

Crystallization of a Fragment of Human Fibronectin: Introduction of Methionine by Site-Directed Mutagenesis to Allow Phasing via Selenomethionine

Daniel J. Leahy,¹ Harold P. Erickson,² Ikramuddin Aukhil,³ Paritosh Joshi,² and Wayne A. Hendrickson¹

¹Department of Biochemistry and Molecular Biophysics, Howard Hughes Medical Institute, Columbia University, New York, New York 10032, ²Department of Cell Biology, Duke University Medical Center, Durham, North Carolina 27710, and ³Department of Periodontics, University of North Carolina, Chapel Hill, North Carolina 27514

ABSTRACT Crystals of a fragment of human fibronectin encompassing the 7th through the RGD-containing 10th type III repeats (FN7–10) have been produced with protein expressed in *E. coli*. The crystals are monoclinic with one molecule in the asymmetric unit and diffract to beyond 2.0 Å Bragg spacings. A mutant FN7–10 was produced in which three methionines, in addition to the single native methionine already present, have been introduced by site-directed mutagenesis. Diffraction-quality crystals of this mutant protein have been grown in which methionine was replaced with selenomethionine. The introduction of methionine by site-directed mutagenesis to allow phasing from selenomethionyl-substituted crystals is shown to be feasible by this example and is proposed as a general approach to solving the crystallographic phase problem. Strategies for selecting propitious sites for methionine mutations are discussed. © 1994 Wiley-Liss, Inc.

Key words: X-ray crystallography, extracellular matrix, multiwavelength anomalous diffraction (MAD)

INTRODUCTION

Fibronectin is a 500-kDa glycoprotein found in plasma and extracellular matrices throughout the body. Through interactions with cells and other extracellular matrix (ECM) molecules, fibronectin plays roles in cell adhesion, cell morphology, thrombosis, cell migration, embryonic differentiation, and phagocytosis.¹ Fibronectin is composed of homologous repeats of three types of domains.^{2–4} These domains fold independently,^{5–7} and electron micrographs of fibronectin reveal a “string-like” molecule that appears to be a concatenation of these subdomains.^{8,9} Domains homologous to the ~90 amino acid type III repeats of fibronectin (FN-III domains) have been found in 2% of animal proteins sequenced to date¹⁰ including other ECM proteins, cell-surface receptors, enzymes, and cytoplasmic muscle pro-

teins. In many proteins these domains occur as multiple consecutive repeats.

Fibronectin has been shown to interact with heparin, collagen, fibrin, and cell-surface receptors of the integrin family, and these interactions have been mapped to specific segments of fibronectin corresponding to one or a few modules. In some cases, short stretches of amino acids within a domain have been implicated in specific interactions. The 10th type III domain of fibronectin (FN10) is the major site of interaction with the $\alpha_5\beta_1$ integrin, and the sequence RGD within this domain plays a central role in the interaction. Peptides containing the RGD sequence will themselves bind to integrins and will block cell adhesion to fibronectin or to FN10. However, cell adhesion to the isolated domain FN10 is relatively weak, but is greatly enhanced by the presence of the adjacent domains FN8 and FN9.^{12–14} The cell adhesion site on fibronectin for binding $\alpha_5\beta_1$ thus includes domains FN8–10.

In a previous study we determined the atomic structure of TNfn3, an FN-III domain from tenascin that has an RGD in the same location as it occurs in FN10.⁵ The RGD was on a tight turn between β -strands F and G. In another study, Main et al. determined the atomic structure of FN10 by NMR.⁶ The RGD in this structure was on a very disordered loop. We proposed a model for a segment of several FN-III domains⁵ in which it was clear that the RGD would be at the interface between FN9 and FN10, and would perhaps be stabilized by this interface into a particular conformation. We have now obtained diffraction quality crystals of the recombinant segment FN7–10, which should allow us to determine the structural arrangement of repeating FN-III domains and the complete binding site for $\alpha_5\beta_1$, including the RGD loop as it may be constrained in this larger segment.

Received December 23, 1993; revision accepted January 6, 1994.

Address reprint requests to Daniel J. Leahy at his present address: Department of Biophysics and Biophysical Chemistry, Johns Hopkins University School of Medicine, Baltimore, MD 21205.

The original FN7-10 crystals were grown with protein that contained two amino acid substitutions relative to the native sequence. These substitutions resulted from errors in the amplification of the FN7-10 gene segment by the polymerase chain reaction (PCR), and we have thus designated this protein FN7-10/P. We have produced a new protein with these errors corrected to the native sequence (FN7-10/N) as well as a protein mutated to contain three additional methionines (FN7-10/M). These methionines replaced leucine or isoleucine at conserved hydrophobic sites and were introduced to allow solution of this structure by the technique of multiwavelength anomalous diffraction (MAD) using selenomethionyl-substituted protein.¹⁵ FN7-10 retains the two mutations found in FN7-10/P. Both FN7-10/N and FN7-10/M produced crystals in the same space group with very similar cell constants as the FN7-10/P crystals. A potential mercury derivative of FN7-10/P that is not isomorphous to FN7-10/P crystals but is isomorphous to FN7-10/M crystals has also been produced.

MATERIALS AND METHODS

Mutagenesis

Aukhil et al. have amplified the the DNA sequence encoding the 7th through the 10th FN-III domains of fibronectin by PCR¹⁶ and subcloned this sequence into the pET11b expression plasmid.¹⁷ Three contiguous DNA fragments encompassing this FN7-10 sequence were created by restriction endonuclease digestion and individually subcloned into M13mp18 or M13mp19. Site-directed mutagenesis was carried out on these fragments according to the method of Kunkel et al.¹⁸ To correct the PCR-induced mutations to the native sequence, the sequence YERHE in the F-G loop of domain 8 of FN7-10/P was changed to YEQHE, and the sequence PANSK in the G strand of domain 10 changed to PASSK. To create a version of FN7-10 that contained three additional methionines (but retained the two PCR-induced mutations) the sequence GVLTV in strand B of domain 7 was changed to GVMTV, the sequence TGIDF in strand A of domain 9 was changed to TGMDF, and the sequence TSLLI in domain 10 was changed to TSMLI (see Fig. 1). All mutated fragments were sequenced in their entirety prior to reconstructing complete FN7-10 sequence in pET11b. The amino acid sequence corresponding to the expressed FN7-10/N fragment minus an initiating methionine is shown in Figure 1 with all mutated sites indicated.

Bacterial Growth

The pET plasmids were first transformed into *E. coli* strain DH5 α and purified by a miniprep. The bacterial strain 834, which is the parent of BL21(DE3) and is a methionine auxotroph, was provided by Dr. William Studier, Brookhaven. BL21

(DE3) was transformed with the pET plasmids for FN7-10/P and FN7-10/N, and 834 was transformed with FN7-10/M, using 10–50 times the normal amount of plasmid (these strains are difficult to transform). Single colonies were selected, grown in 5 ml cultures in LB medium, and freezer stocks with 20% glycerol were made at log-phase. These freezer stocks appear to be stable for years at -80°C , and were used for all subsequent cultures.

FN7-10/P and FN7-10/N were grown in LB culture, the cultures mostly being 800 ml LB in a 2-liter flask. "Baffled flasks" (Fisher Scientific) gave 20–50% higher yields of bacteria and expression protein than did regular Ehrlenmyer flasks in several comparisons. The 800 ml cultures were infected in the morning with 15 ml of an overnight culture, from which the bacteria were centrifuged and resuspended in LB. Cultures were induced with 0.4 mM IPTG at an $A_{600} = 0.7$ –1.0, and bacteria were harvested after 3–4 hr.^{16,17}

FN7-10/M was grown in M9 medium¹⁹ supplemented with 1 mg/liter thiamine, 250 mg/liter $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 14 mg/liter $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$, 5 mg/liter FeCl_3 , 10 g/liter glucose, 50 μg /liter ampicillin, 18 amino acids (missing Met and Cys) at 50 μg /ml each, and selenomethionine at 50 μg /ml. A 15 ml culture was infected from the freezer stock 24–36 hr before starting the main culture (the bacteria do not seem to need any gradual adjustment to the selenomethionine medium). This culture usually reached saturation 20–30 hr after infection, and was left shaking for 6–12 hr until ready for use. Of this culture 5 ml was centrifuged, resuspended in fresh medium, and added to the 800 ml large culture at 8:00 pm. Growth apparently began quickly with a doubling time of 2 hr. The cultures had $A_{600} = 0.52$ at 7:25 am and 0.78 at 9:00 am, when they were induced with IPTG. Samples were taken for SDS-PAGE at 2–4 hr intervals until the bacteria were harvested at 5:00 pm (8 hr after induction). The A_{600} at this time was 1.3 for the regular Ehrlenmeyer and 1.9 for the baffled flask. The SDS-PAGE showed substantially more expression protein at 8 hr than at 4 hr, and more in the baffled flask at all times. We have not yet tested whether additional expression could be obtained by extending the culture time beyond 8 hr.

Protein Purification

All FN7-10 proteins behaved identically during purification. Bacterial lysis and initial purification were as previously described.¹⁶ Briefly, the protein was precipitated from the bacterial supernatant at 40% saturated $(\text{NH}_4)_2\text{SO}_4$, and chromatographed on mono Q, where it eluted at 0.24 M NaCl. SDS-PAGE indicated the protein was >90% pure following mono Q. For all crystallization attempts the protein was further purified by chromatography on mono S. Initially we did this at pH 3.5,¹⁶ but more recently

	<i>P</i> * *		* * <i>W</i>		* * *
FN7	PLSPPTNLHLEANPD		TGV L TVSWERSTTP		DITGYRITTTPTNGQQ
FN8	AVPPPTDLRFTNIGP		-DTMRVTWAPPPSI		DLTNFLVRYSPVKNEE
FN9	GLDSPTG I DFSDITA		-NSFTVHWIAPRA-		TITGYRIRHHPEHFSG
FN10	VSDVPRDLEVVAATP		-TS L LISWDAPAV-		TVRYRITYGETGNS
TN3	RLDAPSQIEVKDVTD		-TTALITWFKPLA-		EIDGIELTYGIKDVPG
SA	4305505157267		550201074083		8091020220238599
	↔ A		↔ B		↔ C
	* *		* <i>L P</i> <i>Y</i> * *		*
FN7	QGNSLEEVV HADQSSCTFDN		LSPGLEYNVSVYTVK--DD--		KESVPISDTIIP
FN8	DVAELSI- -PSDNAVVLTN		LLPGTEYVVS SVSVY--E Q --		HESTPLRGRQKT
FN9	RPREDRVP- -HSRNSITLTN		LTPGTEYVVSIVALN--GR--		EESPLLIGQQST
FN10	PVQEFVTP- -GSKSTATISG		LKPGVDYTITVYAVTGRGDSP		A S SKPISINYRT
TN3	DRTTIDLT- -EDENQYSIGN		LKPDTEYEVS LISRR--GD--		MSSNPAKETFTT
SA	35542707 4745638049		075915020102033 58		650953643060
	↔ C'		↔ E		↔ G
			↔ F		

Fig. 1. The complete amino acid sequence of native FN7-10 is shown with the amino acid sequences of the 4 type III repeats aligned with the sequence of a related type III repeat from tenascin (TNfn3) for which a crystal structure is available. The expressed FN7-10 consisted of these amino acids plus an initiating methionine at the N-terminus. Also shown is the fractional solvent accessibility (SA) for each of the side chains in the crystal structure of TNfn3 expressed as an integer value after multiplying by 10. Thus, 0.82 gives 8, and a side chain with the value of 5 for SA has between 45 and 55% of its available surface area exposed to solvent. The β -strands are shown labeled alphabetically as de-

scribed for the TNfn3 structure.⁵ Shown in bold type face and enclosed in boxes are the three residues mutated to methionine in FN7-10/M (strands A and B) and the sites of the two PCR-induced mutations (strands F and G). Hydrophobic amino acids that are buried (SA < 2) and conserved in most FN-III domains are indicated by asterisks; amino acids that are highly conserved in FN-III domains are indicated by letters. Consideration of these positions greatly aided selection of sites for methionine mutations as well as being generally useful for the alignment of FN-III domain sequences.

we have used pH 4.2. The peak fractions from mono Q were dialyzed into 0.02 M formic acid, adjusted to pH 4.2 and chromatographed in this buffer with a linear gradient of NaCl. FN7-10 eluted at 0.4 M NaCl at pH 4.2, and at 0.8 M NaCl at pH 3.5. In these recent experiments the yield of protein, indicated by A_{278} and the calculated $\epsilon = 1.05$,¹⁶ was 20 mg/liter of culture following mono Q, but only 7.6 mg/liter after mono S. We do not understand the reason for the apparent loss of 60% of the protein on the mono S column. The yield of protein was identical for FN7-10/N grown in LB, and for FN7-10/3M grown in M9 with selenomethionine.

Crystallization

Native and mutant forms of FN7-10 were crystallized under similar conditions. In all cases, purified protein was dialyzed into deionized H₂O and concentrated to ~7.5 mg/ml. Crystals were grown from hanging drops by the method of vapor diffusion. The reservoir buffer consisted of 5-10% PEG 3350, 20 mM Li₂SO₄, and 10 mM cacodylate pH 5.9. One microliter of protein solution was mixed 1:1 with a 1:1 dilution of reservoir buffer with deionized H₂O and equilibrated over undiluted reservoir buffer. For microseeding experiments, crystals were suspended in

reservoir buffer containing 5% PEG 3350 and crushed in a tissue homogenizer. Dilutions of the resulting suspension of microcrystals were mixed 1:1 with the protein solution and equilibrated as hanging drops over the reservoir buffer containing 10% PEG 3350.

Preparation of Mercury Acetate-Soaked FN7-10/P Crystals

A potential mercury-derivatized crystal of FN7-10/P was prepared by soaking crystals in a freshly prepared 0.2 mM solution of mercury acetate in reservoir buffer for 3 days prior to the data collection.

Data Collection and Reduction

Data from single crystals of FN7-10/P and mercury acetate-soaked FN7-10/P were collected at room temperature on a San Diego Multiwire Systems area detector using CuK α radiation and reduced using the programs ROTAVATA and AGROVATA from the CCP4 program package.²⁰ Data from single crystals of FN7-10/N and FN7-10/M were collected at 90 K on Fuji HR-III imaging plates at the X4A beamline of the National Synchrotron Light Source at Brookhaven National Laboratory. The crystals were soaked in reservoir buffer brought

to 20% ethylene glycol over 4 days in daily 5% increments prior to flash freezing immediately before data collection. Freezing loops were made after the fashion of Jeruzalmi and Steitz (manuscript in preparation). The FN7-10/N data were collected in 2.0° oscillations with no overlap between adjacent images and 0.9793 Å radiation while data from the FN7-10/M crystal were collected in 2.8° oscillations with a 0.8° overlap between adjacent oscillations with 0.9793, 0.9791, and 0.9641 Å radiation. The image plates were scanned on a Fuji BA-2000 scanner, and the data were processed using the program DENZO (written and supplied by Z. Otwinoski, Yale University) and reduced with ROTAVATA and AGROVATA.

RESULTS

Crystals of both native and mutant forms of FN7-10 appear and grow to dimensions of up to 0.15 × 0.15 × 0.8 mm within 2–3 days. Large, single, and well-formed crystals of FN7-10/N and FN7-10/P grew reproducibly, but crystals of FN7-10/M typically grew highly twinned with or without selenomethionine substitution. All crystals share a similar morphology and belong to space group $P2_1$ with 1 molecule in the asymmetric unit, but the cell constants vary up to 15% depending on the form of FN7-10 crystallized (see Table I). Microseeding slightly improved the frequency with which suitable single crystals of FN7-10/M could be obtained. No differences in crystal quality or form were seen when FN7-10/N or FN7-10/P crystals were used as seeds to grow FN7-10/M crystals. Prior to performing the mutagenesis, FN7-10/P crystals were soaked in heavy-atom solutions in attempts to produce derivatized crystals for phasing via isomorphous replacement techniques. Mercury, platinum, uranium, samarium, gold, and lead compounds were all tried at various concentrations, but in all but one case the crystals either cracked or no changes in the diffraction pattern were observed. The single exception was mercury acetate which resulted in crystals with up to 10% changes in cell constants that were thus unsuitable for isomorphous replacement (see Table I).

Crystals of all forms of FN7-10 except mercury-soaked FN7-10/P could be found that diffracted well to 2.2 Å Bragg spacings with some reflections observable beyond 1.8 Å d spacings. Mercury-soaked FN7-10/P crystals developed cracks, became highly mosaic (>0.8°), and diffracted at best to 3.5 Å. Crystals of FN7-10/M were also frequently highly mosaic and did not diffract to the 2.2 Å limit in all cases. FN7-10/N crystals all tolerated the transfer to 20% ethylene glycol well as did smaller FN7-10/M crystals, but larger crystals of FN7-10/M cracked when exposed too rapidly to high concentrations of ethylene glycol. Flash freezing resulted in an increase in mosaicity and a ~5% decrease in the

TABLE I. Unit Cell Dimensions for Crystals of Various Forms of the FN7-10 Protein*

Protein	a (Å)	b (Å)	c (Å)	β (deg.)	Mode
FN7-10/P	58.26	71.80	58.10	102.8	AD
FN7-10/P (Hg)	58.20	73.96	53.19	105.4	AD
FN7-10/M	58.39	73.80	53.74	105.0	PR
FN7-10/M (frozen)	58.64	70.50	53.26	104.4	IP
FN7-10/N	64.75	64.23	58.59	103.0	IP
FN7-10/N (frozen)	64.40	60.70	58.50	103.2	IP

*All crystals are space group $P2_1$. FN7-10/P denotes FN7-10 protein with the 2 PCR-induced mutations described in the text, FN7-10/M denotes FN7-10 protein with 3 introduced methionines as well as the PCR-induced mutations, and FN7-10/N denotes FN7-10 protein with the native sequence. The location of each of these mutations in the native amino acid sequence is shown in Figure 1. Cell constants derived from crystals frozen at 90 K (frozen) or soaked with 0.2 mM mercury acetate (Hg) are indicated. The mode by which the cell constants were determined are denoted as follows: AD, determined from data collected on a San Diego Multiwire Systems Area Detector; PR, determined by precession photography; IP, determined from data collected on image plates at the X4A beamline of the National Synchrotron Light Source at Brookhaven National Laboratory.

b-axis for both FN7-10/N and FN7-10/M crystals. Due to the small supply of FN7-10/M crystals and initial difficulties freezing them appropriately, the FN7-10/M data were collected from a crystal that had a mosaicity of ~0.8° and did not diffract well beyond 2.8 Å d spacings. Data collection statistics for frozen FN7-10/N and FN7-10/M crystals are shown in Table II.

Attempts were made to solve the FN7-10/N structure by molecular replacement using the TNfn3 and FN10 structures as search models. Although three rotation function solutions could be found that were consistently the highest peaks when the resolution, Patterson vector length, and atoms included in the search model were varied, no consistent set of translation functions could be obtained.

DISCUSSION

Our inability either to solve the FN7-10 structure by molecular replacement techniques or find suitable heavy-atom derivatives led us to consider alternative methods of solving this structure. In appropriate expression systems selenomethionine can be substituted for methionine and, if a sufficient number of methionines are present, allow phasing by the techniques of multiwavelength anomalous diffraction (MAD).¹⁵ If isomorphous crystals of selenomethionyl-substituted protein can be grown, the method of multiple isomorphous replacement (MIR) may also be applied. While produced in an expression system that allows easy substitution with selenomethionine, native FN7-10 has insufficient methionines (one) to allow MAD phasing of selenomethionyl-substituted crystals. We therefore decided to

TABLE II. Data Collection Statistics for FN7-10/N and FN7-10/M*

D_{\min} (Å)	R_{merge}	R_{cum}	N_{ref}	Mult.	%
FN7-10N					
6.20	0.072	0.072	792	1.6	77.8
4.43	0.055	0.059	1547	1.7	88.4
3.63	0.053	0.056	1963	1.6	87.5
3.15	0.057	0.057	2327	1.6	88.3
2.82	0.055	0.056	2435	1.6	81.7
2.58	0.076	0.057	2583	1.4	78.6
2.39	0.096	0.058	2624	1.4	73.8
2.23	0.180	0.060	2559	1.4	67.1
2.11	0.176	0.062	2589	1.4	63.9
2.00	0.300	0.064	2800	1.3	65.5
Total	—	0.064	22219	1.5	75.0
FN7-10/M					
8.53	0.032	0.032	315	2.0	89.1
6.15	0.048	0.040	562	2.1	93.9
5.06	0.045	0.042	734	2.2	95.9
4.40	0.048	0.044	846	2.2	93.8
3.94	0.067	0.050	961	2.3	94.2
3.60	0.089	0.058	1046	2.3	92.9
3.34	0.107	0.065	1138	2.3	93.1
3.13	0.181	0.076	1197	2.3	91.2
2.95	0.221	0.086	1235	2.2	88.5
2.80	0.274	0.094	1281	2.1	86.9
Total	—	0.094	9315	2.2	91.5

*Agreement statistics for data collected from frozen crystals of native FN7-10 (FN7-10/N) and FN7-10 with 3 additional methionines and 2 PCR-induced mutations (FN7-10/M). In both cases data were collected on image plates at the X4A beamline of the National Synchrotron Light Source at Brookhaven National Laboratory with 0.9793 Å radiation. The data were indexed and integrated with the program DENZO and reduced using the programs ROTAVATA and AGROVATA. Shown are $R_{\text{merge}} (\Sigma |I - \langle I \rangle| / \Sigma \langle I \rangle)$ for each resolution shell, the cumulative R_{merge} to each resolution shell (R_{cum}), the number of independent reflection measured for each resolution shell (N_{ref}), the average multiplicity of the measurements in each resolution shell (Mult.), and the percentage of theoretically possible reflections measured (%) for each resolution shell. Bijvoet pairs were not considered separately for this analysis.

introduce additional methionines by site-directed mutagenesis.

In several crystal structure determinations where heavy-atom derivatives were not forthcoming with standard methods, cysteine residues have been introduced to serve as sites for binding heavy atoms.^{21,22} Propitious sites for cysteine mutations are amino acid positions with side chains exposed to solvent and thus accessible to heavy-atom reagents. Sites with hydrophilic but otherwise unconserved amino acids are thus most often selected for cysteine mutations. Buried side chains, however, are preferable for methionine mutations. Buried methionine side chains are more likely to be well-ordered and thus provide better phasing power from selenomethionine, and changes in the protein that do not alter the solvent-exposed surface are less likely to dis-

TABLE III. Dayhoff Mutation Probability Matrix*

X	p(X → M)	p(M → X)	p(L → X)	p(I → X)
M	98845	98845	350	253
L	350	90	98328	122
I	253	113	212	98722
V	201	71	161	504
T	123	49	25	134
F	18	10	230	41
Other [†]	15	7	33	11

*Elements of the Dayhoff mutation probability matrix pertinent to methionine substitutions are shown. Probability values are those computed by Jones et al.²³; they are scaled by a factor of 10^5 .

[†]Average of probabilities for all remaining amino acids, namely, A, R, N, D, C, Q, E, G, H, K, P, S, W, and Y. Individual probabilities in these cases are all under 100×10^{-5} .

rupt crystal-lattice contacts and hinder isomorphous crystallization of the mutant protein. Fortunately, methionine side chains are compatible with a hydrophobic environment. Analyses of the frequency of substitution of the various amino acids in related protein sequences show that methionine is most frequently exchanged with leucine, isoleucine, and, to a lesser extent, valine.²³⁻²⁵ Pertinent elements of the Dayhoff mutation probability matrix computed by Jones et al.²³ are reproduced in Table III.

The likelihood that conservative mutations in the hydrophobic core of a protein will significantly affect the overall protein structure appears small. Structural and functional studies of proteins in which hydrophobic core residues have been systematically replaced show that methionine for leucine, isoleucine, or valine mutations have little or no effect on protein stability or function.^{26,27} Indeed, the crystal structure of a T4 lysozyme mutant with the more drastic substitution of lysine for a buried methionine showed only minor local changes in the protein structure.²⁸

From the Dayhoff matrix (Table III), methionine is the most likely replacement for leucine and the second most probable change from isoleucine (after valine), and no other amino acids are more likely to change to methionine than leucine and isoleucine. To select sites in FN7-10 to mutate to methionine we thus took advantage of sequence alignments of many related FN-III repeats to identify sites that were uniformly hydrophobic with a strong bias for leucine, isoleucine, and methionine (see Fig. 1). We also verified that the amino-acid side chains at homologous positions in the crystal structure of TNfn3 were buried in the hydrophobic core of that protein. In the sequence alignment in Figure 1 we also indicate the position of hydrophobic sites that are buried in TNfn3 (0-20% solvent accessibility) and conserved in most FN-III domains. We have found that the position of these conserved hydrophobic sites

and the few highly conserved amino acids indicated in Figure 1 aid considerably in alignment of new FN-III domains.

Our ability to purify and crystallize FN7-10/M in a fashion similar to FN7-10/P justifies our assumption that the methionine mutations were unlikely to significantly alter the protein structure. The moderate changes in cell constants of the FN7-10/M crystals relative to other FN7-10 crystals render them unsuitable for MIR phasing, although these changes provide no obstacle to MAD phasing techniques. The changes in the FN7-10/N cell constants relative to those of FN7-10/P and the fact that soaking FN7-10/P crystals with mercury acetate results in changes in the cell constants to values virtually identical to the FN7-10/M crystal cell constants (see Table I) suggest that the variant FN7-10/M cell constants are not the result of a unique alteration in the structure of FN7-10 due to the introduced methionines but probably result from an alternative packing arrangement of FN7-10 in the crystal. These cell constant changes also indicate that, if necessary, mercury-derivatized FN7-10/M crystals may provide useful phasing information from isomorphous replacement.

CONCLUSION

We have produced diffraction-quality crystals of a fragment of human fibronectin encompassing the 7th through the 10th type III domains. Solution of the atomic structure of this fragment will enable the visualization of a complete RGD-containing integrin binding site and the interdomain relationships between adjacent type III domains.

We have found that the introduction of methionine by site-directed mutagenesis is a feasible method for introducing phasing centers into protein crystals and propose the substitution of methionine for leucine or isoleucine at conserved hydrophobic sites as a general approach to solving the crystallographic phase problem for proteins.

ACKNOWLEDGMENTS

We thank Drs. Wei Yang and Craig Ogata for assistance with synchrotron data collection. This work was supported by NIH Grants R37-CA47056 (HPE), R01-DE07801 (IA), and R01-GM34102 (WAH), and by the Howard Hughes Medical Institute (WAH). D.J.L. was a fellow of the Aaron Diamond Foundation.

REFERENCES

1. Hynes, R.O. "Fibronectins." New York: Springer-Verlag, 1990.
2. Petersen, T.E., Thøgersen, H.C., Skorstengaard, K., Vibe-Pedersen, K., Sahl, P., Sottrup-Jensen, L., Magnusson, S. Partial primary structure of bovine plasma fibronectin. Three types of internal homology. *Proc. Natl. Acad. Sci. U.S.A.* 80:137-141, 1983.
3. Umezawa, K., Kornblihtt, A.R., Baralle, F.E. Isolation and characterization of cDNA clones for human liver fibronectin. *FEBS Lett.* 186:31-34, 1985.
4. Schwarzbauer, J.E., Patel, R.S., Fonda, D., Hynes, R.O. Multiple sites of alternative splicing of the rat fibronectin gene transcript. *EMBO J.* 6:2573-2580, 1987.
5. Leahy, D.J., Hendrickson, W.A., Aukhil, I., Erickson, H.P. Structure of a fibronectin type III domain from tenascin phased by MAD analysis of the selenomethionyl protein. *Science* 258:987-991, 1992.
6. Main, A.L., Harvey, T.S., Baron, M., Boyd, J., Campbell, I.D. The three-dimensional structure of the tenth type III module of fibronectin: An insight into RGD-mediated interactions. *Cell* 71:671-678, 1992.
7. Baron, M., Norman, D., Willis, A., Campbell, I.D. Structure of the fibronectin type 1 module. *Nature (London)* 345:642-646, 1990.
8. Engel, J., Odermatt, E., Engel, A., Madri, J.A., Furthmayr, H., Rohde, H., Timpl, R. Shapes, domain organizations and flexibility of laminin and fibronectin, two multifunctional proteins of the extracellular matrix. *J. Mol. Biol.* 150:97-120, 1981.
9. Erickson, H.P., Carrell, N., McDonagh, J. Fibronectin molecule visualized in electron microscopy: A long, thin flexible strand. *J. Cell Biol.* 91:673-678, 1981.
10. Bork, P., Doolittle, R. Proposed acquisition of an animal protein domain by bacteria. *Proc. Natl. Acad. Sci. U.S.A.* 89:8990-8994, 1992.
11. Pierschbacher, M.D., Ruoslahti, E. Cell attachment activity of fibronectin can be duplicated by small synthetic fragments of the molecule. *Nature (London)* 309:30-33, 1984.
12. Aota, S., Nagai, T., Yamada, K.M. Characterization of regions of fibronectin besides the arginine-glycine-aspartic acid sequence required for adhesive function of the cell-binding domain using site-directed mutagenesis. *J. Biol. Chem.* 266:15938-15943, 1991.
13. Bowditch, R.D., Halloran, C.E., Aota, S., Obara, M., Plow, E.F., Yamada, K.M., Ginsberg, M.H. Integrin $\alpha_{IIb}\beta_{3}$ (platelet GPIIb-IIIa) recognizes multiple sites in fibronectin. *J. Biol. Chem.* 266:23323-23328, 1991.
14. Kimizuka, F., Ohdate, Y., Kawase, Y., Shimojo, T., Taguchi, Y., Hashino, K., Goto, S., Hashi, H., Kato, I., Sekiguchi, K., Titani, K. Role of type III homology repeats in cell adhesive function within the cell-binding domain of fibronectin. *J. Biol. Chem.* 266:3045-3051, 1991.
15. Hendrickson, W.A., Horton, J.R., LeMaster, D.M. Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): A vehicle for direct determination of three-dimensional structure. *EMBO J.* 9:1665-1672, 1990.
16. Aukhil, I., Joshi, P., Yan, Y., Erickson, H.P. Cell- and heparin-binding domains of the hexabrachion arm identified by tenascin expression proteins. *J. Biol. Chem.* 268:2542-2553, 1993.
17. Studier, F.W., Rosenberg, A.H., Junn, J.J., Dubendorff, J.W. Use of T7 RNA polymerase to direct expression of cloned genes. *Meth. Enzymol.* 185:60-89, 1990.
18. Kunkel, T.A., Roberts, J.D., Azkour, R.A. Rapid and efficient site-specific mutagenesis without phenotype selection. *Meth. Enzymol.* 154:367-384, 1987.
19. Sambrook, J., Fritsch, E.F., Maniatis, T. "Molecular Cloning, a Laboratory Manual." New York: Cold Spring Harbor Laboratory Press, 1989.
20. "The SERC (UK) collaborative computing project No. 4, A suite of programs for protein crystallography." Warrington, U.K.: Daresbury Laboratory, 1979.
21. Sanderson, M.R., Freemont, P.S., Rice, P.A., Goldman, A., Hatfull, G.F., Grindley, N.D.F., Steitz, T.A. The crystal structure of the catalytic domain of the site specific recombination enzyme $\gamma\delta$ resolvase at 2.7 Å resolution. *Cell* 63:1323-1329, 1990.
22. Nagai, K., Oubridge, C., Jessen, T.H., Li, J., Evans, P.R. Crystal structure of the RNA-binding domain of the U1 small nuclear ribonucleoprotein A. *Nature (London)* 348:515-520, 1990.
23. Jones, D.T., Taylor, W.R., Thornton, J.M. The rapid generation of mutation data matrices from protein sequences. *Comp. Appl. Biosci. (CABIOS)* 8:275-282, 1992.
24. Dayhoff, M.O., Schwartz, R.M., Orcutt, B.C. "Atlas of Protein Sequence and Structure." Washington, DC: National

- Biomedical Research Foundation, Vol. 5, Suppl. 3, 345–352.
25. Gonnet, G.H., Cohen, M.A., Benner, S.A. Exhaustive matching of the entire protein sequence database. *Science* 256:1443–1445, 1992.
 26. Bowie, J.U., Reidhaar-Olson, J.F., Lim, W.A., Sauer, R.T. Deciphering the message in protein sequences: Tolerance to amino acid substitutions. *Science* 247:1306–1310, 1990.
 27. Lim, W.A., Farruggio, D.C., Sauer, R.T. Structural and energetic consequences of disruptive mutations in a protein core. *Biochemistry* 31:4324–4333, 1992.
 28. Dao-pin, S., Anderson, S.E., Baase, W.A., Dahlquist, F.W., Matthews, B.W. Structural and thermodynamic consequences of burying a charged residue within the hydrophobic core of T4 lysozyme. *Biochemistry* 30:11521–11529, 1991.