

NMR Studies of Modular Protein Structures and Their Interactions

Andrew R. Pickford* and Iain D. Campbell

Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK

Received November 12, 2003

Contents

1. Introduction	3557
2. How Modules Are Used in Biology	3557
3. The Role of NMR	3558
3.1. Definition of Binding Sites and Complexes	3558
3.2. Intermodule Dynamics and Orientation	3558
4. The Scope of This Review	3559
5. Fibrillin-1	3559
6. Fibronectin	3560
7. Proteins Involved in the Regulation of Complement Activation	3562
8. Conclusions	3564
9. Acknowledgments	3564
10. References	3564

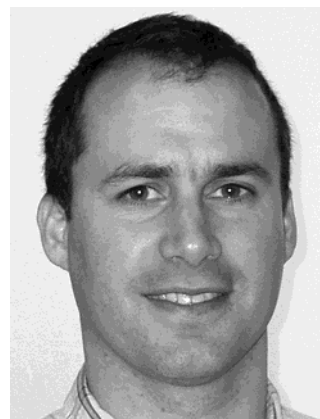
1. Introduction

Although variation at the amino acid sequence level has limitless possibilities, it has become clear that variation at the level of protein folds is quite limited.¹ One of the most striking features of this observation is the identification of repeated protein domains or modules in various genomes. A protein module can be defined as a contiguous sequence that folds to form a compact globular unit; modules appear over and over again in diverse contexts; their number is finite (several thousand), and they are usually less than about 300 amino acids in length.²

There are now excellent databases with compiled information about module distribution, sequences, and structure; see, for example, Interpro (<http://ebi.ac.uk/interpro>) and SMART (<http://smart.embl-heidelberg.de>). In general, about half of all the identified modules have a known structure.³ Determining the folds of those modules that have yet to be characterized is an attractive target for structural genomics projects.⁴ The 15 most common modules in the human genome (<http://ebi.ac.uk/proteome>) are listed in Table 1. It can be seen that all of these have known structure and that NMR has contributed to structural knowledge about all of them, in some cases many times over.

2. How Modules Are Used in Biology

Modules represent convenient and stable protein scaffolds that have been retained during biological evolution. This scaffold is the stable evolutionary



Andrew Pickford received an M.A. (1993) and D.Phil (1997) in biochemistry from the University of Oxford. He is currently a postdoctoral research associate in Iain Campbell's group investigating the molecular basis of fibronectin–collagen interactions.










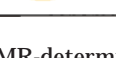


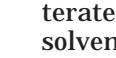

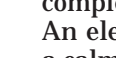


Iain Campbell received a B.Sc (1963) and Ph.D. (1967) from the University of St. Andrews. He is a Professor of Structural Biology at the University of Oxford.

entity; other features, such as the nature of the binding surface and the ways in which the modules are assembled, can evolve and change relatively rapidly to provide a range of functions. Once biology has adopted a convenient module, it can be used in a variety of ways.⁵ One of the main uses of modules seems to be to provide a binding surface to facilitate interactions among a wide range of macromolecules. New modules can be inserted into proteins readily, and they are sometimes used as spacers to present a binding function in the right position. Modular proteins seem to facilitate the construction of large dynamic complexes that assemble and disassemble. Some connections between modules are found to be relatively rigid, while, in other cases, the interface

* To whom correspondence should be addressed.

Table 1. The 15 Most Common Domains in the *Homo sapiens* Genome^a

InterPro	Matches per genome/(protein)	Name	Symbol (SMART)	NMR pdb files
IPR007110	6253 (1060)	Immunoglobulin-like		1tit, 1ci5, 1qsv, 1wit
IPR007087	29334 (962)	Zn-finger, C2H2 type		4znf, 1ard, 1bbo, 1bhi, 1fv5, 1jn7, 1jm7, 1klr, 1ncs, 1njq*
IPR001841	1194 (387)	Zn-finger, RING		1bor, 1chc, 1e4u, 1f62, 1g25, 1iym
IPR006209	3170 (368)	EGF-like domain		1adx, 1dqb, 1egf, 1epg, 1hae, 1gk5, 1hre, 1iox, 1yuf, 1bf9*
IPR001849	1150 (364)	Pleckstrin-like		1awe, 1bak, 1dro, 1fho, 1mph, 1pls, 1pms
IPR000504	1809 (346)	RNA-binding motif RNP-1		1aud, 1d8z, 1d9a, 1dz5, 1fht, 1fj7, 1fnx, 1hd0, 1u2f, 1iqt*
IPR002048	1945 (316)	Calcium-binding EF-hand		1aj4, 1ak8, 1blq, 1boc, 1c7w, 1cdc, 1cff, 1cmf, 1cmg, 1ckk*
IPR001452	2247 (293)	SH3 domain		1a0n, 1aey, 1ark, 1awj, 1awo, 1aww, 1aze, 1pnf, 1azg, 5hck*
IPR001356	1495 (259)	Homeobox		1ahd, 1bw5, 1du6, 1f43, 1ftt, 1ftz, 1hdp, 1hom, 1lfu, 1n3*
IPR003961	1932 (243)	Fibronectin, type III		1ttf, 1bj8, 1bpv, 1c8p, 1j8k, 1k85, 1lwr, 1mf, 2fnb
IPR000210	539 (215)	BTB/POZ		1qqi
IPR001478	1079 (210)	PDZ/DHR/GLGF		1b8q, 1d5g, 1gm1, 1i16, 1iu0, 1kef, 1m5z, 1n7t, 1qlc, 2pdz*
IPR005225	187 (187)	Small GTP-binding protein		1aa9, 1crp
IPR000008	877 (182)	C2 domain		1bci, 1byn, 1k5w, 3rpb
IPR002126	4821 (178)	Cadherin		1suh

^a For the cases with an *, there are more than 10 different NMR-determined structures in the PDB (<http://www.rcsb.org/pdb/>). The information about Interpro domains (<http://www.ebi.ac.uk/interpro>) was derived from <http://www.ebi.ac.uk/proteome/in> November 2003. The SMART symbols were extracted from <http://smart.embl-heidelberg.de/except> Fn3 where the color is changed for convenience.

is highly flexible. Flexible connections are of particular importance in cases where function is regulated by module rearrangement. Generally, although there are exceptions, the conformation of individual modules does not change significantly in different states, although the protein as a whole may undergo large shape changes. Some examples of these aspects of module usage will be given here.

3. The Role of NMR

As can be seen from Table 1 where the common modules in *H. sapiens* are listed, NMR has made a major contribution to the determination of the structures of individual modules in the last 15 years. In many ways, however, its main strength and likely future role will be to provide information that is complementary to that provided by X-ray crystallography. Particularly important, in this regard, is the definition of weak binding interactions and the characterization of intermodule dynamics.

3.1. Definition of Binding Sites and Complexes

With isotope labeled proteins, HSQC and TROSY experiments can rapidly and conveniently monitor perturbations caused by ligand interactions.⁶ Resonances shift or broaden selectively on complex formation. With an assigned spectrum, the induced shifts can be used to identify the location of binding sites on a known structure (chemical shift mapping). This is a powerful tool that is relatively quick and easy to use. TROSY-type experiments have been applied to protein–protein complexes as large as 900 kDa.⁷ Other important mapping methods include cross saturation from a relatively large protein to a deu-

terated protein⁸ and monitoring induced changes in solvent amide exchange rates in the presence of ligand (see, e.g., ref 9).

The ultimate goal in NMR studies of protein ligand complexes is to define the structure of the complex. An elegant early example was the determination of a calmodulin peptide complex¹⁰ that involved domain rearrangement around a peptide ligand. The observation of NOEs between protein and ligand is facilitated by differential isotope labeling, and a variety of elegant NMR methods have been developed to select between labeled and unlabeled protein.¹¹ A recent example from our own work is the determination of a complex between two protein modules from fibronectin and a peptide from a pathogenic bacterium.¹² Better use of chemical shift information and implementation of soft docking programs are likely to having an increasingly important impact in studies of complexes (e.g., ref 13).

3.2. Intermodule Dynamics and Orientation

In crystallography, relative module orientation is necessarily restricted. It has become clear, however, that combinations of modules can have considerable flexibility in solution. NMR can monitor module–module interactions in various ways. Direct contacts across modules can be measured by shift mapping just as with ligand interactions (see, e.g., the case of three contiguous modules from the collagen-binding region of fibronectin, also discussed below¹⁴). Direct intermodule contacts can also be observed by NOEs across the interface, although in many cases these intermodule NOEs are rather sparse.

In recent years, two new methods have been introduced that give longer range orientation infor-

mation. Proteins can be weakly aligned in a magnetic field, using a variety of alignment media.¹⁵ This gives rise to residual dipolar couplings (RDC) that provide long-range information, complementary to that provided by shift maps and NOEs. For example, ^{15}N - ^1H RDCs give information about the orientation of the NH bond angle with respect to parameters that define the average alignment of the protein in the alignment medium (alignment tensors). The relative orientation of a module pair can be derived from analysis of the alignment tensors of the entire molecule and the individual modules.^{16–18} Differences in the magnitudes of the principal values of the alignment tensors of individual modules in a pair can also be used to give information about intermodule flexibility.^{18,19} An interesting recent extension of this methodology was the combined use of RDCs and small angle scattering to define the assembly modes of calmodulin.²⁰

It is also possible to obtain information about intermodule orientation and flexibility by analysis of relaxation data.^{21–23} One useful procedure is to measure T_1 and T_2 relaxation of ^{15}N and/or ^{13}C atoms in isolated modules and modules in a module pair.^{24–26} By using a set of residues that is representative of the overall motions, one can obtain estimates of the diffusion parameters of the various module combinations, under identical experimental conditions. Comparison of the diffusion tensors of the isolated modules with the diffusion tensors of the individual modules in the module pair allows qualitative assessment of intermodule motion. More quantitative information can be obtained by analyzing data acquired at multiple magnetic field strengths.²⁷ An interesting recent example of the use of this methodology was the analysis of the dynamics of two KH modules in complex with DNA.²⁸

4. The Scope of This Review

Because structure determination of individual modules is relatively advanced and because there are a very large number of examples in the literature, this aspect is not discussed in any detail here. Our main intention is to discuss the biological context of modular proteins and the ways in which NMR can be used to study interactions between modular proteins and their various ligands as well as intermodule orientation and flexibility. To further reduce the scope to manageable proportions, we have restricted the main illustrative examples to literature on extracellular modular proteins. We have also mainly left discussion of NMR methodology to other contributors in this volume.

5. Fibrillin-1

Fibrillin-1 is a major structural component of the 10–12 nm beaded microfibrils in the extracellular matrix of elastic and nonelastic tissues. This 350 kDa protein has a modular architecture consisting mainly of tandem arrays of calcium-binding epidermal growth factor-like modules (cbEGF) interspersed with single transforming growth factor β -binding protein-like (TB) modules.²⁹ Genetic mutations that cause amino

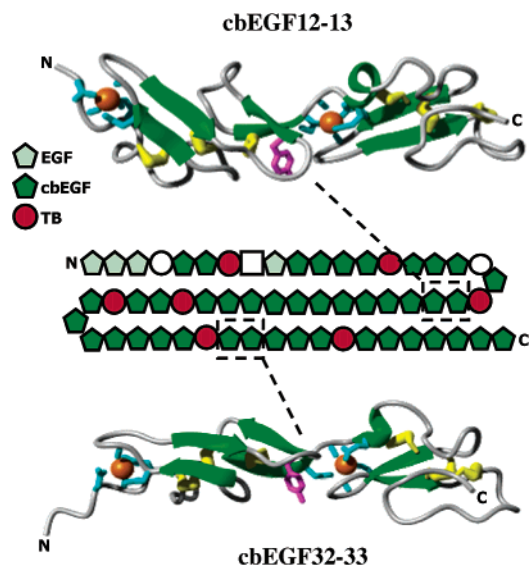


Figure 1. Solution structures of cbEGF module pairs from human fibrillin-1. Ribbon diagrams of the lowest-energy structures of cbEGF12-13³⁸ and cbEGF32-33³¹ are depicted above and below the mosaic structure of fibrillin-1. Regions of secondary structure are shown in green, with disulfide bonds in yellow, Ca^{2+} ions in orange, the calcium ligating residues at the N-terminal end of each module in light blue, and the conserved aromatic residue that forms the intermodule interface in pink.

acid changes within fibrillin-1 have been linked to Marfan syndrome (MFS), a relatively common connective tissue disorder that affects $\sim 1/5000$ individuals in the population.³⁰ To date, no clear genotype–phenotype relationships have been derived for these MFS mutations, but knowledge of the structure, function, and dynamics of fibrillin-1 should help to explain their molecular basis.

Each of the cbEGF modules in fibrillin-1 includes the calcium-binding consensus sequence D/N–X–D/N–E/Q–X_m–D/N*–X_n–Y/F, where m and n are variable and an asterisk indicates possible posttranslational addition of a β hydroxyl group.²⁹ The module also contains six cysteine residues which form three disulfide bonds in a 1–3, 2–4, and 5–6 arrangement. Multiple X-ray crystallographic and NMR studies have shown that the secondary structure of the cbEGF module comprises two regions of double-stranded β -sheet, and a loop between the first and second cysteine residues that is sometimes observed to have α -helical character.³¹ The calcium chelating residues are clustered toward the N-terminal end of the elongated EGF module (Figure 1).

The 69–78 residue TB modules contain eight cysteine residues including an unusual cysteine triplet. The NMR solution structure of TB6 from human fibrillin-1 identified a novel globular fold of six antiparallel β -strands and two α -helices with the cysteine triplet buried in the hydrophobic core.³² The four disulfide bonds are paired in a 1–3, 2–6, 4–7, and 5–8 arrangement, with the 2–6 and 4–7 bridges solvent exposed. An RGD motif in TB4 of fibrillin-1 mediates cell binding via $\alpha\text{V}\beta 3$ integrin and is predicted to be located on the solvent-exposed tip of a β -hairpin based on the coordinates of TB6.³² There have been no reports of calcium binding to the TB module.

Table 2. Calcium Dissociation Constants for Fibrillin-1 Fragments by NMR

fragment	module	K_d value	ref
cbEGF12-13	cbEGF12	1.6 mM	38
	cbEGF13	$\leq 30 \mu\text{M}$	
cbEGF13-14	cbEGF13	3 mM	67
	cbEGF14	$\leq 100 \mu\text{M}$	
cbEGF32	cbEGF32	4.3 mM	33
TB6-cbEGF32	cbEGF32	1.6 mM	68
cbEGF32-33	cbEGF32	9.2 mM	37
	cbEGF33	$350 \mu\text{M}$	

The supermolecular structure and proteolytic susceptibility of the 10–12 nm beaded microfibrils are profoundly affected by calcium binding. An explanation for this behavior requires knowledge of the calcium-binding affinity of individual cbEGF modules, and how this affinity is modulated by module–module interactions. The affinity of cbEGF-containing protein fragments for calcium has been determined by one- and two-dimensional NMR, equilibrium dialysis, intrinsic protein fluorescence, and chromophoric chelator analysis. However, of these, only NMR allows definitive assignment of K_d values to specific modules when studying constructs containing multiple binding sites. As shown in Table 2, NMR-monitored Ca^{2+} -titrations of cbEGF modules from fibrillin-1 have yielded K_d values ranging from $<30 \mu\text{M}$ to 9 mM.^{33–37} In each case, the chemical shift of $\text{H}\delta$ proton resonances of the consensus tyrosine/phenylalanine residue were observed with increasing CaCl_2 concentration, because the calcium on–off rates were fast enough to obtain “fast exchange” NMR conditions. Taken together, the results reveal that the calcium-binding activity of a cbEGF module is dramatically enhanced by the linkage of a second cbEGF module to the N-terminus (e.g., cbEGF12-13 versus cbEGF13), but addition of a TB module to the N-terminus produces only a modest improvement (e.g., TB6-cbEGF32 versus cbEGF32). In cbEGF12-13 and cbEGF32-33, the dissociation constants of the C-terminal modules ($<30 \mu\text{M}$ for cbEGF13 and $350 \mu\text{M}$ for cbEGF33) ensure near-complete occupancy of the sites at physiological calcium levels.

The structural basis for this calcium-binding enhancement in cbEGF module pairs of fibrillin-1 has been investigated through the structure determination of the cbEGF32-33 and cbEGF12-13 module pairs by NMR.^{31,38} The two structures were solved using identical methodology, with ensembles calculated from interproton NOEs and backbone dihedral restraints. In both the cbEGF32-33 and the cbEGF12-13 module pairs, the two modules are organized in a well-defined, rodlike conformation, which is stabilized by hydrophobic packing interactions (Figure 1). The similar arrangement of the two cbEGF module pairs is evident from the intermodule “tilt” and “twist” angles: $30^\circ \pm 15^\circ$ and $152^\circ \pm 13^\circ$ for cbEGF12-13, and $18^\circ \pm 6^\circ$ and $159^\circ \pm 6^\circ$ for cbEGF32-33. The intermodule packing interactions involve a conserved aromatic residue at the open end of the minor β -sheet of cbEGF12 or cbEGF32, packing against the top of the major β -sheet of cbEGF13 or cbEGF33, respectively. In each case, the calcium ion is localized to the module–module interface, where it may play a role in stabilizing the superstructure of the pair.

Recent investigations of the backbone dynamic properties of the cbEGF32-33 and cbEGF12-13 pairs have shown that the vicinity of the cbEGF33 and cbEGF13 calcium-binding sites, localized to the intermodule interfaces, are the most stable regions of the protein fragments.^{38,39} For each protein, longitudinal (T_1) and transverse (T_2) relaxation time constants and the steady-state ^1H – ^{15}N NOEs of all nonoverlapped amide resonances were measured under conditions of saturating calcium and were used as input for model-free analysis of the internal dynamics (Lipari and Szabo, 1989a,b). A reduced NOE value was not observed for the single linker residue in either cbEGF12-13 or cbEGF32-33, indicating that fibrillin-1 cbEGF module pairs possess a rigid intermodule linker when saturated with Ca^{2+} .

For cbEGF32-33, and to a lesser extent for cbEGF12-13, numerous residues in a “lobe” of the N-terminal module (cbEGF32 or cbEGF12), containing the low affinity calcium-binding site and encompassed by the 1–3 and 2–4 disulfide bonds, were found to possess significantly smaller T_2 values (and higher R_{ex} terms) than the rest of the protein, indicative of motion on the microsecond to millisecond time-scale.^{38,39} These effects were not seen for the equivalent lobe of the C-terminal module (cbEGF33 or cbEGF13) which contains the high affinity calcium-binding site. Therefore, the intermodule interface in cbEGF module pairs appears to stabilize the calcium-binding site of the C-terminal module.

Backbone amide groups with high ^1H – ^{15}N NOE and minimal exchange contributions (R_{ex}) were selected for determination of the magnitude and orientation of the diffusion tensors. In each case, the relaxation data were best fit to a prolate, symmetric top model, as expected from the atomic coordinates (Figure 1). For cbEGF32-33, the diffusion tensor was found to align well with the inertia tensor, but had a lower than expected axial ratio (D_{\parallel}/D_{\perp}) of 1.55:1 which was primarily attributed to the significant flexibility observed in the N-terminal lobe of the cbEGF32 module.^{31,39} The cbEGF12-13 module pair, which has significantly higher Ca^{2+} -binding activity than cbEGF32-33, exhibited increased anisotropy with an axial ratio of 1.9:1, indicating that the N-terminal lobe of cbEGF12 in cbEGF12-13 has a more rigid structure.³⁸ This correlation of reduced dynamics with increased calcium binding was supported by a comparison of the holo and apo forms of the cbEGF32-33 pair, which showed increased motion on the microsecond to millisecond time-scale in the absence of calcium, particularly for residues of cbEGF33.³⁹

6. Fibronectin

The extracellular matrix glycoprotein fibronectin is a large, multifunctional molecule involved in adhesion and migration events in a range of important physiological processes such as embryogenesis, wound healing, haemostasis, and thrombosis.⁴⁰ As a soluble dimer in plasma, it is involved in blood coagulation through its affinity for fibrin and platelets. As an insoluble network in the extracellular matrix, it interacts with cell surface receptors and

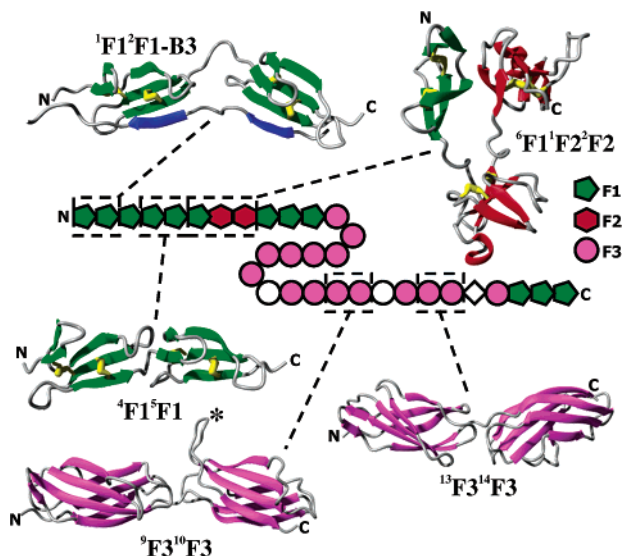


Figure 2. Structures of human fibronectin fragments that have been investigated by NMR. Ribbon diagrams of the lowest-energy solution structures of $^1\text{F1}^2\text{F1-B3}$,¹² $^4\text{F1}^5\text{F1}$,^{41–43} and $^6\text{F1}^1\text{F2}^2\text{F2}$ ¹⁴ are shown together with fragments of the crystal structures of $^7\text{F3}^8\text{F3}^9\text{F3}^{10}\text{F3}^{49}$ and $^{12}\text{F3}^{13}\text{F3}^{14}\text{F3}$.⁶⁶ Each has been mapped onto the mosaic structure of a fibronectin monomer. The secondary structure elements of the F1, F2, and F3 modules and of the B3 peptide are shown in green, red, pink, and blue, respectively. The disulfide bonds of the F1 and F2 modules are colored yellow. The RGD loop of the $^9\text{F3}^{10}\text{F3}$ module pair is marked with an asterisk (*).

with other matrix components such as collagens and protoglycans, thus assisting cell migration and the maintenance of tissue integrity.⁴⁰ Each monomer of fibronectin is composed almost entirely of three types of module (F1, F2, and F3), which are organized into functional domains (Figure 2).

The amino-terminal domain of fibronectin contains five sequential F1 modules, each of which is characterized by a β -sandwich of two antiparallel β -sheets, a double-stranded sheet followed by a triple-stranded sheet. The fibrin-binding activity of this domain requires an intact $^4\text{F1}^5\text{F1}$ module pair; deletion of either module inhibits binding.⁴⁰ The solution structure of $^4\text{F1}^5\text{F1}$ revealed an extended, linear arrangement of the two modules with intimate hydrophobic contacts between them (Figure 2).⁴¹ NMR relaxation data (T_1 , T_2 , and NOE) showed that the module pair behaves as a rigid rod.^{42,43} In contrast, no intermodule NOEs were observed for the $^1\text{F1}^2\text{F1}$ module pair with the result that the modules are effectively uncoupled, joined only by a highly dynamic polypeptide linker.⁴⁴

Staphylococcus aureus and *Streptococcus pyogenes*, two important human pathogens, target host fibronectin in their adhesion to and invasion of host cells. Fibronectin-binding proteins (FnBPs), anchored in the bacterial cell wall, have multiple fibronectin binding repeats in an unfolded region of the protein. Recently, the solution structure of a streptococcal (*S. dysgalactiae*) FnBP peptide (B3) in complex with the $^1\text{F1}^2\text{F1}$ module pair was solved by NMR spectroscopy.¹² The structure determination involved the use of NOEs (including 127 intermolecular distances), backbone dihedral angle restraints from ^{13}C chemical

shifts, and N–H bond vector restraints derived from ^{15}N relaxation rates and ^1H – ^{15}N RDCs. In the family of structures, the individual F1 modules are well defined and retain the structure of the free F1 module.⁴⁴ On binding to the module pair, the bacterial peptide contributes a fourth antiparallel strand to the triple-stranded β -sheet of sequential F1 modules in a “tandem β -zipper” interaction (Figure 2). Despite the use of long-range structural restraints from RDCs and ^{15}N relaxation rates, a degree of variability remains in the intermodule orientation. However, the vast majority of $^1\text{F1}^2\text{F1-B3}$ conformers exhibit an elongated module arrangement, the orientation of which is significantly better defined in the complex with B3 than in the free form, indicating that B3 might tether the two F1 modules. The tandem β -zipper provided the first structural insight into the molecular basis of microbial peptide–fibronectin interactions. It not only revealed the mode of interaction, but also identified the length and nature of $^1\text{F1}$ - and $^2\text{F1}$ -binding motifs in B3. Armed with this information and a multiple sequence alignment of bacterial peptides, the authors proposed a novel “extended tandem β -zipper” model for binding of *S. pyogenes* and *S. aureus* to the fibronectin amino terminal domain, a theory which was subsequently supported by NMR and isothermal titration calorimetry data.¹²

The collagen-binding domain of fibronectin has been mapped to the six module fragment $^6\text{F1}^1\text{F2}^2\text{F2}$ – $^7\text{F1}^8\text{F1}^9\text{F1}$ on the basis of its affinity for heat-denatured collagen (gelatin).⁴⁰ With the exception of $^6\text{F1}$, each of these modules binds gelatin, but with greatly reduced affinity as compared to that of the intact domain. The way in which the individual modules combine to form a biologically active entity has been probed through a “dissect and build” strategy, in which the structural and functional information from overlapping module pairs is combined to construct a model of the intact domain. The solution structure of the $^6\text{F1}^1\text{F2}$ module pair was determined by NMR from NOEs (including 18 intermodule distances) and backbone dihedral angle restraints,⁴⁵ and was later refined using the dependence of ^{15}N relaxation on rotational diffusion anisotropy.²⁶ The intermodule linker was found to be flexible on the picosecond to nanosecond time-scale, as shown by its low heteronuclear NOE. The two modules interact through a small hydrophobic interface involving just two residues from $^6\text{F1}$ and one from $^1\text{F2}$ which lies on the opposite side of the module to the gelatin-binding site. In contrast, the complete lack of intermodule NOEs or significant chemical shift perturbations in the overlapping module pair $^1\text{F2}^2\text{F2}$ indicated that these two F2 modules show no discernible interaction other than their covalent linkage.²⁵ Subsequently, the $^6\text{F1}^1\text{F2}^2\text{F2}$ module triplet was found to possess greatly enhanced gelatin-binding activity over $^1\text{F2}^2\text{F2}$, an enhancement that could not simply be attributed to covalent linkage of the (nonbinding) $^6\text{F1}$ module.¹⁴ A comparison of amide chemical shifts in the $^6\text{F1}^1\text{F2}$, $^1\text{F2}^2\text{F2}$, and $^6\text{F1}^1\text{F2}^2\text{F2}$ fragments revealed that the $^6\text{F1}$ and $^2\text{F2}$ modules interact despite being separated in sequence

by the 1F_2 module. The solution structure of ${}^6F_1{}^1F_2{}^2F_2$ was calculated from NOEs (including 41 6F_1 – 2F_2 intermodule distances) and backbone dihedral angle restraints¹⁴ and revealed an extensive hydrophobic interface between the noncontiguous 6F_1 and 2F_2 modules (Figure 2). This was the first high-resolution study to reveal a compact, globular arrangement of modules in fibronectin and disputed the long-held view that fibronectin is simply a linear “string of beads”.

As can be seen in Table 1, the fibronectin type 3 module, F_3 , is very common and appears in a wide range of proteins. In fibronectin, F_3 modules bind a wide range of ligands, including heparin and integrins, and they are also involved in the formation of fibronectin fibrils (fibrillogenesis). Various studies have shown that ${}^{13}F_3$ and ${}^{14}F_3$ contain a heparin-binding site. Mapping of the heparin-binding sites of ${}^{13}F_3$, ${}^{14}F_3$, ${}^{13}F_3$, and ${}^{14}F_3$ by NMR chemical shift perturbation, isothermal titration calorimetry, and molecular modeling showed that ${}^{13}F_3$ provides the dominant heparin-binding site and that the residues involved are within the first 29 amino acids of ${}^{13}F_3$. This binding site involves positively charged residues that project into the solvent from the ABE face of the triple-stranded β -sheet on ${}^{13}F_3$. In contrast, ${}^{14}F_3$, although previously implicated in heparin binding, does not appear to contribute significantly.⁴⁶

One of many interesting features of fibronectin is that it is more active in binding to integrins when it is in the extracellular matrix than when it is in plasma. Various hydrodynamic and other studies have suggested that this is because it is in a compact “pretzel” form in plasma and an extended form in the matrix (see, e.g., ref 47). An ability to change shape like this requires considerable intermodule flexibility. The 9F_3 ${}^{10}F_3$ pair is mainly responsible for binding to cell surfaces via integrins. Particularly important is ${}^{10}F_3$ that has a loop between β -strands containing the residues Arg·Gly·Asp (RGD). We determined, some time ago, the structure of ${}^{10}F_3$ and showed that the RGD loop was flexible in solution.⁴⁸ For some integrins, such as $\alpha_5\beta_1$, an additional “synergy” site on 9F_3 is needed before full binding and biological activity of fibronectin is realized. There is a crystal structure of the 7F_3 8F_3 9F_3 ${}^{10}F_3$ region that has clear electron density for the RGD loop, indicating that it is rigid in the crystal.⁴⁹ The relative orientation of 9F_3 with respect to ${}^{10}F_3$ is also well-defined in the crystal, although the two modules have a distinctly different relative orientation as compared to other F_3 pairs, and, at 333 Å², the buried surface area between the modules is relatively small as compared to the 7F_3 8F_3 (587 Å²) and 8F_3 9F_3 interfaces (527 Å²). NMR studies of the 9F_3 ${}^{10}F_3$ pairs from human⁵⁰ and mouse⁵¹ fibronectin have shown that the RGD loop remains relatively flexible in solution and that the modules have a significant degree of intermodule flexibility. Enhancing the 9F_3 ${}^{10}F_3$ intermodule flexibility still further by the extending the intermodule linker results in a reduction in integrin-mediated cell attachment and spreading.^{50,52} These structural studies, and those on the proteins that regulate complement activation (below), clearly show the comple-

mentary nature of NMR and X-ray crystal studies.

Another very interesting region that contains F_3 modules is the 1F_3 2F_3 region. This region has been implicated in fibrillogenesis, a process that is initiated by addition of partially or completely denatured F_3 modules or by mechanical tension (reviewed in ref 53). A 76 aa fragment from 1F_3 , termed anastellin, obtained by removing strands A and B promotes assembly of a super fibronectin structure.⁵⁴ The structure of anastellin was recently solved by NMR,⁵⁵ as was the structure of the intact 1F_3 .⁵⁶ Steered MD simulations were combined with these structures to investigate the mechanical unfolding pathway of 1F_3 . It was shown that 1F_3 could unfold to form a stable intermediate, 4 times the length of the native folded state that is a plausible intermediate in fibrillogenesis.⁵⁶

7. Proteins Involved in the Regulation of Complement Activation

The complement system is a group of essential circulating plasma proteins that inactivate and dispose of foreign intruders. This defensive process is achieved through the binding of complement to foreign cell membranes (complement fixation) and their subsequent lysis and phagocytosis. The system relies upon cascades of specific protein–protein interactions and catalytic events that lead to the formation (and destruction) of multiprotein complexes.⁵⁷ The homologous group of proteins that control these events, the regulators of complement activation (RCA), are comprised almost entirely of “complement control protein” (CCP) modules that are approximately 60 amino acid residues in length. The CCP modules are characterized by a consensus sequence that includes four invariant cysteines that stabilize the module’s tertiary structure of short β -sheets and turns. Functional and mutagenesis studies have shown that in most cases two or more neighboring CCP modules form specific binding sites for other molecules. Hence, the orientation in space of a CCP module with respect to its neighbors and the flexibility of the intermodular junction are likely to be critical for function. Recent NMR studies on the conformation and dynamics of CCP module pairs of the RCA proteins decay accelerating factor (DAF), complement receptor 1 (CR1), and vaccinia virus complement control protein (VCP) have shed light on these structure–function relationships.

The GPI-anchored DAF is a ubiquitous RCA that accelerates the dissociation of the C3 and C5 convertases of the classical and alternative pathways and hence protects host cell surfaces from activation of autologous complement.⁵⁷ DAF comprises four tandem CCP modules followed by an O-glycosylated serine/threonine-rich domain. Classical pathway convertase regulation has been mapped to CCP modules 2 and 3. Recently, the structure of the 2nd and 3rd CCP module pair from DAF (DAF~2,3) was determined by solution NMR.^{58,59} The structure determination incorporated restraints from interproton NOEs, backbone dihedral angles, and H^N – N^H residual dipolar couplings (RDCs). The interproton NOE list included six intermodule restraints and 27 module–

linker restraints. The DAF~2,3 structures were relatively elongated with an end-to-end arrangement of well-defined modules that share only a small interface. However, the orientation of one module with respect to the other was not defined well by the data, with two populations of tilt angle that were indistinguishable on the basis of potential energy (agreement with experimental data). Furthermore, the RDC data were inconsistent with a rigid connection between the two modules, as use of a common alignment tensor resulted in a poorer fit to the experimental data. Interestingly, residues within the intermodule linker did not have unusual ^{15}N relaxation properties (long ^{15}N T_1 and T_2 values, and a low $^1\text{H},^{15}\text{N}$ NOE) as might be expected if this region were mobile on the fast (picosecond to nanosecond) or intermediate (microsecond to millisecond) time-scales.

The apparent flexibility of the DAF 2~3 interface in NMR studies disagrees with the recently determined crystal structure of the four CCP module fragment DAF~1,2,3,4.⁶⁰ In this study, three differently packed crystal forms were analyzed, and a total of eight independent models of DAF~1,2,3,4 were built. Each model was found to be a linear molecule with the four CCP modules forming an extended rod with overall dimensions of $160 \times 50 \times 30 \text{ \AA}$. Despite the different crystal packing environments, there was very little variation in the DAF~2,3 intermodule interface. It could be argued that this apparent rigidity is due to the high ionic strength of the mother liquor used in the crystallization process (25-fold higher than that used for the NMR studies) or the high concentration of precipitant, both of which may promote hydrophobic interactions of the type seen in the DAF~2,3 interface. Thus, the crystallization process may have taken a “snapshot” of a molecule that is inherently dynamic in solution. However, an analytical ultracentrifugation study under physiological solution conditions revealed the sedimentation coefficient of DAF~1,2,3,4 to be almost identical to that calculated from the crystal coordinates.⁶⁰ Therefore, the structure in the crystalline environment is probably a good representation of the conformation present in solution.

Activation of complement leads to deposition of C4b and C3b on a target. Once the antigenic particle is coated with these molecules, it is recognized by complement receptor type 1 (CR1 or CD35) on the surface of host blood cells.⁵⁷ CR1 is a 250 kDa transmembrane glycoprotein with an extracellular domain that comprises 30 CCP modules with two functionally distinct sites. Two nearly identical copies of “site 2” (localized to modules 8–10 and 15–17) bind C3b and C4b. The intermodule orientations, module–module interactions, and flexibility of site 2 (CR1~15–17) has been investigated by NMR to better understand the molecular basis of C3b/C4b-recognition.⁶¹ A comparison of the backbone chemical shifts of CR1~15–17 with the overlapping 15,16 and 16,17 module pairs revealed that there were no long-range interactions between the nonsequential modules CR1~15 and CR1~17, thus implying an extended conformation for the triple-module fragment. Hence,

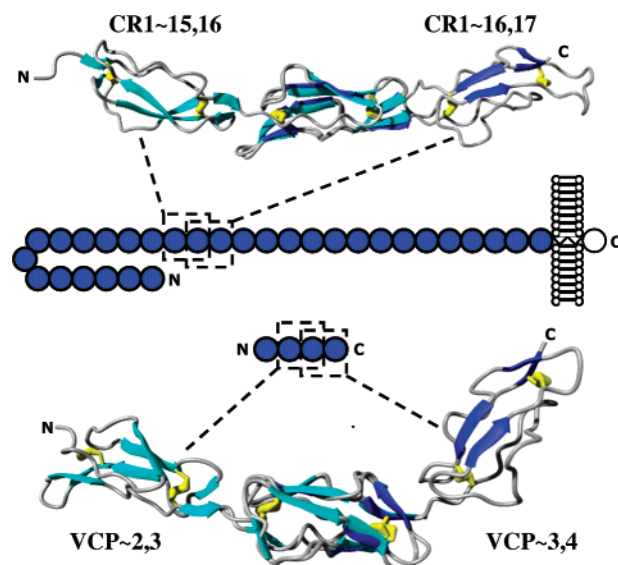


Figure 3. Determination of the structures of CCP module-containing proteins by NMR using a “dissect and build” strategy. Ribbon diagrams of the lowest-energy structures of CR1~15,16 and CR1~16,17,⁶¹ and of VCP~2,3⁶³ and VCP~3,4⁶² are shown mapped onto the mosaic structures of CR1 and VCP, respectively. The overlapping fragments have been overlaid over the backbone heavy atoms of their common module (i.e., CR1~16 or VCP~3). Regions of secondary structure are shown in light (CR1~15,16 and VCP~2,3) or dark (CR1~16,17 and VCP~3,4) blue, with disulfide bonds in yellow.

a strategy was adopted in which the NMR solution structures of the overlapping CR1~15,16 and 16,17 module pairs were combined to “reconstruct” CR1~15–17.⁶¹ For CR1~15,16, the unambiguous NOE restraints included 13 intermodule and 94 module–linker distances giving a relatively well-defined orientation of the two modules. Furthermore, ^1H – ^{15}N RDCs from the two modules were found to be consistent with a common alignment tensor, thus supporting a relatively fixed conformation. The two modules of CCP~15,16 are assembled in a head-to-tail fashion through interactions between the *DE* loop of module 15 and strand *F* of module 16 (Figure 3). In contrast, the unambiguous NOE restraints for CR1~16,17 included only 4 intermodule and 68 module–linker distances with the result that the intermodule interface of CR1~16,17 is poorly defined with few hydrophobic contacts. This lack of definition in intermodule orientation may reflect flexibility between CCPs 16 and 17, but neither the CR1~15–16 nor the CR1~16,17 linker peptides appear to be flexible on the microsecond to millisecond or picosecond to nanosecond time-scales according to long longitudinal (T_1) and transverse (T_2) relaxation times and heteronuclear NOE measurements.⁶¹ In the model of CR1~15–17, reconstructed from the solution structures of the overlapping CR1~15,16 and 16,17 module pairs, the three modules trace a gentle curve with equivalent surfaces of the three modules facing in different directions (Figure 3).⁶¹ The elongated shape of site 2 results in a large solvent-exposed surface area with potential for extensive contacts with binding partners. The structure allowed rationalization of a large body of mutagenesis data and design of new, structure-guided mutagenesis

experiments. These new experiments identified a positively charged surface region on module 15 that is critical for C4b binding. This patch, together with basic side chains of module 16 exposed on the same face of CR1, is required for interaction with C3b.⁶¹

VCP binds host C3b and C4b, thus inhibiting both the classical and the alternative pathways and defending the virus against attack by the host complement system.⁵⁷ VCP is a homologue of mammalian regulators of complement activation, consisting entirely of four CCP modules that show high sequence similarity to modules 15–18 (encompassing site 2) of mammalian CR1. It is the smallest and least complex of the complement control proteins, with no glycosylation or covalent membrane attachment sites and only 244 residues, making it an excellent candidate for complete structural and dynamic analysis by NMR. To date, two overlapping fragments of VCP, the C-terminal module pair VCP~3,4 and the central module pair VCP~2,3, have been characterized. The solution structure of VCP~3,4 was calculated using backbone *J* couplings and an NOE restraint list that included 90 intermodule distances.⁶² These restraints resulted in a relatively well-defined orientation of VCP~3 and VCP~4 with the long axis of module 4 tilted by $59^\circ \pm 4^\circ$ with respect to that of module 3, and twisted with respect to module 3 by $22^\circ \pm 6^\circ$ (Figure 3). The intermodule interface is predominantly hydrophobic with interactions between strand *E* of module 3 and the *CD* and *FG* loops of module 4. The amide proton T_1 relaxation times of VCP~3,4 were consistent with only limited flexibility at the module–module interface on the picosecond to nanosecond time-scale. The linker was as precisely defined as other well-structured regions of the protein. In contrast to the intimate intermodule contact seen in VCP~3,4, there was no extensive interface observed in VCP~2,3; just two intermodule and 40 module–linker NOEs were identified.⁶³ Consequently, only a preferred, elongated orientation of the two modules was observed. The rotational diffusion anisotropy of the module pair (calculated from ¹⁵N relaxation data) indicated that the time-averaged structure is more compact than suggested by the NOEs. These data are consistent with the presence of many intermodule orientations, some of which are kinked, undergoing interconversion on the nanosecond to microsecond time-scale. Thus, the NMR data strongly support the idea of a flexible intermodule hinge in VCP~2,3 at 37 °C. However, at lower temperatures, other biophysical techniques have showed that VCP~2,3 becomes progressively more elongated (unpublished data cited in ref 57), presumably due to the stabilization of a specific intermodular junction. Indeed, it is likely that such interface stabilization was a contributing factor in the recent crystallization of intact VCP at 20 °C.⁶⁴ The crystal structure of VCP at 2.2 Å resolution provided the first atomic coordinates of any intact regulator of complement activation and permitted a critical analysis of the previous NMR results. Five crystallographically independent examples of the molecule within the unit cells of two crystal forms were analyzed, and in each the molecule had an elongated structure with no intramo-

lecular contacts between nonsequential modules, and a near-identical arrangement of CCP modules. This suggests that crystal packing interactions are not responsible for the conformation observed, but rather that a strongly preferred conformation of VCP exists in solution under the conditions used to grow the crystals. The interface between modules 3 and 4 is very similar to that modeled from NMR data previously⁶² with 408 Å² of buried surface area. The tilt and twist angles between modules 3 and 4 of the NMR structure of VCP~3,4 are comparable to those seen in the crystal structure, thus validating the dissect-and-build approach in this case. In contrast, with 281 Å² of buried surface area and a well-defined modular orientation, the 2–3 interface in the crystal structure is very different from that characterized by solution NMR.^{63,65} However, as in the case of DAF~2,3, the higher ionic strength and precipitant content of the mother liquor used for crystallization may have induced intermodular hydrophobic contacts between VCP modules 2 and 3. Furthermore, given the known temperature dependence of the VCP~2,3 conformation (above), and the difference in temperature of the two studies (37 °C for NMR cf. 20 °C for crystallography), the discrepancy in conformation is not surprising in this case.

8. Conclusions

Several modular extracellular proteins and their various interactions have been discussed. We have tried to illustrate how modular proteins can readily assemble into dynamic, functional protein complexes. Studies of extracellular proteins constructed from numerous repeated modules allow some general features to be recognized. One type of module, for example, F3, can provide a wide range of binding functions, although the modules have very similar structures. In general, the modules do not change conformation themselves but change their orientation with respect to other modules to change overall shape and activity. This regulation property can be facilitated by having functional binding sites on more than one module. Changes in module orientation can be induced by ligand binding or environmental change, and in some cases new “cryptic” sites may be exposed by module rearrangement or applied mechanical tension. NMR is particularly well suited to investigate intermodule dynamics and weak ligand interactions.

9. Acknowledgments

We thank the Wellcome Trust and BBSRC for their support.

10. References

- (1) Koonin, E. V.; Wolf, Y. I.; Karev, G. P. *Nature* **2002**, *420*, 218.
- (2) Bork, P.; Downing, A. K.; Kieffer, B.; Campbell, I. D. *Q. Rev. Biophys.* **1996**, *29*, 119.
- (3) Copley, R. R.; Doerks, T.; Letunic, I.; Bork, P. *FEBS Lett.* **2002**, *513*, 129.
- (4) Staunton, D.; Owen, J.; Campbell, I. D. *Acc. Chem. Res.* **2003**, *36*, 207.
- (5) Campbell, I. D. *Immunol. Rev.* **1998**, *163*, 11.
- (6) Zuiderweg, E. R. P. *Biochemistry* **2002**, *41*, 1.

- (7) Fiaux, J.; Bertelsen, E. B.; Horwich, A. L.; Wuthrich, K. *Nature* **2002**, *418*, 207.
- (8) Nishida, N.; Sumikawa, H.; Sakakura, M.; Shimba, N.; Takahashi, H.; Terasawa, H.; Suzuki, E.; Shimada, I. *Nat. Struct. Biol.* **2003**, *10*, 53.
- (9) Hoellerer, M. K.; Noble, M. E. M.; Labesse, G.; Campbell, I. D.; Werner, J. M.; Arold, S. T. *Structure* **2003**, *11*, 1207.
- (10) Ikura, M.; Clore, G. M.; Gronenborn, A. M.; Zhu, G.; Klee, C. B.; Bax, A. *Science* **1992**, *256*, 632.
- (11) Breeze, A. L. *Prog. Nucl. Magn. Reson. Spectrosc.* **2000**, *36*, 323.
- (12) Schwarz-Linek, U.; Werner, J. M.; Pickford, A. R.; Gurusiddappa, S.; Kim, J. H.; Pilka, E. S.; Briggs, J. A. G.; Gough, T. S.; Hook, M.; Campbell, I. D.; Potts, J. R. *Nature* **2003**, *423*, 177.
- (13) Dominguez, C.; Boelens, R.; Bonvin, A. *J. Am. Chem. Soc.* **2003**, *125*, 1731.
- (14) Pickford, A. R.; Smith, S. P.; Staunton, D.; Boyd, J.; Campbell, I. D. *EMBO J.* **2001**, *20*, 1519.
- (15) Bax, A.; Kontaxis, G.; Tjandra, N. *Nuclear Magnetic Resonance of Biological Macromolecules, Pt B*; Academic Press: San Diego, CA, 2001; Vol. 339.
- (16) Fischer, M. W. F.; Losonczi, J. A.; Weaver, J. L.; Prestegard, J. H. *Biochemistry* **1999**, *38*, 9013.
- (17) Dosset, P.; Hus, J. C.; Marion, D.; Blackledge, M. *J. Biomol. NMR* **2001**, *20*, 223.
- (18) Ulmer, T. S.; Werner, J. M.; Campbell, I. D. *Structure* **2002**, *10*, 901.
- (19) Braddock, D. T.; Cai, M. L.; Baber, J. L.; Huang, Y.; Clore, G. M. *J. Am. Chem. Soc.* **2001**, *123*, 8634.
- (20) Mattinen, M. L.; Paakkonen, K.; Ikonen, T.; Craven, J.; Drakenberg, T.; Serimaa, R.; Waltho, J.; Annala, A. *Biophys. J.* **2002**, *83*, 1177.
- (21) Bruschweiler, R.; Liao, X. B.; Wright, P. E. *Science* **1995**, *268*, 886.
- (22) Clore, G. M.; Gronenborn, A. M.; Szabo, A.; Tjandra, N. *J. Am. Chem. Soc.* **1998**, *120*, 4889.
- (23) Ghose, R.; Fushman, D.; Cowburn, D. *J. Magn. Reson.* **2001**, *149*, 204.
- (24) Fushman, D.; Xu, R.; Cowburn, D. *Biochemistry* **1999**, *38*, 10225.
- (25) Smith, S. P.; Hashimoto, Y.; Pickford, A. R.; Campbell, I. D.; Werner, J. M. *Biochemistry* **2000**, *39*, 8374.
- (26) Hashimoto, Y.; Smith, S. P.; Pickford, A. R.; Bocquier, A. A.; Campbell, I. D.; Werner, J. M. *J. Biomol. NMR* **2000**, *17*, 203.
- (27) Baber, J. L.; Szabo, A.; Tjandra, N. *J. Am. Chem. Soc.* **2001**, *123*, 3953.
- (28) Braddock, D. T.; Louis, J. M.; Baber, J. L.; Levens, D.; Clore, G. M. *Nature* **2002**, *415*, 1051.
- (29) Handford, P. A.; Downing, A. K.; Reinhardt, D. P.; Sakai, L. Y. *Matrix Biol.* **2000**, *19*, 457.
- (30) Dietz, H. C.; Pyeritz, R. E. *Hum. Mol. Genet.* **1995**, *4*, 1799.
- (31) Downing, A. K.; Knott, V.; Werner, J. M.; Cardy, C. M.; Campbell, I. D.; Handford, P. A. *Cell* **1996**, *85*, 597.
- (32) Yuan, X. M.; Downing, A. K.; Knott, V.; Handford, P. A. *EMBO J.* **1997**, *16*, 6659.
- (33) Handford, P.; Downing, A. K.; Rao, Z. H.; Hewett, D. R.; Sykes, B. C.; Kielty, C. M. *J. Biol. Chem.* **1995**, *270*, 6751.
- (34) Whiteman, P.; Downing, A. K.; Handford, P. A. *Protein Eng.* **1998**, *11*, 957.
- (35) Kettle, S.; Cardy, C. M.; Yuan, X. M.; Downing, A. K.; Handford, P. *Matrix Biol.* **1997**, *16*, 81.
- (36) Smallridge, R. S.; Whiteman, P.; Doering, K.; Handford, P. A.; Downing, A. K. *J. Mol. Biol.* **1999**, *286*, 661.
- (37) Knott, V.; Downing, A. K.; Cardy, C. M.; Handford, P. *J. Mol. Biol.* **1996**, *255*, 22.
- (38) Smallridge, R. S.; Whiteman, P.; Werner, J. M.; Campbell, I. D.; Handford, P. A.; Downing, A. K. *J. Biol. Chem.* **2003**, *278*, 12199.
- (39) Werner, J. M.; Knott, V.; Handford, P. A.; Campbell, I. D.; Downing, A. K. *J. Mol. Biol.* **2000**, *296*, 1065.
- (40) Potts, J. R.; Campbell, I. D. *Curr. Opin. Cell Biol.* **1994**, *6*, 648.
- (41) Williams, M. J.; Phan, I.; Harvey, T. S.; Rostagno, A.; Gold, L. L.; Campbell, I. D. *J. Mol. Biol.* **1994**, *235*, 1302.
- (42) Phan, I. Q. H.; Boyd, J.; Williams, M. J.; Kieffer, B. M.; Campbell, I. D. *J. Cell. Biochem.* **1995**, *34*.
- (43) Phan, I. Q. H.; Boyd, J.; Campbell, I. D. *J. Biomol. NMR* **1996**, *8*, 369.
- (44) Potts, J. R.; Bright, J. R.; Bolton, D.; Pickford, A. R.; Campbell, I. D. *Biochemistry* **1999**, *38*, 8304.
- (45) Bocquier, A. A.; Potts, J. R.; Pickford, A. R.; Campbell, I. D. *Struct. Folding Des.* **1999**, *7*, 1451.
- (46) Sachchidanand; Lequin, O.; Staunton, D.; Mulloy, B.; Forster, M. J.; Yoshida, K.; Campbell, I. D. *J. Biol. Chem.* **2002**, *277*, 50629.
- (47) Hynes, R. O. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 2588.
- (48) Main, A. L.; Harvey, T. S.; Baron, M.; Boyd, J.; Campbell, I. D. *Cell* **1992**, *71*, 671.
- (49) Leahy, D. J.; Aukhil, I.; Erickson, H. P. *Cell* **1996**, *84*, 155.
- (50) Spitzfaden, C.; Grant, R. P.; Mardon, H. J.; Campbell, I. D. *J. Mol. Biol.* **1997**, *265*, 565.
- (51) Copie, V.; Tomita, Y.; Akiyama, S. K.; Aota, S.; Yamada, K. M.; Venable, R. M.; Pastor, R. W.; Krueger, S.; Torchia, D. A. *J. Mol. Biol.* **1998**, *277*, 663.
- (52) Altroff, H.; van der Walle, C. F.; Asselin, J.; Fairless, R.; Campbell, I. D.; Mardon, H. J. *J. Biol. Chem.* **2001**, *276*, 38885.
- (53) Geiger, B.; Bershady, A.; Pankov, R.; Yamada, K. M. *Nat. Rev. Mol. Cell Biol.* **2001**, *2*, 793.
- (54) Yi, M.; Ruoslahti, E. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 620.
- (55) Briknarova, K.; Akerman, M. E.; Hoyt, D. W.; Ruoslahti, E.; Ely, K. R. *J. Mol. Biol.* **2003**, *332*, 205.
- (56) Gao, M.; Craig, D.; Lequin, O.; Campbell, I. D.; Vogel, V.; Schulten, K. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 14784.
- (57) Kirkitadze, M. D.; Barlow, P. N. *Immunol. Rev.* **2001**, *180*, 146.
- (58) Uhrinova, S.; Lin, F.; Ball, G.; Bromek, K.; Uhrin, D.; Medof, M. E.; Barlow, P. N. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 4718.
- (59) Uhrinova, S.; Lin, F.; Uhrin, D.; Medof, M. E.; Barlow, P. N. *J. Biomol. NMR* **2002**, *23*, 167.
- (60) Lukacik, P.; Roversi, P.; White, J.; Esser, D.; Smith, G. P.; Billington, J.; Williams, P. A.; Rudd, P. M.; Wormald, M. R.; Harvey, D. J.; Crispin, M. D. M.; Radcliffe, C. M.; Dwek, R. A.; Evans, D. J.; Morgan, B. P.; Smith, R. A. G.; Lea, S. M. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 1279.
- (61) Smith, B. O.; Mallin, R. L.; Krych-Goldberg, M.; Wang, X. F.; Hauhart, R. E.; Bromek, K.; Uhrin, D.; Atkinson, J. P.; Barlow, P. N. *Cell* **2002**, *108*, 769.
- (62) Wiles, A. P.; Shaw, G.; Bright, J.; Perczel, A.; Campbell, I. D.; Barlow, P. N. *J. Mol. Biol.* **1997**, *272*, 253.
- (63) Henderson, C. E.; Bromek, K.; Mullin, N. P.; Smith, B. O.; Uhrin, D.; Barlow, P. N. *J. Mol. Biol.* **2001**, *307*, 323.
- (64) Murthy, K. H. M.; Smith, S. A.; Ganesh, V. K.; Judge, K. W.; Mullin, N.; Barlow, P. N.; Ogata, C. M.; Kotwal, G. J. *Cell* **2001**, *104*, 301.
- (65) Kirkitadze, M. D.; Henderson, C.; Price, N. C.; Kelly, S. M.; Mullin, N. P.; Parkinson, J.; Dryden, D. T. F.; Barlow, P. N. *Biochem. J.* **1999**, *344*, 167.
- (66) Sharma, A.; Askari, J. A.; Humphries, M. J.; Jones, E. Y.; Stuart, D. I. *EMBO J.* **1999**, *18*, 1468.
- (67) Whiteman, P.; Downing, A. K.; Smallridge, R.; Winship, P. R.; Handford, P. A. *J. Biol. Chem.* **1998**, *273*, 7807.
- (68) Kettle, S.; Yuan, X. M.; Grundy, G.; Knott, V.; Downing, A. K.; Handford, P. A. *J. Mol. Biol.* **1999**, *285*, 1277.

CR0304018

