

**Structure of a Fibronectin Type III Domain from Tenascin Phased by MAD  
Analysis of the Selenomethionyl Protein**



Daniel J. Leahy; Wayne A. Hendrickson; Ikramuddin Aukhil; Harold P. Erickson

*Science*, New Series, Vol. 258, No. 5084 (Nov. 6, 1992), 987-991.

Stable URL:

<http://links.jstor.org/sici?sici=0036-8075%2819921106%293%3A258%3A5084%3C987%3ASOAFTI%3E2.0.CO%3B2-8>

*Science* is currently published by American Association for the Advancement of Science.

---

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/aaas.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

---

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

nature. Microorganisms have used antibiotic inactivation to overcome natural (16) and drug-related defenses of the host (17). The mechanism of resistance encoded by *HMI* demonstrates that the same strategy can be used by hosts to block microorganism growth. Genes that encode resistance to specific diseases may be engineered in the laboratory, as has been done for a virulence factor (18).

## REFERENCES AND NOTES

- R. P. Scheffer and A. J. Ullstrup, *Phytopathology* **55**, 1037 (1965).
- J. M. Liesch *et al.*, *Tetrahedron* **38**, 45 (1982); M. L. Gross *et al.*, *Tetrahedron Lett.* **23**, 5381 (1982); J. D. Walton, E. D. Earle, B. W. Gibson, *Biochem. Biophys. Res. Commun.* **107**, 785 (1982); M. Kawai, D. H. Rich, J. D. Walton, *ibid.* **111**, 398 (1983).
- O. C. Yoder, *Annu. Rev. Phytopathol.* **18**, 103 (1980); J. B. Rasmussen and R. P. Scheffer, *Physiol. Mol. Plant Pathol.* **32**, 283 (1988); S. J. Wolf and E. D. Earle, *Plant Science* **70**, 127 (1991).
- R. B. Meeley and J. D. Walton, *Plant Physiol.* **97**, 1080 (1991).
- R. B. Meeley, G. S. Johal, S. P. Briggs, J. D. Walton, *Plant Cell* **4**, 71 (1992).
- S. P. Briggs, *Curr. Top. Plant Biochem. Physiol.* **6**, 59 (1987).
- G. S. Johal and S. P. Briggs, *Maize Genetics Cooperation Newsletter* **64**, 37 (1990); W. D. Beavis and D. Grant, *Theor. Appl. Genet.* **82**, 636 (1991).
- Total DNA from leaf tissue of maize was isolated by the urea extraction method [S. L. Dellaporta, J. Wood, J. B. Hicks, *Plant Mol. Biol. Rep.* **1**, 18 (1983)]. Southern blots were prepared as described [P. Athma and T. Peterson, *Genetics* **128**, 163 (1991)]. For RFLP analysis, DNA was transferred to nylon membranes (MSI from Fisher) and the hybridizations were performed as above but without the addition of formamide. Probes were made from gel-purified DNA fragments and labeled by random priming (Amersham). The *Mu1*-specific probe was an internal 650-bp *Ava* I-Bst EII fragment isolated from pAVB5 [V. L. Chandler and V. Walbot, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 1767 (1986)]. The *Mu3* probe was the Hind III-Xba I fragment isolated from pKO121 (K. K. Oishi, personal communication). The *Spm* probes were a Ban II-Xba I fragment from pBx1 and a Xba I-Sal I fragment from pXS2.3 [containing probes A and B, respectively, described in K. C. Cone, R. J. Schmidt, B. Burr, F. A. Burr, in *Plant Transposable Elements*, O. Nelson, Ed. (Plenum, New York, 1988)].
- DNA isolated from homozygous mutant seedlings was digested with Sst I or Xho I and the appropriate sized DNA fragments (as judged from the Southern blots) were purified by preparative gel electrophoresis and electroelution into dialysis tubing. This purified DNA was ligated to preannealed Sst I or Xho I cut arms from the bacteriophage vector  $\lambda$  *sep6-lac5* [E. Meyerowitz and D. Hogness, *Cell* **28**, 165 (1982)]. Packaging into Gigapack Gold (Stratagene) and screening of the libraries were carried out according to the manufacturer's instructions. All clones were subcloned into Bluescript SK<sup>+</sup> (Stratagene) and maintained in the SURE strain (Stratagene) of *Escherichia coli*.
- G. S. Johal and S. P. Briggs, unpublished results.
- P. Sisco, personal communication.
- Total RNA was isolated from 5- to 6-day-old etiolated seedlings [P. Chomczynski and N. Sacchi, *Anal. Biochem.* **162**, 156 (1987)]. Poly(A)<sup>+</sup> RNA was enriched (polyATract, Promega). Samples (~15  $\mu$ g) of poly(A)<sup>+</sup> RNA were denatured with the use of formaldehyde, fractionated in a 1.3% agarose gel, and blotted onto Hybond-N. DNA probes were radiolabeled by random priming and the blots were hybridized and washed as described for Southern blots.
- The cDNA library was prepared from 14-day-old, light-grown seedlings of the inbred strain B73 as described [A. Barkan and R. A. Martienssen, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 3502 (1991)].
- R. N. Perham, N. S. Scrutton, A. Berry, *BioEssays* **13**, 515 (1991).
- To synthesize the first strand of cDNA, 20  $\mu$ g of total RNA was reverse-transcribed with the use of 0.2 pmol of primer 1 (5'-TCGTCGATGAAGTCTCTGTACCGAC-3', corresponding to nucleotides 512 to 487 of the cloned cDNA) as described by O. Ohara, R. L. Dorit, W. Gilbert, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 5673 (1989). The product was amplified by PCR with 10 pmol of primer 2 (5'-TOGGCTCTGGCTGGTCAGGAAGCTC-3', corresponding to nucleotides 98 to 123 of the cloned cDNA) and 10 pmol of primer 1. PCR conditions were as described by Perkin-Elmer Cetus, except that the reactions were performed in 20% glycerol [R. J. Henry and K. Oono, *Plant Mol. Biol. Rep.* **9**, 139 (1991)]. Reactions were heated to 94°C for 3 min, then cycled 40 times for 1 min at 94°C, 1 min at 60°C, and 2 min at 72°C; and finally extended for 10 min at 72°C. The gel-purified PCR products were reamplified and directly sequenced with the use of synthetic sequencing primers.
- V. P. W. Miao and H. D. VanEtten, *Appl. Environ. Microbiol.* **58**, 801 (1992).
- H. C. Neu, *Science* **257**, 1064 (1992).
- H. Anzai, K. Yoneyama, I. Yamaguchi, *Mol. Gen. Genet.* **219**, 492 (1989).
- Mutator* stocks used in this work were from D. Robertson, personal communication.
- D. M. Shah, R. C. Hightower, R. B. Meagher, *J. Mol. Appl. Genet.* **2**, 111 (1983).
- M. Beld, C. Martin, H. Huits, A. R. Stuitje, A. G. Gerats, *Plant Mol. Biol.* **13**, 491 (1989).
- Ac/Ds* stocks were from I. Greenblatt, plasmid pAVB5 was from L. Taylor, pBx1 and pXS2.3 were from K. Cone, pKO121 was from V. Chandler, pMac1 was from R. Meagher, the cDNA library was from A. Barkan, and  $\lambda$  *sep6-lac5* was from R. Martienssen. We thank J. Duvick and R. Bensen for critical reading of the manuscript and A. LaMotte for secretarial support.

8 June 1992; accepted 3 September 1992

## Structure of a Fibronectin Type III Domain from Tenascin Phased by MAD Analysis of the Selenomethionyl Protein

Daniel J. Leahy, Wayne A. Hendrickson, Ikramuddin Aukhil, Harold P. Erickson

Fibronectin type III domains are found in many different proteins including cell surface receptors and cell adhesion molecules. The crystal structure of one such domain from the extracellular matrix protein tenascin was determined. The structure was solved by multiwavelength anomalous diffraction (MAD) phasing of the selenomethionyl protein and has been refined to 1.8 angstrom resolution. The folding topology of this domain is identical to that of the extracellular domains of the human growth hormone receptor, the second domain of CD4, and PapD. Although distinct, this topology is similar to that of immunoglobulin constant domains. An Arg-Gly-Asp (RGD) sequence that can function for cell adhesion is found in a tight turn on an exposed loop.

Domains with amino acid (aa) sequence similarity to the type III repeats of fibronectin (FN-III domains) are found in a wide variety of proteins including adhesion molecules, cytokine receptors, muscle-related proteins, collagens, and other extracellular matrix (ECM) proteins. Although most of these ~90-aa domains have no established function, certain of them act as ligands or receptors at the cell surface. For example, the tenth FN-III repeat of fibronectin (FNfn10) interacts with integrins through its RGD sequence motif (1). The interaction between fibronectin and integrins is

believed to anchor cells to the ECM and may also provide cells with important environmental cues.

Tenascin is a large ECM protein made up of six identical subunits in a hexameric structure [reviewed in (2)]. Each subunit comprises a string of small, globular domains, including 8 to 15 FN-III domains. Tenascin has a specific pattern of expression during embryonic development and is abundant in many tumors, but it is present in restricted locations in normal adult tissues. The functions of tenascin are still unclear, but roles in tissue growth and restructuring are suggested. The third FN-III domain (TNfn3) of both human and chick tenascin has an RGD sequence in the same location as that in FNfn10. Although the cell adhesion activity of native tenascin is still controversial (2, 3), the isolated TNfn3 domain promotes strong adhesion and spreading of endothelial cells, apparently

D. J. Leahy, Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY 10032.

W. A. Hendrickson, Department of Biochemistry and Molecular Biophysics, Howard Hughes Medical Institute, Columbia University, New York, NY 10032.

I. Aukhil, Department of Periodontics, University of North Carolina Dental School, Chapel Hill, NC 27514.  
H. P. Erickson, Department of Cell Biology, Duke University Medical School, Durham, NC 27710.

**Table 1.** Values for the number of independent reflections measured, the percentage of the theoretically possible reflections measured (completeness), the percentage of measured reflections for which  $|F| > 3\sigma$ , and the  $R_{\text{sym}} [\sum_i \Sigma_j |I_i(\mathbf{h})| - \langle I(\mathbf{h}) \rangle / \Sigma_j \Sigma_i |I_i(\mathbf{h})|]$  where  $I_i(\mathbf{h})$  is the  $i^{\text{th}}$  measurement and  $\langle I(\mathbf{h}) \rangle$  is the weighted mean of all measurements of  $I(\mathbf{h})$  are shown for each wavelength. Bijvoet pairs were not distinguished for these statistics. The values shown are for the resolution range 30.0 to 1.8 Å, whereas the values in parentheses represent the resolution range 30.0 to 3.0 Å.

Wave-length (Å)	Independent reflections	Completeness (%)	$F > 3\sigma$ (%)	$R_{\text{sym}}$
0.9872	7961 (1985)	90.8 (98.7)	81.9 (97.4)	6.7 (4.8)
0.9792	7950 (1988)	90.7 (98.8)	81.1 (97.4)	6.7 (4.7)
0.9790	7952 (1991)	90.7 (99.0)	82.0 (97.6)	6.9 (4.9)
0.9717	7942 (1992)	90.6 (99.0)	78.3 (97.4)	7.2 (5.0)

mediated by the vitronectin receptor (the  $\alpha_v\beta_3$  integrin) (4). We have undertaken a MAD crystallographic study of TNfn3 to provide a structural basis for understanding RGD dependent interactions and to gain insight into this large class of related protein domains.

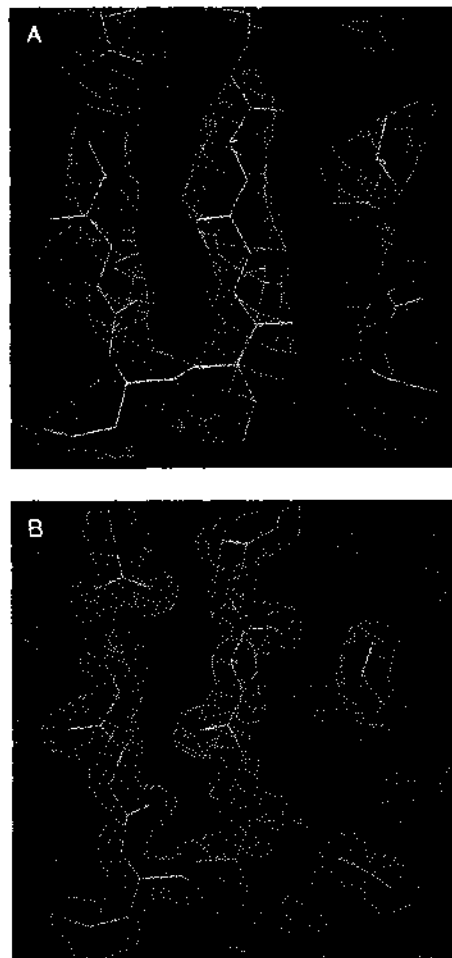
The 91-aa TNfn3 domain was expressed in *Escherichia coli* in both native and selenomethionine (Se-Met) substituted forms (5–7). Se-Met substituted TNfn3 was produced so that the anomalous diffraction signal from Se could be used to solve the crystallographic phase problem (8, 9). A fraction of purified TNfn3 containing the N-formyl-Met was used for crystallization (10).

Diffraction data for MAD analysis were collected on imaging phosphor plates at the Cornell High Energy Synchrotron Source (CHESS) (Table 1) (11). The data were processed with the program DENZO (12), reduced with the CCP4 package (13, 14), and analyzed to extract phases with the MADSYS programs (15). Both the anomalous diffraction differences (see Table 2) and the agreement of  $|F_T|$ ,  $|F_A|$ , and  $\Delta\phi$  values produced by the MAD phasing procedure for multiple measurements of the same reflection indicated that the phasing power of the anomalous signal deteriorated for reflections beyond 3.8 Å  $d$  (crystal lattice) spacings. Solution of the Se site structure by both Patterson and direct methods techniques followed by least-squares refinement revealed two partially occupied, high- $B$  value Se sites separated by 3.0 Å (16). The proximity of the Se sites indicated that both sites must arise from the same Se-Met side chain. Ultimately, both sites were found to correspond to conformers of Met<sup>860</sup>; the amino-terminal Met is disordered and does not contribute to the structure.

The presence of clear solvent channels in electron density maps calculated at 4.0 Å resolution allowed identification of the correct space-group enantiomorph as  $P4_32_12$ . The decrease in phasing power beyond 3.8 Å  $d$ -spacings led us to attempt to improve

the electron density maps by solvent flattening and density truncation as in (17) by using a hand drawn mask, but this procedure produced little or no discernible improvement in the electron density maps, presumably because of the low solvent content of the crystals.

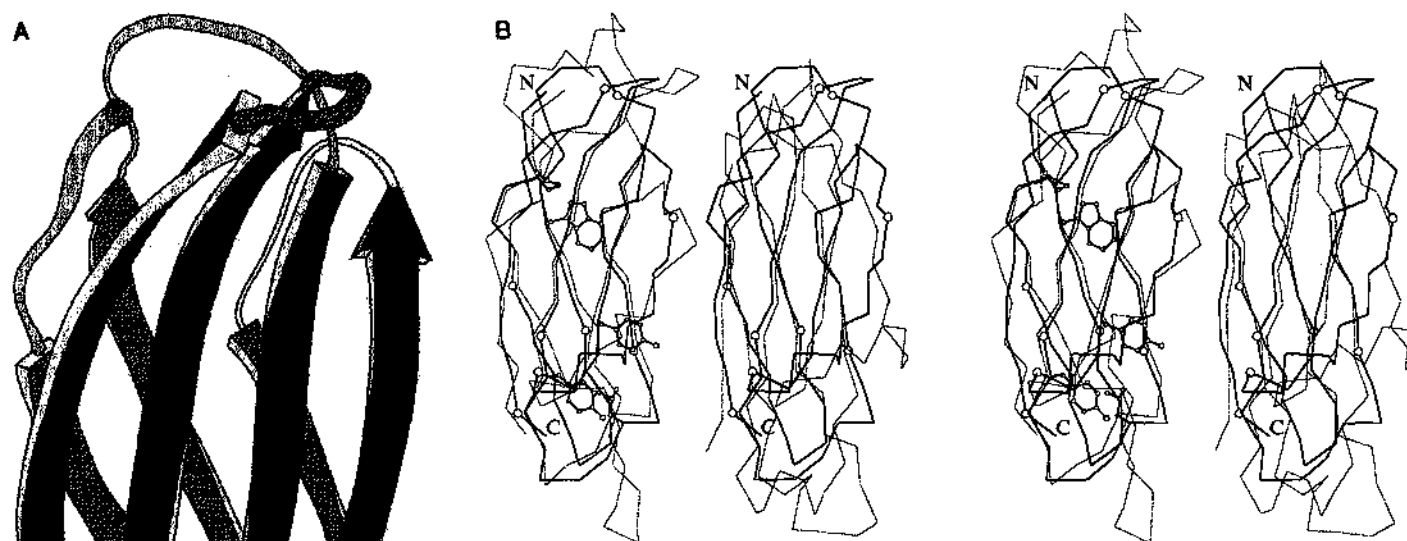
Maps made at 4.0, 3.5, 3.0, and 3.0 Å after solvent flattening were all used to produce a trace of the  $\alpha$  carbon (C $\alpha$ ) backbone. For the unflattened 3.0 Å map, all data between 3.8 and 3.0 Å for which  $|F_A| < 2\sigma$  were omitted from the map calculation (~45% of the data in this resolution range). As we were finishing the chain tracing, the secondary structure of FNfn10 as determined by nuclear magnetic resonance (NMR) was reported (18). To ensure an independent structure determination, we did not consult this result until after partially refining our structure, at which time the structural results were found to agree. The TNfn3 model has now been refined (19) to an  $R$  value  $(\sum |F_o| - |F_c|) / \sum |F_o|$ , where  $F_o$  and  $F_c$  are the observed and calculated structure factors, respectively) of 0.196 for data 10.0 to 1.8 Å with  $|F| > 3\sigma$  (0.202 for all data). The root-mean-square (rms) deviations from ideal values are 0.012 Å for bond lengths, 2.9° for angles, and 2.4° for peptide bond dihedral angles. The model consists of the carbonyl atoms of R802, residues L803 to T891, and 74 water molecules (789 atoms total) (20). Restrainted isotropic temperature factors have been refined for each atom. The rms difference between  $B$  factors of main chain atoms is 1.5 and 1.8 Å<sup>2</sup> for covalent bond (1-2) and angle (1-3) related atoms, respectively. All residues have energetically acceptable Ramachandran angles, and the side chains of five residues (D841, V871, M880, T888, and T890) have been modeled with alternative conformations. Two loop regions (L827 to E829 and R877 to M880) exhibit high  $B$  factors and less well defined density. Sufficient density is nevertheless present to provide confidence in the main chain positions, but side chain



**Fig. 1.** (A) A region of the original MAD-phased electron density encompassing strands A, B, and E is shown with the final atomic model of TNfn3. The MAD-phased map was calculated with data from 10.0 to 3.0 Å for which all reflections in the range 3.8 to 3.0 Å for which  $|F_A| > 2\sigma$  were omitted and is contoured at 1.0 $\sigma$ . (B) The final model for this region is shown with a map calculated with  $(2|F_o| - |F_c|)$  coefficients for data 10.0 to 1.8 Å and contoured at 1.0 $\sigma$ .

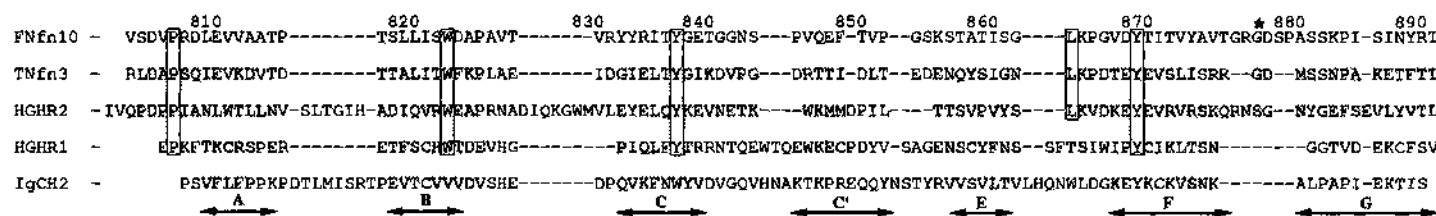
positions in these regions are ambiguous. A representative view of the electron density is shown in Fig. 1.

The structure of TNfn3 has overall dimensions of ~40 by 17 by 28 Å and consists of seven  $\beta$  strands arranged into two  $\beta$  sheets, one of four and one of three strands. A ribbon diagram of the TNfn3 structure is shown in Fig. 2A. The topology of the strands is identical to that of the D2 domain of CD4 (21), the bacterial chaperonin PapD (22), and the extracellular domains of human growth hormone receptor (hGHR1 and hGHR2) (23). We have thus labeled the strands A, B, C, C', E, F, and G by analogy with these structures, and the position of these strands in the aa sequence is shown in Fig. 3. The  $\beta$ -sheet designations from the NMR sec-



**Fig. 2.** (A) Ribbon drawing of the TNfn3 structure. (B) Stereoplot of the  $\alpha$ -carbon ( $C_{\alpha}$ ) backbones of TNfn3, hGHR2, and Ig CH2 following superposition is shown. The left image is TNfn3 (solid lines) superimposed with hGHR2 (dashed lines) (24). The residues that are highly conserved in FN-III domains (P806, W823, Y837, L863, and Y869 in TNfn3) are also shown. The right image is TNfn3 (solid lines) superimposed with IgCH2 (dashed lines) (25). In both superpositions, every tenth residue of TNfn3 is marked with an open circle, and the amino- and carboxyl-termini are labeled. TNfn3 was initially aligned with these structures manually with the interactive graphics program QUANTA (Polygen, Wallham, Massachusetts) by superimposing conserved residues and aligning the  $\beta$  strands. The  $C_{\alpha}$ s that were superimposed by this method were then used in a least-squares alignment, and the  $C_{\alpha}$ s that were within 2.5 Å of one another after this alignment were used in a subsequent least-squares alignment. The  $C_{\alpha}$ s within 2.5 Å of one another following the second alignment are shown in Fig. 3. The rms distance between  $C_{\alpha}$ s following superposition was 1.21 Å for 77 residues of TNfn3 and hGHR2, 1.24 Å for 44 residues of

TNfn3 and hGHR1, and 1.30 Å for 46 residues of TNfn3 and IgCH2. One sheet (strands C', C'', F, and G) of D2 of CD4 superimposes well with the corresponding sheet of TNfn3, but the other sheet (strands A, B, and E) does not superimpose well regardless of the orientation of the C', C'', F, and G sheet.



**Fig. 3.** Sequence alignments of FNfn10, TNfn3, hGHR1, hGHR2, and IgCH2. The latter three domains were aligned with TNfn3 after superposition of the structures as described in the legend to Fig. 1, whereas the FNfn10 sequence was aligned with TNfn3 by inspection. Residues whose  $C_{\alpha}$ s are less than 2.5 Å away from the corresponding residue in TNfn3

after superposition of the structures are shaded. Residues that are highly conserved in FN-III domains are enclosed in boxes. The FNfn10 RGD sequence is marked with an asterisk, and the positions of the  $\beta$  strands in TNfn3 are shown below the sequences. The numbering is for tenascin according to (6).

**Table 2.** Anomalous diffraction differences and scattering factors for a representative crystal of TNfn3. Observed diffraction ratios represent  $(\Delta|F|^2)^{1/2}/(|F|^2)^{1/2}$  where  $\Delta|F|^2$  is the absolute value of the Bijvoet difference at one wavelength (diagonal elements) or the dispersive difference between intersecting wavelengths for off-diagonal elements. The Bijvoet differences for centric reflections are shown in parentheses for each

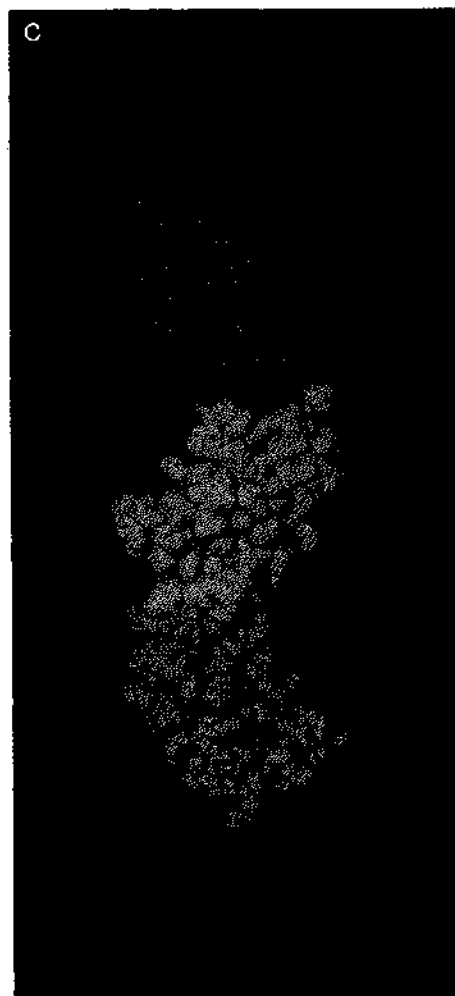
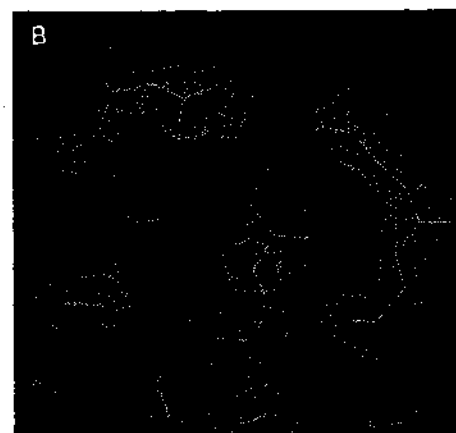
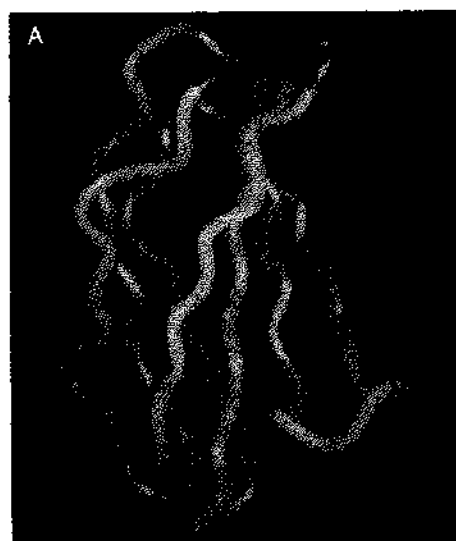
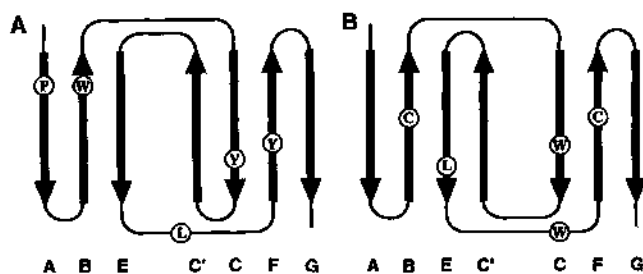
wavelength and serve as an estimate of the noise in the anomalous signals;  $f'$  for the pre-edge wavelength (0.9872 Å) was fixed at the theoretical value, 0.506 (13), whereas the other scattering factors were refined following multiple MADLSQ cycles as described in (14). Comparison of the anomalous diffraction ratios from different resolution blocks shows the loss of signal with resolution due to disorder of the Se site.

Wave-length (Å)	Observed ratio (30.0 > d > 6.0 Å)				Observed ratio (6.0 > d > 3.2 Å)				Observed ratio (3.2 > d > 2.5 Å)				Scattering factors (e)	
	0.9872	0.9792	0.9790	0.9717	0.9872	0.9792	0.9790	0.9717	0.9872	0.9792	0.9790	0.9717	f'	f''
0.9782	0.013 (0.012)	0.041	0.028	0.017	0.014 (0.012)	0.019	0.018	0.021	0.032 (0.033)	0.039	0.038	0.044	-4.54	0.506
0.9792		0.033 (0.019)	0.018	0.044		0.021 (0.017)	0.015	0.022		0.044 (0.042)	0.041	0.048	-9.17	4.29
0.9790			0.049 (0.016)	0.033			0.023 (0.014)	0.021			0.044 (0.041)	0.044	-7.93	5.39
0.9717				0.040 (0.019)				0.022 (0.014)				0.045 (0.038)	-2.94	4.27

ondary structure analysis of FNfn10 (18) are also compatible with this tertiary fold-

ing pattern. The topology of these domains is similar to that of immunoglobulin

**Fig. 4.** Schematic diagrams of (A) FN-III domains and (B) IgCH2 domains with the positions of highly conserved residues indicated. The three-dimensional relationship of the sheets can be reconstructed for each structure by closing the two sheets as if they were an open book lying face up on the page. In all cases, the conserved residues point into the hydrophobic core. For FN-III domains, the conserved Pro and Trp contact one another as do the Leu and the Tyr from the F strand. For the IgCH2 domains, the two cysteines form a disulfide bond that contacts one side of the conserved C-strand Trp whereas the Leu contacts the other side of this Trp. The C-strand Tyr of FN-III domains is in a position analogous to the C-strand Trp of IgCH2 domains, but the side chains of these two residues point in opposite directions in the aligned structures.



**Fig. 5.** (A) A representation of the  $C\alpha$  backbone of TNfn3 is shown. The amino terminus is in blue, the carboxyl terminus is in yellow, and the RGD loop is in red. This view is orthogonal to the view in Fig. 2A. (B) The electron density in the RGD-loop region is shown. The density was calculated with  $(2F_o - F_c)$  coefficients and is contoured at  $1.0\sigma$ . (C) A space-filling model for how FN-III domains might align in a repeated array. Different colored TNfn3 domains are used as the building blocks, and the RGD loop on the white domain has been colored red. The FN-III domain on the amino-terminal side of the white domain is in blue, and the domain carboxyl terminal to the white domain is in yellow. The model was produced by using the constraints described in the text; (A) was made with the program Insight II (Biosym, San Diego, California), and (C) was made with the program QUANTA (Polygen).

(Ig) constant domains (24) with the notable exception of the "sheet switching" of the C' strand. A stereoplot of an  $C\alpha$  trace of TNfn3 superimposed with both the hGHR2 and IgCH2 domains in Fig. 2B illustrates the remarkable structural similarity of these domains. The homology of hGHR2 to the FN-III motif has previously been noted by Parthy (25) and Bazan (26) on the basis of limited aa identity. In addition, Bazan suggested a more general homology of the FN-III motif to the hGHR1 and similar Ig-like domains in other receptors. The identification of a class of domains with the C' strand switched, however, has only been deduced from the recent x-ray structures and the secondary structure of FNfn10 determined by NMR.

Alignment of the structures of TNfn3 with hGHR and Ig constant domains implies the alignment of the aa sequences that is shown in Fig. 3. Although several highly conserved hydrophobic core residues of TNfn3 and hGHR match precisely, there is no obvious aa similarity between Ig-constant domains and FN-III domains. Schematic diagrams of FN-III and IgCH2 domains with the positions of highly conserved residues indicated are shown in Fig. 4. The distinctly different pattern of conserved residues and packing of the hydrophobic core observed between FN-III and Ig-constant domains raises the possibility that, despite similar structures, these two domain types have independent origins.

FN-III domains are frequently found in a repeated array, which when visualized by electron microscopy appears as an extended and relatively straight strand. The 15 FN-III repeats of tenascin form the thick segment of the hexabrachion arm, with a repeat spacing of 32 Å (7). Analysis of tenascin FN-III sequences and the TNfn3 structure indicates that residues R802 to T891 of TNfn3 comprise one FN-III repeat, and that in the hexabrachion arm repeats of this size are concatenated either directly or through a single aa insertion. The TNfn3 structure reveals a distance of 33.5 Å between the  $C\alpha$  of R802 and T891, implying a repeat distance of 36 Å or more along the amino- to carboxyl-terminal axis. In order to achieve the 32 Å spacing between domains, the amino- to carboxyl-terminal axes of successive domains must thus be tilted relative to one another. Moreover, since the amino and carboxyl termini are located on the same side of the domain, it seems likely that alternative domains are rotated by  $\sim 180^\circ$ . Such an approximate twofold screw axis relates D1 and D2 of CD4 (21) and may serve to reduce the flexibility of the FN-III arrays, producing a straight rather than

arbitrarily curved strand. A model for how these domains may align themselves that is consistent with the above constraints is shown in Fig. 5C.

The RGD sequence in TNfn3 occurs in an extended type II'  $\beta$ -hairpin loop formed between the F and G strands by residues R877 to M880. Type II' turns typically require Gly in the second position (27), suggesting that this conformation may be important for RGD-dependent cell adhesion. Electron density for the RGD-loop region of TNfn3 is shown in Fig. 5B. The single TNfn3 domain, as well as larger tenascin segments with additional domains on the carboxyl-terminal end, support adhesion and spreading of endothelial cells through the vitronectin receptor in an RGD-dependent manner (7). Tenascin segments with additional domains on the amino terminus give reduced cell adhesion and no spreading. Since the RGD loop on TNfn3 is adjacent to the amino terminus (see Fig. 5A), it seems likely that the domain TNfn2 partially blocks the accessibility of this loop.

The RGD motif of FNfn10 is in the same location as in TNfn3, but the FNfn10 loop contains four extra residues, apparently two residues each on either side of the RGD sequence (see Fig. 3). These extra residues must serve to extend the active FNfn10 RGD sequence out from the domain. If two residues were added to both the F and G strands of TNfn3, in each case continuing the  $\beta$  sheet, the RGD loop could be extended up to 7 Å. Alternatively, these extra amino acids may be inserted as loop rather than the  $\beta$  sheet, increasing the flexibility as well as extending the RGD unit. This extension may be essential to provide exposure of the RGD loop in the intact fibronectin molecule, where it interacts with the  $\alpha_5\beta_1$  integrin. The less extended RGD loop of human tenascin apparently does not interact with this integrin (7). Other active RGD sequences in known structures occur on flexible loops as seen in the snake venom "disintegrins" kistrin (28) and echistatin (29) and in the foot and mouth disease virus (30). The presentation of the RGD sequence on a flexible loop, extended from the protein surface, appears to be a common motif of the RGD-containing cell adhesion molecules.

## REFERENCES AND NOTES

- M. D. Pierschbacher and E. Ruoslahti, *Nature* **309**, 30 (1984); E. Ruoslahti and M. D. Pierschbacher, *Science* **238**, 491 (1987); R. O. Hynes, *Cell* **69**, 11 (1992).
- H. P. Erickson and M. A. Bourdon, *Annu. Rev. Cell Biol.* **5**, 71 (1989).
- D. R. Friedlander, S. Hoffman, G. M. Edelman, *J. Cell Biol.* **107**, 2329 (1988); J. Spring, K. Beck, R. Chiquet-Ehrismann, *Cell* **59**, 325 (1989); M. A. Bourdon and E. Ruoslahti, *J. Cell Biol.* **108**, 1149 (1989); V. A. Lightner and H. P. Erickson, *J. Cell Sci.* **95**, 263 (1990).
- P. Joshi, I. Aukhil, H. P. Erickson, *J. Cell Biol.* **115**, 134a (1991); in preparation.
- To express TNfn3, the DNA segment corresponding to amino acids 801 to 891 of human tenascin (6) was amplified by the polymerase chain reaction and ligated into the pET-11b expression vector [F. W. Studier, A. H. Rosenberg, J. J. Dunn, J. W. Dubendorff, *Methods Enzymol.* **185**, 60 (1990)]. In order to produce Se-Met-substituted TNfn3, the TNfn3-pET11b plasmid was used to transform the Met auxotrophic *E. coli* strain 834DE3 (the parent strain of BL21DE3) provided by F. W. Studier. The transformed bacteria were grown in 1 liter of M9 medium supplemented with 1 mg/ml thiamine and containing 40 mg/liter Se-Met in place of Met. The doubling time in this medium was 2.5 hours. Cells were harvested 10 hours after induction with IPTG (isopropyl- $\beta$ -D-thiogalactopyranoside), and the protein was purified by ion exchange chromatography (7). Se-Met substitution was complete as verified by aa composition analysis.
- D. E. Nies, T. J. Hemesath, J.-H. Kim, J. R. Gulcher, K. Stefansson, *J. Biol. Chem.* **266**, 2818 (1991).
- I. Aukhil, P. Joshi, Y. Z. Yan, H. P. Erickson, unpublished results.
- W. A. Hendrickson, J. R. Horton, D. M. LeMaster, *EMBO J.* **9**, 1665 (1990); W. Yang, W. A. Hendrickson, R. J. Crouch, Y. Satow, *Science* **249**, 1398 (1990).
- W. A. Hendrickson, *Science* **254**, 51 (1991).
- Crystals of TNfn3 were grown in hanging drops equilibrated by vapor diffusion to 18 to 22% polyethylene glycol (PEG) 3350 with 10 mM citric acid, pH 3.0, and a protein concentration of 10 to 20 mg/ml. The crystals diffract to beyond 1.8 Å  $d$ -spacings, are in space group  $P4_22_2$  with  $a = b = 49.78$  Å and  $c = 71.04$  Å, and have a calculated solvent content of 32% for one molecule in the asymmetric unit. Crystals of Se-Met TNfn3 suitable for diffraction could not be grown with Se-Met protein alone and had to be produced by seeding Se-Met protein with native TNfn3 crystals. TNfn3 has one Met (position 880) in addition to the introduced amino-terminal Met.
- Diffraction data were collected on Fuji HR-III imaging phosphor plates with oscillation photography at the F1 beam line at CHESS, modified with a channel-cut monochromator to achieve high-energy resolution (B. W. Batterman *et al.*, in preparation). Data were collected from three crystals cooled to 4°C and aligned such that Bijvoet pairs could be collected simultaneously. Complete diffraction data sets were collected for four wavelengths at or near the K absorption edge of Se. Wavelengths producing peak I' (0.9792 Å) and peak I'' (0.9790 Å) scattering factors were determined from x-ray absorption spectra, measured by fluorescence with a scintillation detector. Data from two additional "remote" wavelengths were collected to provide dispersive (due to I') differences for the MAD analysis. These remote wavelengths are at plus and minus 100 eV from the absorption edge, limited by the F1 beam-line configuration (B. W. Batterman *et al.*, in preparation). In order to minimize differences in the data sets from the effects of radiation damage, each oscillation range (typically 2.3°) was collected at all four wavelengths prior to moving to the next oscillation range.
- DENZO was written and supplied to us by Z. Otwinowski (Yale University). This program reduces diffraction data measured by oscillation photography and is adapted for image plates. It corrects for background, refines alignment parameters, performs profile fitted integrations, and applies the Lorentz and polarization factor corrections.
- ROTAVATA and AGROVATA [CCP4, *The SERC (UK) Collaborative Computing Project No. 4, A Suite of Programs for Protein Crystallography* (Daresbury Laboratory, Warrington, U.K., 1979)] were used to calculate and apply scale factors, respectively, to each wavelength. These programs were modified as described in (14).
- W. I. Weis, R. Kahn, R. Fourme, K. Drickamer, W. A. Hendrickson, *Science* **254**, 1608 (1991).
- MADSYS is a system of programs for phase determinations from multiwavelength data. Included among its programs are ANOSCL and WLSCL for local scaling of Bijvoet and dispersive data, respectively, MADLSQ to solve the phasing equations, MERGIT to merge redundant evaluations, ASLSQ to refine the structure of anomalous scattering centers, and MADFAZ to produce phases and figure of merit weights. These programs may be obtained from W.A.H.
- The two Se sites were refined with  $B$  factors of 25 and 40 Å<sup>2</sup> and occupancies of 0.58 and 0.36, respectively. The  $R$  value  $(\sum |F_o| - |F_c|) / \sum |F_o|$ , where  $F_o$  is that determined from MADLSQ for these sites was 0.412 from 30.0 to 3.0 Å and 0.351 in the resolution shell from 30.0 to 4.3 Å with a cutoff of  $2\sigma$  on  $F_o$  (57.8 and 73.5% of the data, respectively).
- W. A. Hendrickson, G. L. Klippenstein, K. B. Ward, *Proc. Natl. Acad. Sci. U.S.A.* **72**, 2160 (1975).
- M. Baron *et al.*, *Biochemistry* **31**, 2068 (1992).
- The  $C\alpha$  coordinates were read off mini-maps, and a complete model built with the program FRODO [T. A. Jones, *Methods Enzymol.* **115**, 157 (1985)]. The initial model was refined at 3.0 Å with conjugate-gradient minimization in X-PLOR [A. T. Brünger, *X-PLOR Manual, Version 3.0* (Yale University, New Haven, CT, 1992)] against the 0.9872 Å data with Bijvoet mates merged and the Se scattering factor adjusted to account for the real component of the anomalous scattering. Following rebuilding with the program FRODO, the model was subjected to a simulated annealing refinement at 2.5 Å resolution [A. T. Brünger, A. Krukowski, J. Erickson, *Acta Crystallogr.* **A46**, 585 (1990)]. Several more rounds of conventional refining and model building followed during which the resolution was gradually extended to 1.8 Å. Throughout refinement, acidic groups (Glu, Asp, and the carboxyl terminus) were modeled as protonated (pH 3.0).
- Abbreviations for the amino acid residues are: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
- S.-E. Ryu *et al.*, *Nature* **348**, 419 (1990); J. Wang *et al.*, *ibid.*, p. 411.
- A. Holmgren and C.-I. Branden, *ibid.* **342**, 248 (1989).
- A. M. deVos, M. Ultsch, A. A. Kossiakoff, *Science* **255**, 306 (1992).
- J. Deisenhofer, *Biochemistry* **20**, 2361 (1981).
- L. Pathy, *Cell* **61**, 13 (1990).
- J. F. Bazan, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 6934 (1990).
- J. S. Richardson, *Adv. Prot. Chem.* **34**, 167 (1981).
- M. Adler, R. A. Lazarus, M. S. Dennis, G. Wagner, *Science* **253**, 445 (1991).
- V. Saudek, R. A. Atkinson, J. T. Pelton, *Biochemistry* **30**, 7369 (1991).
- R. Acharya *et al.*, *Nature* **337**, 709 (1989).
- D. T. Cromer and D. Liberman, *J. Chem. Phys.* **53**, 1891 (1970).
- We thank the CHESS organization for facilitating the experiment, particularly K. Finkelstein and S. Shastri for their help with the beam line; H. E. Aronson, W. Yang, and W. I. Weis for assistance with data collection; W. I. Weis and C. Ogata for programs; F. W. Studier for bacterial strain 834DE3; R. Yarmolinsky for preparing Fig. 1A; and A. de Vos and A. Kossiakoff for the  $C\alpha$  coordinates of hGHR. Supported in part by NIH grant GM-34102 and NSF grant DMB 8917570 to W.A.H., NIH grant DE-07801 to I.A., and NIH grant CA-47056 to H.P.E. D.J.L. is a fellow of the Helen Hay Whitney Foundation. Atomic coordinates have been deposited in the Brookhaven Protein Data Bank.

12 June 1992; accepted 14 August 1992