¹H-n.m.r. studies of the fibronectin 13 kDa collagen-binding fragment

Evidence for autonomous conserved type I and type II domain folds

Keith L. CONSTANTINE,* Shelesa A. BREW,† Kenneth C. INGHAM† and Miguel LLINÁS*‡

*Department of Chemistry, Carnegie Mellon University, Pittsburgh, PA 15213, U.S.A.,

and †Biochemistry Laboratory, American Red Cross Research and Development, Rockville, MD 20855, U.S.A.

A ¹H-n.m.r. study of a 117-residue (13 kDa) gelatin-binding fragment of human fibronectin, which contains the sixth (from the *N*-terminus) type I domain and the first type II domain, was undertaken. The resolution of the ¹H-n.m.r. spectrum indicates that the domains are independent and mobile relative to each other. Analysis of two-dimensional ¹H-n.m.r. experiments recorded at 500 MHz afforded spin-system identifications for all aromatic and a number of aliphatic residues. Utilizing the fact that phenylalanine residues occur only in the type II portion of this fragment, many spin systems were localized to either the type I or the type II module via analysis of two-dimensional nuclear-Overhauser-effect (NOESY) experiments. This allowed unambiguous assignment of the two tryptophan residues, as they occur singly in each domain. Patterns of NOESY connectivities are found to be consistent with known type I and type II domain structures; this affords a number of tentative sequence-specific assignments. For both domains, evidence of conserved hydrophobic cores and secondary-structure elements is obtained. In addition, ¹H-n.m.r.-monitored thermal-melting studies demonstrate conclusively that the domains are independently folded and that the type I domain has high thermal stability relative to the type II domain. This is consistent with the results of calorimetric studies, and also confirms the localization of spin systems determined from the NOESY data.

INTRODUCTION

Fibronectin (Fn) is a large (approx. 500 kDa), multifunctional/multidomain adhesive glycoprotein found in blood plasma and on various cell surfaces (Ruoslahti, 1988; Yamada, 1989; Hynes, 1990). It is a two-chain mosaic protein (Patthy, 1985; Holland & Blake, 1987; Baron *et al.*, 1991) in which both of the nearly identical chains are composed of three types of repeating units, namely type I, type II and type III domains, all of which appear in other proteins (Petersen *et al.*, 1989; Baron *et al.*, 1991). The various domains carry binding determinants specific for actin, DNA, fibrin, heparin, collagen, cell-surface receptors and other biological macromolecules (Yamada, 1989). Thus fibronectin plays key roles in the interactions between cells and the extracellular matrix.

Interactions with collagen and gelatin have been localized to a 42 kDa gelatin-binding fragment (GBF) of Fn (Hahn & Yamada, 1979; Forastieri & Ingham, 1985). This fragment contains (from the *N*- to the *C*-terminus) the sixth type I, first type II, second type II, seventh type I, eighth type I and ninth type I domains. The type II domains are unique to the 42 kDa GBF, whereas type I domains also occur in other regions of Fn. In order to localize the gelatin-binding site(s) further, subfragments of the 42 kDa GBF were prepared (Ingham *et al.*, 1989). Affinity chromatography on gelatin-Sepharose revealed a strongly binding 21 kDa GBF containing the eighth and ninth type I domains and a weakly binding 13 kDa GBF containing the sixth type I (Fn/I₆) and the first type II (Fn/II₁) domains (Fig. 1). In addition, the type II domains of human fibronectin, expressed in

Escherichia coli as β -galactosidase fusion proteins, have been shown to bind specifically to immobilized collagen (Bányai *et al.*, 1990). Thus it appears that separate gelatin-binding sites are located on the type I and type II domains of the Fn 42 kDa GBF.

We have previously reported (Constantine *et al.*, 1991a) a sequence-specific assignment of the ¹H-n.m.r. spectrum and preliminary models of a gelatin-binding (Bányai *et al.*, 1990) fibronectin-like type II domain, i.e. the isolated second or b-domain of the bovine seminal-fluid protein PDC-109 (PDC-109/b). Ligand-binding studies and the refined solution structure of PDC-109/b have recently been presented (Constantine *et al.*, 1991b, 1992). These studies revealed that PDC-109/b is composed of two antiparallel β -sheets and two irregular loops. A solvent-exposed aromatic surface was shown to structure a binding site for phenethylamine and leucine analogue ligands.

The solution structure of the seventh Fn type I domain (Fn/I₇) has been reported (Baron *et al.*, 1990). This study revealed that Fn/I₇ is constructed primarily of antiparallel β -sheets, with a well-defined hydrophobic core. In addition, calorimetric experiments (Litvinovich *et al.*, 1991) demonstrated that all six domains of the Fn 42 kDa GBF are independently folded, and that three of the type I domains are highly thermostable, with Fn/I₆ being most resistant to thermal denaturation. Although the domains are autonomous, evidence for additional stabilizing domain-domain interactions among certain module pairs in the Fn 42 kDa GBF was uncovered.

In the present paper we report a ¹H-n.m.r. study of the Fn 13 kDa GBF, with an emphasis placed on uncovering spectroscopic evidence for the degree of physical independence of the

Abbreviations used: GBF, gelatin-binding fragment; COSY, phase-sensitive two-dimensional chemical-shift correlated spectroscopy; Fn, fibronectin; Fn/I₆, sixth fibronectin type I domain; Fn/I₇, seventh fibronectin type I domain; Fn/II₁, first fibronectin type II domain; NOESY, phase-sensitive two-dimensional nuclear-Overhauser-effect spectroscopy; pH*, glass electrode pH reading uncorrected for deuterium isotope effects; RELAYED-COSY, absolute-value-relayed COSY; PDC-109/b, PDC-109 domain b; TOCSY, phase-sensitive two-dimensional total correlation spectroscopy.

‡ To whom correspondence should be addressed.

two domains. In the process, by reference to recently completed ^1H -n.m.r. studies of type I (Baron *et al.*, 1990) and type II (Constantine *et al.*, 1991a,b, 1992) domains, we have been able to identify and assign the aromatic spin systems and a number of the alanine, leucine, threonine and valine aliphatic spin systems. Spatial relationships among the identified spin systems were characterized via NOESY experiments. These studies reveal that the solution conformations of Fn/I₆ and Fn/II₁ are similar to the reported Fn/I₇ and PDC-109/b structures respectively. Thus type I and type II domain secondary and tertiary structures are shown to be largely conserved among these homologous domain pairs. Domain-folding independence has been demonstrated via ^1H -n.m.r. monitored thermal-melting experiments. Our studies lay the groundwork necessary for a sequential assignment and detailed structural characterization of this fragment, and for investigations of its ligand-binding properties.

MATERIALS AND METHODS

Sample preparation

Purification of Fn and the isolation of the 13 kDa GBF has been described elsewhere (Ingham *et al.*, 1989). For ^1H -n.m.r. experiments, Fn 13 kDa GBF samples were dissolved to concentrations of 2 mM in $^2\text{H}_2\text{O}$ and in H_2O . Samples in $^2\text{H}_2\text{O}$ were prepared with pH* values of 5.2, 5.7 and 6.2; pH values for samples in H_2O were 4.1, 5.6 and 6.6. Before preparation of the $^2\text{H}_2\text{O}$ samples, the exchange-labile protons were replaced with deuterons by allowing the sample to equilibrate for several hours at room temperature in $^2\text{H}_2\text{O}$ followed by extensive freeze-drying. 1,4-Dioxan was added to all samples as an internal chemical-shift reference, and was assumed to resonate 3.77 p.p.m. low-field relative to the sodium trimethylsilyl[2,2,3,3- $^2\text{H}_4$]propionate signal (DeMarco, 1977). For samples dissolved in H_2O , 10% (v/v) $^2\text{H}_2\text{O}$ was added to serve as the field/frequency lock. $^2\text{H}_2\text{O}$ (99.96%) was purchased from MSD Isotopes, St. Louis, MO, U.S.A.

N.m.r. spectroscopy

^1H -n.m.r. experiments were conducted with a Bruker AM-500 spectrometer equipped with an Aspect 3000 computer interfaced to an Aspect X-32 workstation. Two-dimensional spectra were recorded at 27 °C and 37 °C. The H_2O or residual H_2O resonance was suppressed by gated low-power irradiation during relaxation delays and mixing times (τ_m).

One-dimensional spectra were recorded with quadrature detection and resolution was enhanced via Gaussian multiplication. The two-dimensional experiments were recorded with standard pulse sequences and phase cycling schemes. Relaxation delays of 1.4–1.6 s were employed for all acquisitions. Experiments recorded in the phase-sensitive mode included COSY (Marion & Wüthrich, 1983), TOCSY (Braunschweiler & Ernst, 1983; Bax & Davis, 1985) and NOESY (Bodenhausen *et al.*, 1984) spectra. For the last-mentioned experiments, τ_m values of 100, 200 and 250 ms were employed, with τ_m randomly varied within $\pm 10\%$ to quench zero-quantum coherences. For TOCSY experiments, τ_m values between 50 and 80 ms were used. RELAYED-COSY (Wagner, 1983) spectra were recorded in magnitude mode with delays of 25 ms between the 90° and 180° mixing pulses.

Spectral widths for the two-dimensional spectra in ω_2 were 5400 Hz ($^2\text{H}_2\text{O}$) or 6200 Hz (H_2O), with the ω_1 spectral widths set to half these values. Typically, 512 or 768 equally spaced t_1 experiments were collected and Fourier-transformed to yield, after zero-filling, 2048 and 1024 real data points in ω_2 and ω_1 respectively. Before Fourier transformation, the time domain data matrices were multiplied in both directions by unshifted sine-bell weighting functions (COSY and RELAYED-COSY),

by 45°-shifted sine-bell squared filtering functions (NOESY) or by Gaussian windows (TOCSY). The frequency domain data matrices were baseline-corrected in both dimensions by the use of third-degree polynomials.

RESULTS AND DISCUSSION

Preliminary characterization

The primary sequence of the Fn 13 kDa GBF is shown in Fig. 1. The fragment consists of a 17-residue *N*-terminal tail, 35 residues composing the Fn/I₆ domain proper (defined here to include all amino acid residues from the first to the last half-cysteine), a 17-residue linker region, the Fn/II₁ domain proper (42 residues) and a six-residue *C*-terminal tail. We have adopted the following numbering scheme, which readily facilitates comparison to other type I and type II domains. The *N*-terminal tail, linker region and *C*-terminal tail residues are denoted 1(N)–17(N), 1(L)–17(L) and 1(C)–6(C) respectively. The Fn/I₆ and Fn/II₁ residues are numbered, including alignment deletions (Fig. 1), 1(I₆)–41(I₆) and 1(II₁)–42(II₁) respectively. This division is not meant to exclude the possibility of folded conformations in the tail and linker regions. Alignments (Petersen *et al.*, 1989) between the domains of the Fn 13 kDa GBF and known domain structures (Baron *et al.*, 1990; Constantine *et al.*, 1992) are shown in Fig. 2.

Fig. 1. Primary structure of the Fn 13 kDa GBF

Residues are labelled according to the standard one-letter amino acid code. Numbering of residues is done separately for the *N*-terminal peptide tail (N), Fn/I₆ (I₆), linker (L), Fn/II₁ (II₁) and *C*-terminal peptide tail (C) regions. Deletions have been added to facilitate alignment with homologous type I and type II domain sequences (Petersen *et al.*, 1989).

Fn/I ₆	¹ C V T D S - G V V ¹⁰ Y S V G M Q W L K T ²⁰ Q - - G N K Q M L C T C L G N G - - - V S C ⁴⁰
Fn/I ₇	C T T N E - G V M Y R I G D Q W D K Q H D M G - H M M R C T C V G N G R G E W T C
Fn/II ₁	¹ C V L P F T Y N G R T F Y S C T T E G R ²⁰ Q D G H L W C S T T S N Y E Q D Q K Y S F C ⁴⁰
PDC-109/b	C V F P F I Y G G K K Y E T C T K I G S M - W M S W C S L S P N Y D K D R A W K Y C

Fig. 2. Alignments between the domains of the Fn 13 kDa GBF and known domain structures

Fig. 3. ¹H-n.m.r. spectrum of the Fn 13 kDa GBF

The spectrum shown has been resolution-enhanced via Gaussian multiplication. Protein concentration was approx. 2 mM in ¹H₂O/²H₂O (9:1, v/v), pH 6.6, at 37 °C.

Fig. 4. TOCSY spectrum of the Fn 13 kDa GBF: aromatic region

Cross-peaks involving tyrosine and phenylalanine aromatic resonances are labelled. Three-bond tryptophan aromatic connectivities are traced with continuous lines. Broken lines indicate non-overlapping four-bond and five-bond connectivities. Sequence-specific assignments for most of the aromatic spin systems are given in Table 1. The spectrum was recorded with $\tau_m = 50$ ms. Protein concentration was approx. 2 mM in ²H₂O, pH* 6.2, at 37 °C.

A 500 MHz one-dimensional spectrum of the Fn 13 kDa GBF in H₂O is presented in Fig. 3. The aromatic and aliphatic spectral regions are relatively sharp and well resolved for a protein of this size. The spectral resolution points to highly mobile individual modules. Furthermore, these characteristics indicate that the 13 kDa GBF is not substantially aggregated in pure aqueous solution, and suggest that domain-domain interactions are unlikely to be significant (Oswald *et al.*, 1989; Teuten *et al.*, 1991).

In contrast with the 13 kDa GBF, PDC-109/b aggregates extensively in pure H₂O; a 10% (v/v) solution of acetone in H₂O is required to break up PDC-109/b aggregation (Constantine *et al.*, 1991a). Evidence has been obtained that indicates that the PDC-109/b Trp-23 and Met-24 residues are involved in the aggregation (Constantine *et al.*, 1992). These residues are replaced by Gly-23 and His-24 in type II₁, and this may account for the lack of aggregation for the Fn 13 kDa GBF. The occurrence of low-field-shifted C_(α)H resonances (5–6 p.p.m.), high-field-shifted methyl resonances (–0.2–0.6 p.p.m.) and widespread aromatic resonances (5.3–8.6 p.p.m.) in the Fn 13 kDa GBF spectrum indicates folded conformations for large segments of this fragment.

Many of the resonances involving exchangeable protons are quite broad (Fig. 3). This has impeded our attempts to derive a complete sequential assignment (Wüthrich, 1986) of the protein.

Spin-system identification

The Fn 13 kDa GBF contains two tryptophan, eight tyrosine, three phenylalanine and four histidine residues (Fig. 1). Fig. 4 shows the aromatic region of a 500 MHz TOCSY spectrum ($\tau_m = 50$ ms) of this fragment dissolved in ²H₂O. All ring proton spin systems (excluding histidine residues) can be located in this display. Analysis of the aromatic regions of COSY and RELAYED-COSY spectra recorded on samples in ²H₂O, and NOESY spectra recorded on both ²H₂O and H₂O samples, allowed specific identification of all ring proton resonances. In addition, most aromatic-ring proton spin systems were connected to their corresponding backbone spin systems via NOEs between the ring protons and the C_(β)H–C_(β)H' geminal pairs. The chemical shifts of the aromatic residue spin systems are reported in Table 1. The large chemical shift spread of the aromatic signals indicates aromatic clustering. The Phe-II aromatic resonances appear broad and could not be connected to the backbone spin system. A very similar situation holds for the Phe-5 ring proton signals from PDC-109/b. In PDC-109/b, this residue is buried (Constantine *et al.*, 1992). Evidence is presented below that the Fn 13 kDa Phe-II spin system is indeed Phe-5(II₁). The histidine residues do not give rise to any identifiable inter-residue NOESY cross-peaks. This precludes their assignment by reference to the previous studies (Baron *et al.*, 1990; Constantine *et al.*, 1991a,b, 1992).

The primary structure of the Fn 13 kDa GBF (Fig. 1) contains three alanine, six leucine, 14 threonine and nine valine residues. The ¹H spin systems of these residue types give rise to readily distinguishable cross-peak patterns in two-dimensional scalar

Table 1. Chemical shifts of the Fn 13 kDa GBF aromatic residues

Chemical shifts (± 0.01 p.p.m.) were measured at 37 °C, pH* (pH) 6.2 (6.6). A dash indicates that the resonance has not been located. Assignments are given in square brackets.

Residue	Chemical shift (p.p.m.)			
	NH	C _(α) H	C _(β) H	Other
His-I	—	—	3.24, 3.02	H2 8.46; H4 6.96
His-II	—	—	—, —	H2 8.42; H4 7.23
His-III	—	4.82	3.33, 3.21	H2 8.37; H4 7.27
His-IV	—	—	—, —	H2 8.20; H4 —
Phe-I [F ¹² (II ₁)]	—	4.31	3.53, 2.65	H2,6 7.25; H3,5 7.54; H4 6.90
Phe-II [F ⁵ (II ₁)]	—	—	—, —	H2,6 6.54; H3,5 7.37; H4 6.19
Phe-III [F ⁴¹ (II ₁)]	7.96	5.70	2.81, 2.66	H2,6 7.20; H3,5 7.15; H4 7.30
Trp-I [W ²⁶ (II ₁)]	9.63	5.86	3.38, 2.93	NH1 9.95; H2 7.33; H4 6.66; H5 5.28; H6 5.63; H7 6.38
Trp-II [W ¹⁶ (I ₆)]	—	—	3.19, 3.04	NH1 9.80; H2 6.79; H4 6.34; H5 6.22; H6 6.48; H7 7.09
Tyr-I [Y ⁷ (II ₁)]	—	4.19	2.56, 2.15	H2,6 6.95; H3,5 7.33
Tyr-II [Y ³³ (II ₁)]	—	—	—, —	H2,6 7.40; H3,5 6.40
Tyr-III	—	4.58	3.06, —	H2,6 7.17; H3,5 6.87
Tyr-IV	—	4.61	2.98, —	H2,6 7.12; H3,5 6.81
Tyr-V [Y ⁴ (N)]	—	4.23	3.37, 2.59	H2,6 7.01; H3,5 6.72
Tyr-VI [Y ¹³ (II ₁)]	—	—	3.10, 2.94	H2,6 6.89; H3,5 6.67
Tyr-VII [Y ³⁹ (II ₁)]	9.22	5.79	3.12, 2.87	H2,6 6.58; H3,5 6.35
Tyr-VIII [Y ¹⁰ (I ₆)]	—	—	—, —	H2,6 6.33; H3,5 5.99

Table 2. Chemical shifts of the Fn 13 kDa GBF aliphatic residues

Chemical shifts (± 0.01 p.p.m.) were measured at 37 °C, pH* (pH) 6.2 (6.6), except for Val-VII NH (pH 4.1). A dash indicates that the resonance has not been located. Assignments are given in square brackets.

Residue	Chemical shift (p.p.m.)			
	NH	C _(α) H	C _(β) H	Other
Ala-I	—	4.20	1.58	
Ala-II	—	4.51	1.39	
Ala-III	—	4.11	1.48	
Leu-I	—	—	—, —	C _(γ) H ₃ 1.68; C _(δ) H ₃ 0.89; C _(δ') H ₃ 0.74
Thr-I [T ¹⁶ (II ₁)]	8.55	4.81	3.54	C _(γ) H ₃ —0.20
Thr-II [T ⁶ (II ₁)]	8.11	4.53	3.70	C _(γ) H ₃ 0.77
Thr-III [T ¹¹ (II ₁)]	8.79	5.62	4.12	C _(γ) H ₃ 1.19
Thr-IV	—	4.94	4.08	C _(γ) H ₃ 1.18
Thr-V	7.75	4.44	4.14	C _(γ) H ₃ 1.29
Thr-VI	—	—	3.68	C _(γ) H ₃ 1.05
Thr-VII	—	—	4.58	C _(γ) H ₃ 1.17
Thr-VIII	—	—	4.07	C _(γ) H ₃ 1.17
Thr-IX	—	—	3.81	C _(γ) H ₃ 1.48
Thr-X	—	—	4.11	C _(γ) H ₃ 1.13
Thr-XI	—	—	4.40	C _(γ) H ₃ 0.93
Thr-XII	—	—	4.17	C _(γ) H ₃ 1.10
Thr-XIII	—	—	4.23	C _(γ) H ₃ 1.18
Thr-XIV	—	—	4.09	C _(γ) H ₃ 1.44
Val-I [V ⁸ (I ₆)]	7.11	3.69	1.58	C _(γ) H ₃ 0.70; C _(γ') H ₃ 0.28
Val-II [V ³ (N)]	—	4.48	1.13	C _(γ) H ₃ 0.71; C _(γ') H ₃ 0.56
Val-III	8.61	3.90	1.45	C _(γ) H ₃ 0.78; C _(γ') H ₃ 0.68
Val-IV	—	3.53	1.83	C _(γ) H ₃ 0.90; C _(γ') H ₃ 0.82
Val-V [V ³⁹ (I ₆)]	—	4.57	2.08	C _(γ) H ₃ 1.25; C _(γ') H ₃ 1.06
Val-VI	—	—	1.44	C _(γ) H ₃ 0.92; C _(γ') H ₃ 0.86
Val-VII [V ⁹ (I ₆)]	7.97	4.55	1.73	C _(γ) H ₃ 0.81; C _(γ') H ₃ 0.69

correlated spectra (Wüthrich, 1986). Figs. 5 and 6 show aliphatic regions of a COSY spectrum of the Fn 13 kDa GBF. A preliminary analysis of this spectrum and additional COSY, RELAYED-COSY and TOCSY spectra (not shown) yielded a number of spin-system identifications (Table 2). All three alanine and all 14 threonine spin systems were located (Fig. 5). Seven of the nine valine and one of the six leucine spin systems were

identified (Fig. 6). Several of the identified methyl proton signals resonate at positions significantly removed from typical random-coil values (Bundi & Wüthrich, 1979). These include (Table 2) the Thr-I C_(γ)H₃, the Val-I C_(γ')H₃ and the Val-II C_(γ')H₃, which exhibit high-field shifts indicative of ring current shielding effects, and the Val-V C_(γ)H₃, which exhibits a low-field shift indicative of ring-current deshielding.

Fig. 5. COSY spectrum of the Fn 13 kDa GBF: alanine and threonine connectivities

Cross-peaks connecting Ala $C_{\beta}H_3$ (δ_2) and Ala $C_{\alpha}H$ (δ_1) resonances, as well as those connecting Thr $C_{\gamma}H_3$ (δ_2) and Thr $C_{\beta}H$ (δ_1) resonances, are labelled. Thr-I has been assigned as Thr¹⁶(II₁). Experimental conditions are the same as for Fig. 4.

Fig. 6. COSY spectrum of the Fn 13 kDa GBF: leucine and valine connectivities

Cross-peaks connecting Leu $C_{\beta}H_3$ – $C_{\alpha}H_3$ (δ_2) and Leu $C_{\gamma}H$ (δ_1) resonances, as well as those connecting Val $C_{\gamma}H_3$ – $C_{\beta}H_3$ (δ_2) and Val $C_{\beta}H$ (δ_1) resonances, are labelled. Val-I, Val-V and Val-VII have been assigned as Val⁸(I₆), Val³⁹(I₆) and Val⁹(I₆) respectively. Experimental conditions are the same as for Fig. 4.

NOE analysis and assignments

A complete sequential assignment of the Fn 13 kDa GBF spectrum is handicapped by significant broadening of the backbone amide NH resonances. A total of only 65 NH– $C_{\alpha}H$ COSY cross-peaks have been located for all protein solution conditions thus far examined. Therefore sequence-specific resonance assignments for the Fn 13 kDa GBF are, for the present, based mainly on comparative spectroscopy with reference to the reported type I (Baron *et al.*, 1990) and type II (Constantine *et al.*, 1992) domain structures.

Fig. 7. NOESY spectrum of the Fn 13 kDa GBF: aromatic low-field $C_{\alpha}H$ region

Selected cross-peaks are labelled with arabic numerals, with the following assigned δ_2/δ_1 correlations: 1, Y-I(H3,5)/F-I(H3,5); 2, Y-I(H3,5)/Y-II(H2,6); 3, Y-VII(H2,6)/F-III(H4); 4, Y-VII(H2,6)/Y-II(H2,6); 5, Y-II(H3,5)/Y-I(H3,5); 6, Y-VII(H3,5)/F-III(H4); 7, Y-VII(H3,5)/F-III(H3,5); 8, Y-II(H3,5)/Y-I(H2,6); 9, Y-VII(H3,5)/Y-II(H3,5); 10, Y-II(H3,5)/Y-VII(H2,6); 11, W-I(H4)/Y-VII(H2,6); 12, F-I(H3,5)/W-I(H6); 13, Y-II(H2,6)/W-I(H6); 14, Y-I(H3,5)/W-I(H6); 15, F-III(H2,6)/W-I(H_α); 16, Y-I(H2,6)/W-I(H6); 17, F-I(H4)/W-I(H6); 18, W-I(H4)/Y-VII(H_α); 19, Y-II(H3,5)/Y-VII(H_α); 20, Y-II(H3,5)/W-I(H6). The NOESY was recorded with $\tau_m = 250$ ms. Protein concentration was approx. 2 mM in 2H_2O , pH* 5.2, at 37 °C.

Fig. 7 shows a portion of a NOESY spectrum ($\tau_m = 250$ ms) of the Fn 13 kDa GBF; specifically, the region displaying aromatic–aromatic and low-field $C_{\alpha}H$ –aromatic connectivities. A number of inter-residue NOESY cross-peaks have been identified. NOEs are observed between Trp-I $C_{\alpha}H$ and the Phe-III ring protons, and between Trp-I H6 and Phe-I H3.5. Since phenylalanine residues occur only within Fn/II₁ (Fig. 1), we can identify Trp-I as Trp-26(II₁) and, by elimination, Trp-II as Trp-16(I₆). Other aromatic residues interacting with one of the phenylalanine rings include Tyr-I, Tyr-VI and Tyr-VII; this demonstrates that these residues lie within Fn/II₁, as observable NOEs arising from interdomain contacts between hydrophobic residues are highly unlikely, given the noted domain independence and relative mobility. Furthermore, Tyr-II can be localized to Fn/II₁ since it gives rise to NOEs to Trp-26(II₁), Tyr-I and Tyr-VII.

As discussed above, the contacts revealed by the NOEs in Fig. 7 demonstrate the presence of a cluster of aromatic residues in Fn/II₁. This is consistent with the results obtained for PDC-109/b (Constantine *et al.*, 1991a, 1992). The aromatic cluster in PDC-109/b contains residues 7, 26, 33 and 39 (Constantine *et al.*, 1992). Modelling of PDC-109/b revealed that these residues are partially exposed to the solvent; it was subsequently demonstrated that these residues interact with leucine analogues (Constantine *et al.*, 1992). The numerous NOESY cross-peaks between Trp-26(II₁), Tyr-I, Tyr-II and Tyr-VII (Fig. 7) indicate

Fig. 8. NOESY spectra of PDC-109/b and the Fn 13 kDa GBF: aromatic-methyl region

(a) NOESY connectivities involving the PDC-109/b Thr-16 $C_{(\gamma)}H_3$ resonance (δ_1) and assigned (Constantine *et al.*, 1991a) aromatic resonances (δ_2). (b) Aromatic (δ_2)-methyl (δ_1) NOESY cross-peaks for the Fn 13 kDa GBF. Selected resonance positions are labelled, and connectivities are indicated by broken traces. The PDC-109/b NOESY was recorded with $\tau_m = 300$ ms. Protein concentration was approx. 3.5 mM in 2H_2O , pH* 4.2, at 27 °C. Conditions for the Fn 13 kDa GBF NOESY are the same as for Fig. 7.

that these residues form a similar cluster. A more detailed comparison below yields tentative sequence-specific assignments.

Aromatic-methyl NOE connectivities are shown in Fig. 8. Fig. 8(a) shows connectivities between the Thr-16 $C_{(\gamma)}H_3$ and various aromatic resonances for PDC-109/b. Fig. 8(b) contains analogous connectivities for the Fn 13 kDa GBF. The high-field shift of the Thr-I $C_{(\gamma)}H_3$ resonance (Table 2) is similar to what is observed for PDC-109/b (Constantine *et al.*, 1991a). NOEs involving the Thr-I side chain are very similar to those of Thr-16 in PDC-109/b. The comparative evidence strongly suggests that Thr-I corresponds to Thr-16(II₁). Also, Phe-II H4 gives rise to a very intense cross-peak to Thr-I $C_{(\gamma)}H_3$; this is analogous to the intense cross-peak between Phe-5 H4 and Thr-16 $C_{(\gamma)}H_3$ observed for PDC-109/b. These data, along with the broadness of the Phe-II H2,6 and H3,5 resonances, suggests that Phe-II is Phe-5(II₁). For PDC-109/b, the Thr-16 $C_{(\gamma)}H_3$ resonance also produces NOEs on the side chains of residues 7 and 12 (Fig. 8a). Thus we can tentatively assign Tyr-I to Tyr-7(II₁) and Phe-I to Phe-12(II₁). By elimination Phe-III was assigned to Phe-41(II₁). This provides the information necessary for assigning the remaining aromatic spin systems of Fn/II₁. As mentioned above, residues 7, 26, 33 and 39 of PDC-109/b form an aromatic cluster. It is also known that the rings of residues 39 and 41 are in close proximity in PDC-109/b (Constantine *et al.*, 1992). The numerous NOEs between Phe-41(II₁) and Tyr-VII establish the latter as Tyr-39(II₁). The remaining cluster residue, Tyr-II, can be assigned to Tyr-33(II₁). By elimination, the remaining unassigned Fn/II₁ aromatic residue, Tyr-VI, is assigned to Tyr-13(II₁). These assignments are consistent with the known PDC-109/b structure (Constantine *et al.*, 1992); the self-consistency of the information indicates that the overall distribution of non-polar

Fig. 9. NOESY spectrum of the Fn 13 kDa GBF: $C_{(\alpha)}H-C_{(\alpha)}H$ region

$C_{(\alpha)}H-C_{(\alpha)}H$ connectivities are indicated by *. Those involving identified spin systems are labelled Res² (δ_2)-Res¹ (δ_1). Several of the labelled cross-peaks are diminished owing to proximity to the irradiated residual $^1H^2HO$ transition; they appear more intense under other experimental conditions. Experimental conditions are the same as for Fig. 7.

side chains, and, by inference, the tertiary fold, is largely conserved between PDC-109/b and Fn/II₁.

Fig. 8(b) shows additional aromatic-methyl connectivities for the Fn 13 kDa GBF. Of particular interest are the cross-peaks between Trp-16(I₆) ring proton resonances and the Val-V $C_{(\gamma)}H_3$ and $C_{(\gamma)}H_3$ signals. In Fn/I₇, the Trp-16 ring lies against the side chain of Trp-39 (Baron *et al.*, 1990). This latter residue is replaced by Val-39(I₆) in the Fn/I₆. The NOEs mentioned above, along with the low-field shifts of the Val-V methyl groups (Table 2), strongly suggest that Val-V is Val-39(I₆). Fn/I₆ contains only one tyrosine residue. As discussed in the next section, thermal melting studies show that Tyr-VIII is Tyr-10(I₆). Assuming that conserved residues adopt similar conformations in Fn/I₆ and Fn/I₇, the NOEs observed between the Tyr-10(I₆) ring proton resonances and the Val-I $C_{(\gamma)}H_3$ signal indicates that the latter is Val-8(I₆). This assignment is confirmed by a strong sequential NOE (not shown) between the Val-I $C_{(\gamma)}H_3$ and the Val-VII NH, which identifies the latter as Val-9(I₆).

Of the remaining unassigned tyrosine ring spin systems, only that of Tyr-V gives rise to identifiable inter-residue NOEs. These interactions involve the $C_{(\gamma)}H_3$ group of Val-III. It is therefore likely that Tyr-V and Val-III correspond to Tyr-4(N) and Val-3(N) respectively.

Evidence for conserved secondary-structural features has also been obtained. Strong $C_{(\alpha)}H-C_{(\alpha)}H$ NOEs are characteristic of antiparallel β -sheet (Wüthrich, 1986). Fig. 9 shows the 3.7–5.9 p.p.m. (δ_1) to 4.2–5.9 p.p.m. (δ_2) region of a NOESY spectrum of the Fn 13 kDa GBF. Twelve $C_{(\alpha)}H-C_{(\alpha)}H$ cross-peaks have thus far been located. This is just slightly less than the number expected assuming conserved secondary structure, as seven such cross-peaks have been observed for both PDC-109/b (Constantine *et al.*, 1991a) and Fn/I₇ (Baron *et al.*, 1990). Several

Fig. 10. Thermal melting of Fn/II₁ in the Fn 13 kDa GBF

One-dimensional spectra (resolution-enhanced) of the Fn 13 kDa GBF at various temperatures showing the aromatic low-field C_αH and high-field CH₃ regions. Selected resonances are labelled. (a) 53 °C. (b) 58 °C. (c) 64 °C. (d) 69 °C. (e) Fn/II₁ refolded at 43 °C. Sample concentration was 2 mM in 100% ²H₂O, pH* 5.7.

of these interactions in the Fn 13 kDa GBF have been assigned. C_αH–C_αH cross-peaks are observed between Trp-26(II₁) and Phe-41(II₁), and between Tyr-39(II₁) and a residue with C_αH, C_βH and C_βH' signals at 5.00, 4.50 and 3.61 p.p.m. respectively. If this latter resonance is assigned as Ser-28(II₁), then these cross-peaks define an antiparallel β-sheet that is analogous to that found in PDC-109/b. A strong C_αH–C_αH cross-peak is observed between Thr-II and Thr-III. These residues have been localized to Fn/II₁ via NOEs to assigned aromatic residues. This result is consistent with these residues being Thr-6(II₁) and Thr-11(II₁), as a C_αH–C_αH cross-peaks between residues 6 and 11 is observed in PDC-109/b (Constantine *et al.*, 1991a). In addition, NOEs are observed between the Thr-III C_γH₃ group and the ring protons of Tyr-13(II₁), indicating that Thr-III is Thr-11(II₁) and Thr-II is Thr-6(II₁). A cross-peak is evident between the Val-39(I₆) C_αH and a C_αH resonance at 5.10 p.p.m. On the basis of the work of Baron *et al.*, (1990), C_αH–C_αH NOEs are expected between Val-39(I₆) and Cys-30(I₆).

Several backbone resonances have been identified and assigned based on sequential C_αH–NH NOESY connectivities observed for protein samples at 37 °C at pH 4.1. These include Ser-25(II₁) (NH resonance at 8.10 p.p.m., C_αH resonance at 4.39 p.p.m.), Cys-27(II₁) (NH resonance at 8.79 p.p.m., C_αH resonance at 4.07 p.p.m.) and Lys-38(II₁) (NH resonance at 8.16 p.p.m., C_αH resonance at 5.32 p.p.m.). Overall, these data support the presence of conserved antiparallel β-sheets between the homologous type II domains.

Thermal unfolding

Throughout the previous discussion, it has been assumed that

Fn/I₆ and Fn/II₁ are independently folded. The independence of domains can be demonstrated by ¹H-n.m.r. thermal melting studies (Teuten *et al.*, 1991). Fig. 10 shows the aromatic low-field-shifted C_αH and high-field-shifted methyl regions of one-dimensional spectra of the Fn 13 kDa GBF recorded at different temperatures. The melting profile is presented in Figs. 10(a)–10(d). It is apparent that Fn/II₁ has completely unfolded at 69 °C, whereas Fn/I₆ has not been significantly affected up to this temperature. From these data, the midpoint of the melting transition is estimated to be 61 ± 2 °C (pH* 5.7) for Fn/II₁, which is consistent with the value of 65.7 °C measured calorimetrically at pH 7.4 (Litvinovich *et al.*, 1991). The minor difference may be due to differences in pH and buffers employed.

These studies confirm the major conclusions of the thermodynamic investigations, which were that (a) the individual domain modules are independently folded and (b) Fn/I₆ has high thermal stability. Indeed, Fn/I₆ was observed to melt at 74.3 °C only in the presence of 6 M-guanidinium chloride (Litvinovich *et al.*, 1991). Reversibility of Fn/II₁ unfolding is demonstrated in Fig. 10(e), which shows the re-appearance of the folded Fn/II₁ resonances upon lowering the temperature. The thermal melting study also helps to confirm many of our assignments. For example, the Tyr-VIII ring proton resonances remain at their shifted (folded) positions at all temperatures, demonstrating that Tyr-VIII is Tyr-10(I₆). We have also observed (results not shown) that Fn/II₁ is more susceptible to biochemical degradation via enzymic proteolysis than is Fn/I₆, as Fn/II₁ was degraded after about 1 week at 37 °C in our initial protein preparation, whereas Fn/I₆ remained intact. This result led to the identification of the high thermal stability of Fn/I₆ (Litvinovich *et al.*, 1991).

CONCLUSIONS

The spin systems of all aromatic residues and many of the aliphatic alanine, leucine, threonine and valine spin systems of the Fn 13 kDa GBF have been identified via analysis of 500 MHz two-dimensional ¹H-n.m.r. spectra. Sequence-specific assignments for all the aromatic residues as well as several of the aliphatic residues within Fn/I₆ and Fn/II₁ have been obtained via comparison to known type I and type II domain conformations. This approach is based on an initial assumption that the tertiary structures among reasonably homologous domains will be largely conserved. The self-consistency of the results obtained confirms this hypothesis.

We have demonstrated that Fn/I₆ and Fn/II₁ contain hydrophobic clusters and secondary-structure elements (antiparallel β-sheets) that are very similar to those found in Fn/I₇ and PDC-109/b respectively. For PDC-109/b it was determined that the aromatic cluster composed of residues 7, 26, 33 and 39 is involved in hydrophobic ligand binding (Constantine *et al.*, 1992). Therefore Fn/II₁ may also possess a similar exposed hydrophobic surface. In addition, our results demonstrate that it should be feasible to model unknown type I and type II domain structures on the basis of known structures (Baron *et al.*, 1991).

The observations reported here afford a starting point for continuing investigations of the Fn 13 kDa GBF. A complete sequential assignment of this protein will require the establishment of conditions wherein the majority of the backbone amide NH resonances remain relatively sharp. Recombinant techniques and isotopic labelling should also be helpful.

This research was supported by U.S. Public Health Service, National Institutes of Health, Grants HL-21791, HL-29409 and HL-30200. We thank Dr. László Patthy for providing samples of PDC-109/b. K.L.C. is a recipient of a National Institutes of Health predoctoral training grant.

REFERENCES

- Bányai, L., Trexler, M., Koncz, S., Gyenes, M., Sipos, G. & Patthy, L. (1990) *Eur. J. Biochem.* **193**, 801–806
- Baron, M., Norman, D. G., Willis, A. & Campbell, I. D. (1990) *Nature (London)* **345**, 642–646
- Baron, M., Norman, D. G. & Campbell, I. D. (1991) *Trends Biochem. Sci.* **16**, 13–17
- Bax, A. & Davis, D. G. (1985) *J. Magn. Reson.* **65**, 355–360
- Bodenhausen, G., Kolger, H. & Ernst, R. R. (1984) *J. Magn. Reson.* **58**, 370–388
- Braunschweiler, L. & Ernst, R. R. (1983) *J. Magn. Reson.* **53**, 521–528
- Bundi, A. & Wüthrich, K. (1979) *Biopolymers* **18**, 285–297
- Constantine, K. L., Ramesh, V., Bányai, L., Trexler, M., Patthy, L. & Llinás, M. (1991a) *Biochemistry* **30**, 1663–1672
- Constantine, K. L., Patthy, L. & Llinás, M. (1991b) *J. Cell. Biochem. Suppl.* **15G**, 59
- Constantine, K. L., Madrid, M., Bányai, L., Trexler, M., Patthy, L. & Llinás, M. (1992), in the press
- DeMarco, A. (1977) *J. Magn. Reson.* **26**, 527–528
- Forastieri, H. & Ingham, K. C. (1985) *J. Biol. Chem.* **260**, 10546–10550
- Hahn, L. & Yamada, K. (1979) *Proc. Natl. Acad. Sci. U.S.A.* **76**, 1160–1163
- Holland, S. K. & Blake, C. C. F. (1987) *Biosystems* **20**, 181–206
- Hynes, R. O. (1990) *Fibronectins*, pp. 113–141, Springer-Verlag, New York
- Ingham, K. C., Brew, S. A. & Migliorini, M. M. (1989) *J. Biol. Chem.* **264**, 16977–16980
- Litvinovich, S. V., Strickland, D. K., Medved, L. V. & Ingham, K. C. (1991) *J. Mol. Biol.* **217**, 563–575
- Marion, D. & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun.* **113**, 967–974
- Oswald, R. E., Bogusky, M. J., Bamberger, M., Smith, R. A. G. & Dobson, C. M. (1989) *Nature (London)* **337**, 579–582
- Patthy, L. (1985) *Cell* **41**, 657–663
- Petersen, T. E., Skorstengaard, K. & Vibe-Pedersen, K. (1989) in *Fibronectin* (Mosher, D. F., ed.), pp. 1–24, Academic Press, New York
- Ruoslahti, E. (1988) *Annu. Rev. Biochem.* **57**, 375–413
- Teuten, A. J., Smith, R. A. G. & Dobson, C. M. (1991) *FEBS Lett.* **278**, 17–22
- Wagner, G. (1983) *J. Magn. Reson.* **55**, 151–156
- Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, pp. 126–136, John Wiley and Sons, New York
- Yamada, K. M. (1989) in *Fibronectin* (Mosher, D. F., ed.), pp. 47–121, Academic Press, New York

Received 16 May 1991/28 August 1991; accepted 12 September 1991