271 of FN14.

Overall structure of FN12–14

FN is a large, flexible molecule composed of many domains linked together in a modular fashion. Figure 1 summarizes the organization and properties of these repeats, and presents a structure-based sequence alignment of several type III repeats. FN12–14 consists of three type III repeats joined together in a slight spiral. As shown in Figure 2, the approximate length of FN12–14 ($3C_{\alpha}$ – $271C_{\alpha}$) is 100 Å, and each repeat has dimensions of $\sim 15 \times 20 \times 36$ Å. Individual repeats are the expected seven stranded β -sandwich (Figure 2). Each repeat has a conserved structural framework with a defined hydrophobic core. The β -sandwich structure does not have any disulfide bonds and structural variability resides primarily in the loop and linker regions. Each repeat contains the typical ‘wide’ β -bulge at the beginning of strand A (Chan *et al.*, 1993), and structurally homologous core residues

(Figure 1B). FN12 and FN14 both form strong interactions in the crystal lattice.

FN12 and the FN12–13 linker region

FN12 has no distinct charge character (Figure 3A). Like other type III repeats, it has conserved residues which make the hydrophobic core: Pro5, Trp22 and Tyr32 form the top of the core and pack against each other, whereas Leu62 and Tyr68 form the bottom. FN12 is connected to FN13 by numerous inter-repeat contacts between residues from strands A and B of FN12 (including the hairpin) and residues 167–170 of the FG turn in FN13. The linker residues Leu91, Glu92 and Asn93 are solvent-exposed, whereas Val94 is buried within the hydrophobic core of FN13. The linker region has a total buried surface area of 450 Å² and makes tilt and twist angles of ~ 62 and $\sim 312^\circ$, respectively (Table II). The limited buried surface area (compare with other linkers in Table II) is consistent with the cocked orientation of FN12 relative to FN13 and points to possible flexibility.

FN13 and the FN13–14 linker region

FN13 has a definite positive charge character (Figure 3A). The hydrophobic core consists of Pro97, Phe124 and Trp114 (top end), and Leu153 and Tyr159 (bottom end). There is a patch of hydrophobic residues which are partially exposed, contributed from strands F and G. This patch is next to Arg98, Arg99, Arg101, Arg115, Lys117 and Arg146, which are largely responsible for the overall positive charge character of this repeat. Residues Arg98, Arg99, Arg101 and Arg115 are solvent-exposed, do not form salt bridges or hydrogen bonds with other side chains, and form a charge tetrad with residue spacing reminiscent of the sulfate groups of heparin (Figure 3B).

Within the FN13–14 linker region, residues Ala182 and Ile183 are partially buried, the head group of Asp184 makes two hydrogen bonds to the BC loop of FN14, Ala185 is exposed, Pro186 is buried and Ser187 is exposed

(Figure 3C). Several residues from the FG loop of FN14 and the EF loop of FN13 are at the interface between FN13 and FN14. Gln260 and Lys257 point into the interface, and Pro155 from the EF loop of FN13 is close to Asp184. The numerous strong interactions at the FN13–14 interface imply a sturdy and inflexible linker. This region has a buried surface area of 660 Å², which is the highest in FN segment structures determined to date (see Table II).

FN14

FN14 also has an overall positive charge character (Figure 3A). The hydrophobic core of FN14 is formed from Pro186, Trp203 and Tyr213 (top end), and Leu243 and Tyr249 (bottom end). The main chain takes an unusual conformation in the C' strand of FN14 from Glu226 to Arg230. Within this structure, Val227 points to the hydrophobic core, residues Val228 and Pro229 point to the solvent, and the structure is further stabilized by a salt bridge between Glu226 and Arg230. It may be noted that the region LIGRKKKT (265–271) in FN14 does not form an α -helix-like conformation as previously predicted (Drake *et al.*, 1993). This region forms a typical β -strand (strand G) as part of the usual type III FN core structure.

FN15

The FN15 repeat is disordered and no model could be built for it. It is possible that this disorder is due to the absence of the splice repeat (IIICS/V) of up to 120 amino acid residues which is inserted between FN14 and FN15 in splice variants (Kornblihtt *et al.*, 1985). The sequence of FN15 in the absence of the splice domain (IIICS) lacks strand A of the canonical FN type III fold. There is strong evidence that proteins lacking the IIICS region occur naturally (Hershberger and Culp, 1990). Structural data cannot tell us how the FN15 is folded although there is clearly some flexibility. The biological significance of this phenomenon is not known.

Crystals can be grown using either a FN segment spanning FN12–14, or one spanning FN12–15 (Pechik *et al.*, 1993). These crystals have cell constants almost identical to one another and to the crystals used in this study, although each can be grown using different crystallization reagents. Analysis of the crystal packing of FN12–15 in this study reveals that there is a large hole at the end of FN14, which presumably can equally well contain FN15 (in the FN12–15 crystals) or simply solvent (FN12–14 crystals).

Heparin binding by FN12–14

Heparin–protein interactions are based on ionic contacts between positively charged residues (Arg, Lys) and the

negatively charged groups of heparin (sulfate and carboxylate groups). A wide variety of proteins are known to have heparin binding activity, and the binding sites have been structurally mapped for several proteins, e.g. fibroblast growth factor (DiGabriele *et al.*, 1998). The crystal structure presented here allows a detailed mapping of binding determinants based on a wealth of biochemical evidence for the heparin binding activity of human FN12–14.

Heparin-binding site 1

Residues Arg98, Arg99, Arg101, Arg115, Lys117 and Arg146 [this cluster will be referred to as heparin binding site 1 (HBS-1)] in FN13 have been implicated in heparin binding activity previously (Ingham *et al.*, 1990; Barkalow and Schwarzbauer, 1991; Busby *et al.*, 1995). These residues (contributed from strands A, B and E) project into the solvent from the ABE face of the β sheet. The congregation of six positively charged residues results in a continuous patch of positive charge (Figure 3A and B). Mutagenesis of Arg98 or Arg99 reduces heparin binding drastically (Barkalow and Schwarzbauer, 1991). Single amino acid mutation of any of the six residues to a non-charged residue also affects heparin binding substantially (Busby *et al.*, 1995). Although both Lys117 and Arg146 contribute to the patch of positive charge density, they are involved in side-chain contacts with other solvent exposed residues. Arg146 is slightly removed from the other residues within HBS-1 in line with the mutagenesis data where change of Arg146 to Ser reduces heparin binding only 3-fold, whereas mutation of any of the other five residues reduces it at least 10-fold (Busby *et al.*, 1995). Arg139 lies on the opposite side (GFCC' face) of FN13 relative to HBS-1, and its mutagenesis to Ser has no effect on heparin binding (Busby *et al.*, 1995). In addition, residues in HBS-1 are conserved in FNs from man to frog (Hynes *et al.*, 1984; DeSimone *et al.*, 1985; Norton and Hynes, 1987). Thus the crystallographic, mutagenesis and sequence conservation data together pinpoint the heparin-binding site in FN13 to HBS-1.

Putative heparin-binding site: HBS-2

Biochemical evidence suggests that both FN13 and FN14 are required for complete heparin binding (Ingham *et al.*, 1990; Barkalow and Schwarzbauer, 1991). Peptides from FN14 spanning residues 204–210, 217–235 and 257–271 have been implicated independently (McCarthy *et al.*, 1988, 1990; Drake *et al.*, 1993; Woods *et al.*, 1993). The structure reveals the presence of a cluster of positively charged residues in FN14 suggestive of a minor heparin-

Fig. 3. (A) Surface representations of FN12–14. Top left: GRASP representation of FN12–14, with a backbone worm (green) overlaid with all side chains (yellow) and the surface potential (semi-transparent). The view is almost identical to the top panel of Figure 2. The four lower panels show the potential surface, with the first view corresponding to the topmost image and descending frames corresponding to successive 90° rotations about a horizontal axis. The right hand panels show corresponding views of a GRASP molecular surface representation of FN12–14 (in grey). The coloured patches correspond to putative binding sites for heparin and integrin, as defined in the text. HBS-1 is shown in blue, HBS-2 in green, the IDAPS motif in red and the PRARI sequence in yellow. Note the opposition of the heparin- and integrin-binding surfaces. (B) A stereo view of the basic residues that form, from left to right, HBS-1 (FN13) and HBS-2 (FN14) [the view corresponds to the penultimate view in (A)]. The backbone is shown in cyan and the side-chains of the basic residues in dark blue. The distances between terminal amide groups of Arg98, Arg99, Arg101 and Arg115 are: Arg98–Arg99, ~13 Å; Arg98–Arg115, ~8 Å; Arg99–Arg115, ~7 Å; Arg99–Arg101, ~11 Å; Arg101–Arg115, ~18 Å; and Arg98–Arg101, ~17 Å. Each heparin-binding site is composed of a cluster of solvent-exposed positively charged residues presented on the same face of the molecule. The maximum distance from HBS-1 to HBS-2 is ~60 Å. (C) A stereo view of Asp184 which hydrogen bonds to the BC loop (PRARI) of FN14. The protein backbone is shown in green and residues Asp184, Arg207 and Arg209 (the arginines of the PRARI motif) are shown as ball-and-stick. Mutagenesis data (Figure 4) implicates both arginines in integrin binding.

binding site. This cluster (HBS-2) includes Lys216, Lys219, Arg225, Arg230 and Lys261. Since HBS-1 is on the same face of the molecule as HBS-2 (Figure 3), heparin binding may traverse FN13–14.

Binding studies between FN12–14 and oligosaccharides

of various lengths, show that 12–16 saccharide units (~51–69 Å) are required for optimal binding (Walker and Gallagher, 1996). The maximum distance from HBS-1 (Arg98) to HBS-2 (Arg219) is ~60 Å, providing support for the hypothesis of extended heparin binding to FN12–14.

Table III. R.m.s.ds (Å) and sequence identity (%)

	FN8	FN9	FN10	FN12	FN13	FN14
FN7	1.2 (84) 26	1.1 (85) 30	1.1 (82) 29	1.2 (86) 30	1.4 (85) 26	1.2 (84) 26
FN8	—	0.9 (86) 34	1.0 (86) 20	1.0 (85) 34	1.3 (86) 23	1.0 (86) 31
FN9	—	—	1.0 (85) 25	1.1 (85) 30	1.4 (86) 26	0.9 (86) 39
FN10	—	—	—	1.2 (82) 28	1.3 (82) 28	1.1 (86) 38
FN12	—	—	—	—	1.1 (85) 24	1.0 (84) 31
FN13	—	—	—	—	—	1.1 (85) 29

For each comparison the upper numbers indicates the r.m.s.ds between equivalent C α residues with the number of structurally equivalent residues shown in parentheses. The lower number is the percentage sequence identity. All calculations done using SHP (Stuart *et al.*, 1979).

Integrin $\alpha_4\beta_1$ binding by FN12–14

There have been several reports of binding between FN12–14 and various integrins (Mohri *et al.*, 1996a,b; Mould and Humphries, 1991). Functional data suggest that FN12–14 is able to promote melanoma cell adhesion activity and is sensitive to anti- α_4 and anti- β_1 antibodies, suggestive of direct interaction between FN12–14 and $\alpha_4\beta_1$ (Mould and Humphries, 1991). Similar results were obtained with hematopoietic cell binding to FN12–14 (Mohri *et al.*, 1996a). A recent report suggests that a peptide in FN13 spanning residues 105–119, which form the β -strands A and B, binds to the integrin $\alpha_{IIb}\beta_3$ (Mohri *et al.*, 1996b). From the crystal structure of FN12–14 it is evident that this region not only overlaps with the heparin-binding site (HBS-1), but also contains the two residues (Arg115 and Lys117) from HBS-1 which have been unambiguously implicated in heparin binding. It is unlikely that the integrin $\alpha_{IIb}\beta_3$ shares a binding site with heparin, and clarification of these results requires mutagenesis and binding studies with whole segments of FN repeats.

Using a peptide-based approach, the IDAPS (183–187) sequence in the FN13–14 linker has been suggested as being responsible for integrin binding, and in particular, residue Asp184 (Mould and Humphries, 1991). From the crystal structure of FN12–14 it is clear that Asp184 (Asp3 of FN14; see Figure 1B) pins down the BC loop from FN14 (Figure 3C). Structurally, this is a similar interaction to the one that Asp3 (FN9) makes at the FN8–9 interface (Leahy *et al.*, 1996). Indeed the linker sequence AIDAP (FN14) is structurally homologous to the linker sequence GLDSP (FN9). In both, the residue preceding the aspartate is either a leucine or an isoleucine that is buried in a small hydrophobic core. The other residues in each linker are exposed to solvent, except for the proline, which forms the start of strand A in the next repeat. This structural similarity suggests that linkers like AIDAP and GLDSP are a structural feature of some interfaces in type III segments. In the two cases where structural information is available, these linkers adopt similar conformations and appear rigid, forming a relatively tight inter-repeat interface (see Table III).

The above structural considerations are suggestive of an indirect role for the IDAPS motif in integrin binding. Inspection of the FN14 BC loop sequence that hydrogen bonds to Asp184 (IDAPS) reveals an interesting similarity

Fig. 4. Solid-phase assay to measure the binding of wild-type FN12–14 (WT) and its mutants, R207A (PRARI), R209A (PRARI), and D184E (IEAPS), to immobilized $\alpha_4\beta_1$ integrin. The drastic effect of D184E mutation points to the importance of Asp184 in pinning down the BC loop of FN14. Mutation of either arginine in this loop (PRARI) also affects integrin binding (see text).

to the previously determined synergy sequence in FN9, even though they are in different regions of the repeat. In FN14 the sequence is PRARI, while in FN9 it is PHSRN. The latter has been shown to interact directly with the integrins $\alpha_5\beta_1$ and $\alpha_{IIb}\beta_3$ in synergy with the integrin-binding RGD in FN10 (Obara *et al.*, 1988; Aota *et al.*, 1994; Bowditch *et al.*, 1994; Leahy *et al.*, 1996). PHSRN relies on its arginine for synergy activity. We therefore hypothesized that PRARI might be the site within FN12–14, recognized by integrin $\alpha_4\beta_1$, and the reason disruption of Asp184 inhibits integrin binding to FN12–14 is that it perturbs the structure of the PRARI motif (Figure 3C).

The apparent indirect role of Asp184 and the IDAPS sequence in forming the integrin-binding site in FN14 raises questions about why this sequence in peptide form is able to block cell adhesion (Mould and Humphries, 1991). We propose that this is due to IDAPS acting as a mimic of the integrin-binding LDVPS site in the IIICS

Fig. 5. (A) Orthogonal views of the superposition of FN8-9 on FN13-14. Repeats FN7-10 are shown in red and FN12-14 in green. Note that the relative disposition of repeats FN8-9 is almost identical to that in FN13-14, whereas the relative disposition of FN7 compared with FN8-9 is markedly different to that of FN12 relative to FN13-14. (B) A stereo image of the superposition of the backbone trace of all of the individual type III repeats of FN7-10 and FN12-14: FN7 (orange), FN8 (purple), FN9 (magenta), FN10 (red), FN12 (cyan), FN13 (blue) and FN14 (green). (C) A GRASP representation of the surface potential of FN7-10 showing the negative charge character on one face. The integrin binding RGD peptide in FN10 is the protruding loop, coloured blue, and is well separated from the regions of negative charge.

region (Figure 1A). IDAPS has ~20-fold lower activity in cell adhesion assays than LDVPS-containing peptides, and therefore it could be acting as a weak analogue and binding to the LDVPS-binding site on the integrin. Thus,

even though (i) IDAPS peptides are able to block adhesion (Mould and Humphries, 1991), (ii) mutation of the D of IDAPS abolishes integrin binding (Figure 4), and (iii) antibodies that bind close to IDAPS perturb integrin

binding (M.J.Humphries, unpublished data), its role appears to be indirect. These findings further emphasize the need to be extremely careful in interpreting the effects of synthetic peptides, particularly when they are examples of a common motif, such as RGD or LDV.

Site-directed mutagenesis of each arginine in the PRARI sequence was performed, together with a conservative mutation of IDAPS to IEAPS, and the mutant proteins examined for integrin-binding activity. As shown in Figure 4, mutation of either arginine to alanine caused a substantial loss of integrin-binding activity as measured in an isolated receptor protein–protein assay. This indicates that the PRARI sequence is involved in integrin binding and is a new integrin-binding motif. Sequence comparison data indicate that the PRARI motif is essentially conserved across all FN species, from man to frog (in zebrafish FN the PRARI sequence is replaced by PRSTI; however, it is not known whether this domain binds integrin). It is tempting to speculate that PHSRN/PRARI sequences may be commonplace integrin-binding motifs. A number of observations lend support to this hypothesis.

A recent report implicates residues KLDAPT from the FN5 repeat in binding to activated $\alpha_4\beta_1$ and $\alpha_4\beta_7$ integrins (Moyano *et al.*, 1997). From the structure-based sequence alignment shown in Figure 1B, it seems likely that the sequence KLDAPT forms the linker region between FN4–5 (see Table II). The residues LD may be buried [as seen for the ID in AIDAP (FN14) and LD in GLDSP (FN9)], requiring drastic conformational changes to expose the residues for integrin binding. We consider such changes unlikely, and by analogy with FN13–14 suggest that LDAPT interacts with the BC loop of FN5, which may itself contain the integrin recognition site. Inspection of the sequence of this loop in human FN reveals the sequence of this site to be PRAQI, similar to the PRARI motif of FN14. Furthermore, in FN9 the GLDSP sequence hydrogen bonds to the BC loop sequence PRATI (Leahy *et al.*, 1996); it is possible that this sequence might synergize with RGD in integrin binding. Finally, a sequence reminiscent of PRARI (RQR) has been identified as an integrin-binding peptide using panning of phage display libraries on purified $\alpha_v\beta_5$ (Koivunen *et al.*, 1995). Mutagenesis can now be used to explore the possibility that there are additional integrin-binding PRARI-type motifs in FN.

FN architecture: implications for ligand binding

The individual repeats within FN7–10 and FN12–14 are very similar in terms of the core architecture (see Figure 5; Table II). A superposition of all currently available type III FN repeats is shown in Figure 5B. The relative orientation of various repeats, classically expressed in terms of tilt and twist angles (Bork *et al.*, 1996), represents an important parameter in the overall architecture of FN. As seen in Table II, these angles vary drastically, so that depending on the actual combination employed, the intact FN molecule may adopt a variety of shapes from linear to circular. The buried surface area at the inter-repeat interface varies from 340 to 660 Å², indicating that some interfaces (e.g. FN7–8 and FN12–13) are essentially flexible whereas others (e.g. FN8–9 and FN13–14) are relatively rigid units. In particular, the buried surface area, tilt and twist angles are essentially identical between

FN8–9 (Leahy *et al.*, 1996) and FN13–14 (Table II; Figure 5A). Conversely, when FN8–9 is superimposed on FN13–14, it is apparent that the repeats FN7 and FN12 are cocked in almost opposite orientations (Figure 5A). It is difficult to predict the absolute orientation of FN7–10 and FN12–14 without the structure of intervening (and overlapping) repeats; however, it is evident that inclusion of a few of the more extreme inter-repeat orientations (e.g. FN7–8 and FN13–14) would allow a distinctly non-linear form of FN. Electron microscopic studies indicate that FN takes a compact conformation at low ionic strength and an extended conformation in higher ionic strength solutions; the latter conformation indeed has several predicted bend points (Engel *et al.*, 1981; Erickson *et al.*, 1981; Erickson and Carrell, 1983; Rocco *et al.*, 1983; Lai *et al.*, 1993).

It is probable that there are several pivot points for deformation along the intact FN molecule. Global flexibility of FN may be achieved via the less rigid linker regions whilst maintaining precise relative orientations between other segments. The overall flexibility of FN is clearly manifest without any ligand binding, and it may be a necessary feature of molecules which form the extracellular matrix in cellular systems.

Electrostatics calculations show that the FN7–10 repeats have a definite overall negative charge character (see Figure 5C and compare with 3A). This electrostatic character is in contrast to that of FN12–14. Clearly, intact FN is distinctly polarized. This charge segregation may be critical in orienting and docking the integrin and heparin ligands.

Intact FN has binding surfaces for several molecules, including collagens, fibrin, integrins, heparin, DNA etc. *In vivo*, it may be occupied at multiple sites by multiple ligands. The heparin-binding region (FN12–14) falls in the middle of two other integrin-binding regions (Figure 1A). Whether integrin and heparin are bound simultaneously or alternatively is not yet known, although it is interesting to note that the proposed binding sites for these two ligands lie on opposite faces of FN13–14 repeats (Figure 3A). In addition, both the integrin-binding site, at the junction between the two domains, and the heparin-binding sites, which span both repeats, are likely to require a precise juxtaposition of FN13 and FN14 for binding. Loading or unloading heparin may facilitate integrin binding or vice versa. These possibilities will be addressed by more extensive biochemical and structural analyses of FN and its ligands.

Materials and methods

Native and SeMet protein production

Human FN cDNA encoding repeats FN12–15 was sub-cloned into a pGEX expression vector and the construct transformed into *Escherichia coli* strain BLR(DE3)pLysS for expression. Purification of the wild-type and SeMet followed similar protocols. For production of SeMet protein, bacteria were grown in a defined medium (Hendrickson *et al.*, 1990) without methionine and cysteine. A 10 ml culture was inoculated from frozen stocks and grown for 20 h at 37°C. Protein expression was induced with addition of 0.1 mM IPTG and cell growth continued for 4 h. Cells were harvested by centrifugation, the cell pellet was resuspended in 20 ml Tris-buffered saline (TBS) containing 0.1% Triton, 1 mM PMSF and 20 µg/ml leupeptin, and frozen. On thawing the (now) lysed cells, RNase and DNase (500 µg each) were added and incubated at 4°C. The cell debris was removed by centrifugation and the supernatant

added to 5 ml of pre-swollen glutathione-agarose matrix in TBS (10 mM Tris pH 7.4, 150 mM NaCl) buffer. After incubation the unbound proteins were removed by centrifugation. The beads were washed three times in TBS and twice in a low salt buffer (50 mM NaCl, 10 mM Tris pH 7.4) and suspended in 5 ml of the same buffer. A total of 25 U of human thrombin (Sigma) were added and the beads rotated for 2 h at RT to allow FN12–15 to be cleaved away from glutathione *S*-transferase (GST) protein. The supernatant was then removed and purified further by purification over a heparin-agarose column. The resulting protein was concentrated to 6 mg/ml in a centricon and stored frozen at -20°C .

Crystallization

Crystals of FN12–15 were grown at 20°C using the vapour diffusion method and 0.1 mM HEPES pH 7.5, 0.8 M Na, K Tartrate as reservoir solution. Typically, a 4–6 mg/ml protein solution was mixed in a 1:1 ratio with reservoir solution; protein crystals grew within a week. Crystals of the SeMet-FN12–15 were grown under identical conditions. The typical crystal dimensions are $0.3 \times 0.4 \times 0.5 \text{ mm}^3$. The wild-type crystals diffracted to 2.8 Å in a synchrotron beam and the SeMet-FN12–15 crystals to 3.1 Å.

Data collection

Diffraction data were collected using synchrotron radiation at the ESRF facility (MAD data) or at Daresbury (native data; Table I). Crystals were cryo-cooled by freezing rapidly in mother liquor and 35% glycerol (in the case of SeMet-FN12–15 crystals), or in 35% ethylene glycol (in the case of wild-type crystals). The structure was solved with MAD data which were collected on the SeMet-FN12–15 crystals at beam line BM14 at the ESRF. The Se-absorption edge was determined by a fluorescence scan and four wavelengths were selected for data collection corresponding to the peak (maximum f'' , $\lambda = 0.978732 \text{ Å}$; inflection point, $\lambda = 0.978895 \text{ Å}$; and remotes $\lambda = 0.911586 \text{ Å}$, $\lambda = 0.979673 \text{ Å}$). A total of 180° of data were collected in 1.0° oscillation frames for each wavelength using a MAR345 scanner. For model refinement, a higher resolution native data set was collected at the Daresbury synchrotron beam line 9.6 in 1.0° oscillation steps using the MAR345 scanner. All data were indexed, measured, and scaled with DENZO and SCALEPACK programmes (Otwinowski and Minor, 1993), and subsequently processed with the CCP4 package (CCP4, 1994). Relevant statistics are given in Table I.

Structure determination

The program REVISE (CCP4, 1994) was used to estimate the normalized anomalous scattering magnitude. Modified MAD data from REVISE were used in Patterson analyses and the positions of two SeMet atoms determined. These were confirmed by isomorphous difference Patterson maps. The programme SHARP (La Fortelle *et al.*, 1997) was used for refinement of these sites and phases calculated in the 15–3.5 Å range. The program DM (CCP4, 1994) was used for envelope determination, histogram matching and solvent flattening, which was followed by cycles of phase extension and solvent flipping using the program SOLOMON (CCP4, 1994). The final 3.5 Å map showed unambiguous electron density for three FN repeats. The molecular graphics program 'O' (Jones and Kjeldgaard, 1991) was used for model building.

Refinement and structure analysis

Crystallographic refinement of the structure was carried out using the program X-PLOR (Brunger, 1992) and CNS (Brunger *et al.*, 1998). A C_{α} trace of the model was fitted into the original MAD derived electron density map. Rigid body refinement of the whole molecule and of individual repeats was done using data in the resolution range of 15–3.5 Å. Side chains were built in using $2F_o - F_c$ and original MAD maps, followed by positional refinement to 2.8 Å. After several cycles of refinement, individual b-factors for all atoms were refined, followed by the addition of water molecules. A total of 5% of the native data were reserved for R_{free} analysis. Anisotropic diffraction and bulk solvent corrections were carried out for all data from 15–2.8 Å resolution, and the model was refined against these data. The final model statistics are given in Table I; there are no non-glycine residues in disallowed regions of the Ramachandran plot, as confirmed using the programme PROCHECK (Laskowski *et al.*, 1993). Atomic coordinates for FN12–14 have been deposited with the Protein Data Bank, Brookhaven National Laboratory, CT, accession No. 1FNH.

Integrin binding assay and FN12–14 mutagenesis

Purified $\alpha_4\beta_1$ from Molt-4 cells, prepared as reported previously (Newham *et al.*, 1998), were diluted 1:100 in phosphate-buffered saline

and 100 μl aliquots were added to a 96-well Immulon-4 plate (Dynatech). Plates were incubated overnight at 4°C and wells were then blocked with 200 μl of 5% bovine serum albumen (BSA), 150 mM NaCl, 10 mM Tris pH 7.4, for 2 h at RT. The wells were washed twice with 200 μl of 150 mM NaCl, 25 mM Tris pH 7.4, 1% BSA, 1 mM MnCl_2 and 1 mM MgCl_2 (buffer A) before adding 100 μl aliquots of 0.5–50 $\mu\text{g}/\text{ml}$ of biotinylated protein diluted in the same buffer. Proteins were biotinylated by mixing with equal amounts of sulfo-*N*-hydroxysuccinimido biotin and rotary mixing for 30–40 min at RT. The mixture was then dialyzed against several changes of 150 mM NaCl, 25 mM Tris, pH 7.4 to remove excess biotin. Plates were incubated for 3 h at 37°C to allow binding to occur. The ligand was aspirated, the wells washed with buffer A and bound ligand quantitated by the addition of 100 μl of 1:500 extravidin–peroxidase conjugate (Sigma) diluted in buffer A for 30 min at RT. Wells were washed as before, and binding was visualized by the addition of a colour substrate (ABTS, Sigma), followed by reading at 405 nm. The level of non-specific binding was measured for each protein by determining the level of binding to wells coated with BSA alone. Results are expressed with the background levels subtracted. The level of biotinylation of each FN12–15 protein was determined following immobilization in microtitre wells and sequential incubation with avidin–peroxidase and a colour substrate (under conditions of linear colour development). The experimental data for integrin binding were then corrected for the small variations in biotinylation that were detected. The binding of the FN12–15 protein to purified $\alpha_4\beta_1$ has been shown previously to be completely sensitive to co-incubation with EDTA or anti- α_4 and β_1 monoclonal antibodies (Mould *et al.*, 1994). Mutagenesis was carried out using a standard protocol (Kunkel, 1985): IDAPS to IEAPS: 5'-ACT GCC ATT GAA GCA CCA TCC-3'; PRARI to PAARI: 5'-GGC AGC CGC CAG CTG CCA GGA TTA CC-3'; PRARI to PRAAI: 5'-CCG CCA CGT GCC GCG ATT ACC GGC TAC-3'.

Acknowledgements

We thank the ESRF and EMBL staff responsible for BM14 at the ESRF in Grenoble for their help and advice in data collection, and members of LMB, Oxford, UK, for various discussions. This work was funded by the MRC. A.S. is supported by a St John's College Junior Research Fellowship and by the MRC. D.I.S. and E.Y.J. are supported by MRC and Royal Society, respectively, and J.A.A. and M.J.H. are supported by the Wellcome Trust.

References

- Aota, S., Nomizu, M. and Yamada, K.M. (1994) The short amino acid sequence Pro-His-Ser-Arg-Asn in human fibronectin enhances cell-adhesive function. *J. Biol. Chem.*, **269**, 24756–24761.
- Barkalow, F.J. and Schwarzbauer, J.E. (1991) Localization of the major heparin-binding site in fibronectin. *J. Biol. Chem.*, **266**, 7812–7818.
- Benecky, M.J., Kolvenbach, C.G., Amrani, D.L. and Mossessan, M.W. (1988) Evidence that binding to the carboxyl-terminal heparin-binding domain (Hep II) dominates the interaction between plasma fibronectin and heparin. *Biochemistry*, **27**, 7565–7571.
- Bentley, K.L., Klebe, R.J., Hurst, R.E. and Horowitz, P.M. (1985) Heparin binding is necessary, but not sufficient, for fibronectin aggregation. A fluorescence polarization study. *J. Biol. Chem.*, **260**, 7250–7256.
- Bork, P. and Doolittle, R.F. (1992) Proposed acquisition of an animal protein domain by bacteria. *Proc. Natl Acad. Sci. USA*, **89**, 8990–8994.
- Bork, P. and Doolittle, R.F. (1993) Fibronectin type III modules in the receptor phosphatase CD45 and tapeworm antigens. *Protein Sci.*, **2**, 1185–1187.
- Bork, P., Downing, A.K., Kieffer, B. and Campbell, I.D. (1996) Structure and distribution of modules in extracellular proteins. *Q. Rev. Biophys.*, **29**, 119–167.
- Bottomley, M.J., Robinson, R.C., Driscoll, P.C., Harlos, K., Stuart, D.I., Aplin, R.T., Clements, J.M., Jones, E.Y. and Dudgeon, T.J. (1994) Crystallization and preliminary X-ray diffraction characterisation of both a native and selenomethionyl VLA-4 binding fragment of VCAM1. *J. Mol. Biol.*, **244**, 464–468.
- Bowditch, R.D., Hariharan, M., Tomina, E.F., Smith, J.W., Yamada, K.M., Getzoff, E.D. and Ginsberg, M.H. (1994) Identification of a novel integrin binding site in fibronectin. Differential utilization by β_3 integrins. *J. Biol. Chem.*, **269**, 10856–10863.
- Bozzini, S., Falcione, V., Conaldi, P.G., Visai, L., Biancone, L., Dolei, A., Toniolo, A. and Speziale, P. (1998) Heparin-binding domain of human

- fibronectin binds HIV-1 gp120/160 and reduces virus infectivity. *J. Med. Virol.*, **54**, 44–53.
- Brunger, A.T. (1992) *X-PLOR, Version 3.1. A System for X-Ray Crystallography and NMR*. Yale University Press, New Haven, CT.
- Brunger, A.T. *et al.* (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr.*, **D54**, 905–921.
- Bultmann, H., Santas, A.J. and Peters, D.M. (1998) Fibronectin fibrillogenesis involves the heparin II binding domain of fibronectin. *J. Biol. Chem.*, **273**, 2601–2609.
- Busby, T.F., Argraves, W.S., Brew, S.A., Pechik, I., Gilliland, G.L. and Ingham, K.C. (1995) Heparin binding by fibronectin module III-13 involves six discontinuous basic residues brought together to form a cationic cradle. *J. Biol. Chem.*, **270**, 18558–18562.
- CCP4 (1994) Collaborative Computational Project Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr.*, **D50**, 760–767.
- Chan, A.W., Hutchinson, E.G., Harris, D. and Thornton, J.M. (1993) Identification, classification and analysis of β -bulges in proteins. *Protein Sci.*, **2**, 1574–1590.
- Couchman, J.R., Austria, M.R. and Woods, A. (1990) Fibronectin–cell interactions. *J. Invest. Dermatol.*, **94**, 7s–14s.
- de La Fortelle, E. and Bricogne, G. (1997) Maximum-likelihood heavy-atom parameter refinement in the MIR and MAD methods. In Sweet, R.M. and Carter, C.W., Jr (eds), *Methods in Enzymology, Macromolecular Crystallography*. Vol. 276. Academic Press, New York, NY, pp. 472–494.
- DeSimone, D.W., Spiegel, E. and Spiegel, M. (1985) The biochemical identification of fibronectin in the sea urchin embryo. *Biochem. Biophys. Res. Commun.*, **133**, 183–188.
- Dickinson, C.D., Veerapandian, B., Dai, X.P., Hamlin, R.C., Xuong, N.H., Ruoslahti, E. and Ely, K.R. (1994) Crystal structure of the tenth type III cell adhesion module of human fibronectin. *J. Mol. Biol.*, **236**, 1079–1092.
- DiGabriele, A.D., Lax, I., Chen, D.I., Svahn, C.M., Jaye, M., Schlessinger, J. and Hendrickson, W.A. (1998) Structure of a heparin-linked biologically active dimer of fibroblast growth factor. *Nature*, **393**, 812–817.
- Drake, S.L., Varnum, J., Mayo, K.H., Letourneau, P.C., Furcht, L.T. and McCarthy, J.B. (1993) Structural features of fibronectin synthetic peptide FN-C/H II, responsible for cell adhesion, neurite extension and heparan sulfate binding. *J. Biol. Chem.*, **268**, 15859–15867.
- Engel, J. (1991) Common structural motifs in proteins of the extracellular matrix. *Curr. Opin. Cell Biol.*, **3**, 779–785.
- Engel, J., Odermatt, E., Engel, A., Madri, J.A., Furthmayr, H., Rohde, H. and Timpl, R. (1981) Shapes, domain organizations and flexibility of laminin and fibronectin, two multifunctional proteins of the extracellular matrix. *J. Mol. Biol.*, **150**, 97–120.
- Erickson, H.P. and Carrell, N.A. (1983) Fibronectin in extended and compact conformations. Electron microscopy and sedimentation analysis. *J. Biol. Chem.*, **258**, 14539–14544.
- Erickson, H.P., Carrell, N. and McDonagh, J. (1981) Fibronectin molecule visualized in electron microscopy: a long, thin, flexible strand. *J. Cell Biol.*, **91**, 673–678.
- Esnouf, R.M. (1997) An extensively modified version of MolScript that includes greatly enhanced coloring capabilities. *J. Mol. Graph. Model.*, **15**, 132–134, 112–113.
- Faham, S., Hileman, R.E., Fromm, J.R., Linhardt, R.J. and Rees, D.C. (1996) Heparin structure and interactions with basic fibroblast growth factor. *Science*, **271**, 1116–1120.
- Hansen, L.K., O'Leary, J.J., Skubitz, A.P., Furcht, L.T. and McCarthy, J.B. (1995) Identification of a homologous heparin binding peptide sequence present in fibronectin and the 70 kDa family of heat-shock proteins. *Biochim. Biophys. Acta*, **1252**, 135–145.
- Hata, A. *et al.* (1993) Binding of lipoprotein lipase to heparin. Identification of five critical residues in two distinct segments of the amino-terminal domain. *J. Biol. Chem.*, **268**, 8447–8457.
- Haugen, P.K., McCarthy, J.B., Skubitz, A.P., Furcht, L.T. and Letourneau, P.C. (1990) Recognition of the A chain carboxy-terminal heparin binding region of fibronectin involves multiple sites: two contiguous sequences act independently to promote neural cell adhesion. *J. Cell Biol.*, **111**, 2733–2745.
- Hendrickson, W.A. (1991) Determination of macromolecular structures from anomalous diffraction of synchrotron radiation. *Science*, **254**, 51–58.
- Hendrickson, W.A., Horton, J.R. and LeMaster, D.M. (1990) Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. *EMBO J.*, **9**, 1665–1672.
- Hershberger, R.P. and Culp, L.A. (1990) Cell-type-specific expression of alternatively spliced human fibronectin IIICS mRNAs. *Mol. Cell. Biol.*, **10**, 662–671.
- Hynes, R. (1990) *Fibronectins*. Springer-Verlag, New York, NY.
- Hynes, R.O. (1992) Integrins: versatility, modulation and signaling in cell adhesion. *Cell*, **69**, 11–25.
- Hynes, R.O., Schwarzbauer, J.E. and Tamkun, J.W. (1984) Fibronectin: a versatile gene for a versatile protein. *Ciba Found. Symp.*, **108**, 75–92.
- Ingham, K.C., Brew, S.A. and Atha, D.H. (1990) Interaction of heparin with fibronectin and isolated fibronectin domains. *Biochem. J.*, **272**, 605–611.
- Jones, T.A. and Kjeldgaard, M. (1991) Improved methods for the building of protein models in electron density maps and the location of errors in these models. *Acta Crystallogr.*, **A47**, 110–119.
- Koivunen, E., Wang, B.C., Ruoslahti, E. (1995) Phage libraries displaying cyclic-peptides with different ring sizes—ligand specificities of the rgd-directed integrins. *Biotechnology*, **13**, 265–270.
- Kornblihtt, A.R., Umezawa, K., Vibe Pedersen, K. and Baralle, F.E. (1985) Primary structure of human fibronectin: differential splicing may generate at least 10 polypeptides from a single gene. *EMBO J.*, **4**, 1755–1759.
- Kunkel, T.A. (1985) Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl Acad. Sci. USA*, **82**, 488–492.
- Lai, C.S., Wolff, C.E., Novello, D., Griffone, L., Cuniberti, C., Molina, F. and Rocco, M. (1993) Solution structure of human plasma fibronectin under different solvent conditions. Fluorescence energy transfer, circular dichroism and light-scattering studies. *J. Mol. Biol.*, **230**, 625–640.
- Laskowski, R.A., MacArthur, M.W., Moss, D.S. and Thornton, J.M. (1993) PROCHECK: a program to check stereochemical quality of protein structures. *J. Appl. Crystallogr.*, **26**, 283–291.
- Lattera, J., Norton, E.K., Izzard, C.S. and Culp, L.A. (1983) Contact formation by fibroblasts adhering to heparan sulfate-binding substrata (fibronectin or platelet factor 4). *Exp. Cell Res.*, **146**, 15–27.
- Leahy, D.J., Hendrickson, W.A., Aukhil, I. and Erickson, H.P. (1992) Structure of a fibronectin type III domain from tenascin phased by MAD analysis of the selenomethionyl protein. *Science*, **258**, 987–991.
- Leahy, D.J., Aukhil, I. and Erickson, H.P. (1996) 2.0 Å crystal structure of a four-domain segment of human fibronectin encompassing the RGD loop and synergy region. *Cell*, **84**, 155–164.
- Main, A.L., Harvey, T.S., Baron, M., Boyd, J. and Campbell, I.D. (1992) The three-dimensional structure of the tenth type III module of fibronectin: an insight into RGD-mediated interactions. *Cell*, **71**, 671–678.
- McCarthy, J.B., Chelberg, M.K., Mickelson, D.J. and Furcht, L.T. (1988) Localization and chemical synthesis of fibronectin peptides with melanoma adhesion and heparin binding activities. *Biochemistry*, **27**, 1380–1388.
- McCarthy, J.B., Skubitz, A.P., Qi, Z., Yi, X.Y., Mickelson, D.J., Klein, D.J. and Furcht, L.T. (1990) RGD-independent cell adhesion to the carboxy-terminal heparin-binding fragment of fibronectin involves heparin-dependent and -independent activities. *J. Cell Biol.*, **110**, 777–787.
- Merritt, E.A. and Murphy, M.E.P. (1994) Raster3D version 2.0. A program for photorealistic plots of protein structures. *Acta Crystallogr.*, **D50**, 869–873.
- Mohri, H., Katoh, K., Iwamatsu, A. and Okubo, T. (1996a) The novel recognition site in the C-terminal heparin-binding domain of fibronectin by integrin $\alpha_4\beta_1$ receptor on HL-60 cells. *Exp. Cell Res.*, **222**, 326–332.
- Mohri, H., Tanabe, J., Katoh, K. and Okubo, T. (1996b) Identification of a novel binding site to the integrin $\alpha_{IIb}\beta_3$ located in the C-terminal heparin-binding domain of human plasma fibronectin. *J. Biol. Chem.*, **271**, 15724–15728.
- Mosher, D.F. (1989) *Fibronectin*. Academic Press, San Diego, CA.
- Mould, A.P. and Humphries, M.J. (1991) Identification of a novel recognition sequence for the integrin $\alpha_4\beta_1$ in the COOH-terminal heparin-binding domain of fibronectin. *EMBO J.*, **10**, 4089–4095.
- Mould, A.P., Askari, J.A., Craig, S.E., Garratt, A.N., Clements, J. and Humphries, M.J. (1994) Integrin $\alpha_4\beta_1$ -mediated melanoma cell adhesion and migration on vascular cell adhesion molecule-1 (VCAM-1) and the alternatively spliced IIICS region of fibronectin. *J. Biol. Chem.*, **269**, 27224–27230.

- Mourey,L., Samama,J.P., Delarue,M., Petitou,M., Choay,J. and Moras,D. (1993) Crystal structure of cleaved bovine antithrombin III at 3.2 Å resolution. *J. Mol. Biol.*, **232**, 223–241.
- Moyano,J.V. *et al.* (1997) Fibronectin type III5 repeat contains a novel cell adhesion sequence, KLDAPT, which binds activated $\alpha_4\beta_1$ and $\alpha_4\beta_7$ integrins. *J. Biol. Chem.*, **272**, 24832–24836.
- Mulloy,B., Forster,M.J., Jones,C. and Davies,D.B. (1993) NMR and molecular-modelling studies of the solution conformation of heparin. *Biochem. J.*, **293**, 849–858.
- Newham,P., Craig,S.E., Clark,K., Mould,A.P. and Humphries,M.J. (1998) Analysis of ligand-induced and ligand-attenuated epitopes on the leukocyte integrin $\alpha_4\beta_1$: VCAM-1, mucosal addressin cell adhesion molecule-1 and fibronectin induce distinct conformational changes. *J. Immunol.*, **160**, 4508–4517.
- Nicholls,A., Sharp,K.A. and Honig,B. (1991) Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins*, **11**, 281–296.
- Norton,P.A. and Hynes,R.O. (1987) Alternative splicing of chicken fibronectin in embryos and in normal and transformed cells. *Mol. Cell. Biol.*, **7**, 4297–4307.
- Obara,M., Kang,M.S. and Yamada,K.M. (1988) Site-directed mutagenesis of the cell-binding domain of human fibronectin: separable, synergistic sites mediate adhesive function. *Cell*, **53**, 649–657.
- Otwinowski,Z. and Minor,W. (1993) *DENZO: A Film Processing Program for Macromolecular Crystallography*. Yale University Press, New Haven, CT.
- Patthy,L. (1990) Homology of a domain of the growth hormone/prolactin receptor family with type III modules of fibronectin. *Cell*, **61**, 13–14.
- Pechik,I., Nachman,J., Ingham,K. and Gilliland,G.L. (1993) Crystallization and preliminary X-ray diffraction data of two heparin-binding fragments of human fibronectin. *Proteins*, **16**, 43–47.
- Petersen,T.E., Thøgersen,H.C., Skorstengaard,K., Vibe Pedersen,K., Sahl,P., Sottrup Jensen,L. and Magnusson,S. (1983) Partial primary structure of bovine plasma fibronectin: three types of internal homology. *Proc. Natl Acad. Sci. USA*, **80**, 137–141.
- Pierschbacher,M.D. and Ruoslahti,E. (1984) Variants of the cell recognition site of fibronectin that retain attachment-promoting activity. *Proc. Natl Acad. Sci. USA*, **81**, 5985–5988.
- Rocco,M., Carson,M., Hantgan,R., McDonagh,J. and Hermans,J. (1983) Dependence of the shape of the plasma fibronectin molecule on solvent composition. Ionic strength and glycerol content. *J. Biol. Chem.*, **258**, 14545–14549.
- Ruoslahti,E. (1984) Fibronectin in cell adhesion and invasion. *Cancer Metastasis Rev.*, **3**, 43–51.
- Ruoslahti,E. (1988) Fibronectin and its receptors. *Annu. Rev. Biochem.*, **57**, 375–413.
- Schwarzbauer,J.E. (1991) Fibronectin: from gene to protein. *Curr. Opin. Cell Biol.*, **3**, 786–791.
- Sticht,H., Pickford,A.R., Potts,J.R. and Campbell,I.D. (1998) Solution structure of the glycosylated second type 2 module of fibronectin. *J. Mol. Biol.*, **276**, 177–187.
- Stuart,D.I., Levine,M., Muirhead,H. and Stammers,D.K. (1979) Crystal structure of cat muscle pyruvate kinase at a resolution of 2.6 Å. *J. Mol. Biol.*, **134**, 109–142.
- Walker,A. and Gallagher,J.T. (1996) Structural domains of heparan sulphate for specific recognition of the C-terminal heparin-binding domain of human plasma fibronectin (HEPII). *Biochem. J.*, **317**, 871–877.
- Williams,M.J., Phan,I., Harvey,T.S., Rostagno,A., Gold,L.I. and Campbell,I.D. (1994) Solution structure of a pair of fibronectin type I modules with fibrin binding activity. *J. Mol. Biol.*, **235**, 1302–1311.
- Woods,A., Couchman,J.R., Johansson,S. and Hook,M. (1986) Adhesion and cytoskeletal organization of fibroblasts in response to fibronectin fragments. *EMBO J.*, **5**, 665–670.
- Woods,A., McCarthy,J.B., Furcht,L.T. and Couchman,J.R. (1993) A synthetic peptide from the COOH-terminal heparin-binding domain of fibronectin promotes focal adhesion formation. *Mol. Biol. Cell*, **4**, 605–613.

Received December 1, 1998; revised and accepted January 27, 1999