

IMPLICATIONS OF ATOMIC-RESOLUTION STRUCTURES FOR CELL ADHESION

Daniel J. Leahy

Department of Biophysics and Biophysical Chemistry, Johns Hopkins University
School of Medicine, 725 North Wolfe Street, Baltimore, Maryland 21205; e-mail:
leahy@groucho.med.jhmi.edu

KEY WORDS: cadherin, integrin, fibronectin, immunoglobulin

ABSTRACT

Molecules involved in cell adhesion processes are often both structurally and functionally modular, with subdomains that are members of large protein families. Recently, high-resolution structures have been determined for representative members of many of these families including fragments of integrins, cadherins, fibronectin-like domains, and immunoglobulin-like domains. These structures have enhanced our understanding of cell adhesion processes at several levels. In almost all cases, ligand-binding sites have been visualized and provide insight into how these molecules mediate biologically important interactions. Metal-binding sites have been identified and characterized, allowing assessment of the role of bound ions in cell adhesion processes. Many of these structures serve as templates for modeling homologous domains in other proteins or, when the structure of a fragment consisting of more than one domain is determined, the structure of multidomain arrays of homologous domains. Knowledge of atomic structure also allows rational design of drugs that either mimic or target specific binding sites. In many cases, high-resolution structures have revealed unexpected relationships that pose questions about the evolutionary origin of specific domains. This review briefly describes several recently determined structures of cell adhesion molecules, summarizes some of the main results of each structure, and highlights common features of different systems.

CONTENTS

INTRODUCTION	364
DESCRIPTION AND CLASSIFICATION OF β -SANDWICH DOMAINS INVOLVED	
IN CELL ADHESION	367
<i>Immunoglobulin Domains</i>	367
<i>Fibronectin Domains</i>	371
<i>Cadherin Domains</i>	372
<i>Divergent versus Convergent Evolutionary Relationships</i>	372
IMPLICATIONS OF ATOMIC-RESOLUTION STRUCTURES FOR CELL ADHESION ..	374
<i>CD2</i>	374
<i>VCAM Domains</i>	376
<i>Peripheral Myelin Protein Zero</i>	376
<i>FN-III Domains</i>	378
<i>Cadherin Domains</i>	381
<i>Integrin I Domains</i>	383
SUMMARY AND PERSPECTIVE	386

INTRODUCTION

The aim of this review is to survey the contributions of high-resolution structure determinations to the understanding of the mechanisms by which cells adhere to one another and to components of the extracellular matrix. Several large families of related proteins contain many of the molecules involved in cell adhesion processes, and representative structures from a growing number of these families have recently been solved. These structures include ligand- or receptor-binding fragments of integrins (Lee et al 1995a,b, Qu & Leahy 1995, 1996), cadherins (Shapiro et al 1995a, Nagar et al 1996, Overduin et al 1995, 1996), fibronectin-like domains (Constantine et al 1992, DeVos et al 1992, Main et al 1992, Leahy et al 1992b, 1996, Potts & Campbell 1994, Williams et al 1994, Dickinson et al 1994, Huber et al 1994), and immunoglobulin superfamily (IgSF) members (Driscoll et al 1991, Jones et al 1992, 1995, Withka et al 1993, Bodian et al 1994, Murray et al 1995, Wang et al 1995, Thomsen et al 1996). This list will continue to expand quickly and could easily contain many additional molecules at present, but we limit ourselves to the structures noted above to allow illustration of several general points while preserving focus. Before proceeding to a brief description of these structures and their implications, the use of structure determination as a tool for understanding molecular processes is discussed.

In the last decade, advances in methodology and apparatus have transformed structure determination by X-ray diffraction and nuclear magnetic resonance (NMR) techniques from lengthy and highly specialized endeavors to astonishingly routine procedures. Coupled with constantly improving protein expression technologies, this transformation has resulted in an exponentially increasing number of structures being reported with no slowdown in sight (Bernstein et al

1978). Structure determinations are regularly undertaken at very early stages of biochemical and biological characterization of proteins and play variable and often unexpected roles in the discovery process.

As the structural database grows, it is becoming increasingly apparent that tertiary structure is more conserved than amino acid sequence. Structure determinations frequently uncover unexpected homologies that shed insight into the function or origins of a particular protein. Indeed, of 120 new structures reported in 1995 for which no sequence homology to known proteins could be detected, only 40 were judged to represent new protein folds (Murzin 1996). Some structural homologies such as that found between the integrin I domains (Lee et al 1995a, Qu & Leahy 1995) and members of the Rossman-fold family of proteins (Rossman et al 1974) almost certainly represent cases of divergent evolution, whereas an evolutionary relationship is less clear in cases such as the structurally similar immunoglobulin (Ig), fibronectin type III (FN-III), and cadherin domains (Leahy et al 1992b, Main et al 1992, Overduin et al 1995, Shapiro et al 1995b, Wagner 1995).

Aside from the potential for discovery of unanticipated homologies, two primary objectives motivate structural investigations of cell adhesion molecules. The first involves identification and visualization of ligand- or metal-binding sites. If a binding site can be identified from specific amino acids known to be involved in ligand binding, a structure allows the extent and chemical properties of the site to be assessed. Additional mutagenesis can be carried out to further delineate an interaction site, and drugs that either mimic or bind to the site may be rationally designed. If a ligand-interaction site is unknown, surface-exposed residues can be identified and effects of mutations at these sites assayed. Identification and characterization of ligand-binding sites have been important components of the analysis of each structure described in this review. A second major motivation to determine structures of cell adhesion molecules stems from the involvement of many cell adhesion molecules in signal transduction. High-resolution structures of signaling molecules are being pursued with and without bound ligand, as well as in active and inactive states when such forms of the protein can be induced or isolated. It is hoped that these studies will uncover potential molecular mechanisms underlying signal transduction. Structural studies have yet to definitively establish a mechanism by which a cell adhesion interaction transmits signals, but analysis of atomic resolution structures is likely to be required for an understanding of signal transduction in many cell adhesion molecules, most notably the integrins (Schwartz et al 1995).

Recently, an additional source of insight into cell adhesion processes has emerged from several high-resolution structure determinations. Inspection of crystal lattice interactions in crystals of cadherins (Shapiro et al 1995a, Nagar et al 1996), CD2 (Jones et al 1992, Bodian et al 1994), and peripheral myelin

protein zero (P0) (Shapiro et al 1996) have provided models of how these molecules may interact with either themselves or homologues in a physiological setting. It is not surprising that lattice interactions reflect native interactions for homotypic cell adhesion molecules such as cadherins and P0, but lattice interactions can on occasion provide more fortuitous functional clues. For example, the binding of a glutamic acid side chain from one molecule of the CD11b I domain to a metal bound to an adjacent molecule in the crystal lattice may mimic interactions between integrin I domains and native ligands (Lee et al 1995a).

In addition to information about specific molecules, accumulated insight from many structure determinations has also proven illuminating, e.g. high-resolution structure determinations with important consequences for cell adhesion molecules have confirmed that amino acid sequence repeats often correspond to independently folded structural units. Although this correspondence has been observed for many systems, the large number, prominence, and size of Ig-like structures in the Ig superfamily clearly quickened the realization and acceptance of this idea (Williams & Barclay 1988, Bork et al 1994, Vaughn & Bjorkman 1996). The correspondence of sequence repeats to structural units is of particular relevance to cell adhesion because many components of the extracellular matrix, as well as adhesion receptors on the cell surface, are highly modular (Bork et al 1996). An excellent World Wide Web site maintained by Peer Bork catalogs many of the known modules and shows the domain organization of large numbers of module-containing proteins (Bork 1996; see <http://www.bork.embl-heidelberg.de/Modules/>). It is now routine to assume that homologous regions of amino acid sequence found either within a single protein or between different proteins represent independent structural units. Electron microscopy has also played an important role in confirming the modularity of large proteins by revealing extended molecules and, in some cases, resolving individual domains.

Recognition of the structural modularity of many cell adhesion proteins has simplified experimental approaches to their study. Individual modules can be tested for biological activities, and many functions map to one or a few protein subdomains. If a high-resolution structure for a given module exists, information derived from that structure often readily reveals answers for homologous domains in other proteins. Modeling of homologous domains is not necessarily confined to individual domains, however. Many domains appear in tandem arrays, and knowledge of the prototypical domain structure and how the component domains are likely to assemble into larger arrays allows construction of testable models of these arrays (Constantine et al 1992, Huber et al 1994, Leahy et al 1996, Nagar et al 1996).

DESCRIPTION AND CLASSIFICATION OF β -SANDWICH DOMAINS INVOLVED IN CELL ADHESION

Immunoglobulin Domains

The atomic structures of immunoglobulin (Ig) superfamily members have become so familiar that extensive discussion here would be superfluous. There are several thorough and excellent reviews of structures of Ig superfamily members and the immunoglobulin fold (Amzel & Poljak 1979, Williams & Barclay 1988, Bork et al 1994, Vaughn & Bjorkman 1996). Much of our knowledge of sequence-structure relationships among families of related domains originates from studies of the Ig superfamily, however, and this background has shaped our understanding of other families of cell adhesion molecules. The emergence of the Ig superfamily has not only cemented the notion that regions of homologous sequence often represent independently folded and structurally homologous protein modules but has provided a specific framework for characterizing the strikingly similar fibronectin type III and cadherin domains.

More than 30 years ago immunoglobulins were recognized to contain regions of related sequences, and an evolutionary relationship between these domains was postulated (Hill et al 1966, Singer & Doolittle 1966). Since that time, immunoglobulin-like sequences have been found in numerous proteins, and the concept of a superfamily of domains with a common evolutionary heritage has developed (Williams 1987, Williams & Barclay 1988). Although Ig superfamily members perform diverse biological roles, a primary function of most family members is binding a ligand. Since the classic work in which the initial structures of antibody fragments were determined (Poljak et al 1973, Schiffer et al 1973), atomic resolution structures of Ig or Ig-like domains have been determined from many proteins, including major histocompatibility molecules (Bjorkman et al 1987, Stern et al 1994), T-cell receptors (Bentley et al 1995, Fields et al 1995, Garboczi et al 1996, Garcia et al 1996), vascular cell adhesion molecule-1 (VCAM-1) (Jones et al 1995, Wang et al 1995), CD2 (Driscoll et al 1991, Jones et al 1992), CD4 (Ryu et al 1990, Wang et al 1990, Brady et al 1993), CD8 (Leahy et al 1992a), NCAM (Thomsen et al 1996), the muscle protein titin (Pfuhl & Pastore 1995), telokin (Holden et al 1992), a bacterial pilus-associated protein (Holmgren & Branden 1989), and dozens of additional antibody fragments (Bernstein et al 1978).

Ig domain structure consists of a sandwich of two antiparallel β sheets constructed from strands with a conserved topology (Figure 1). The strands in constant domains are labeled alphabetically in the order in which they occur in the amino acid sequence. A constellation of virtually invariant residues,

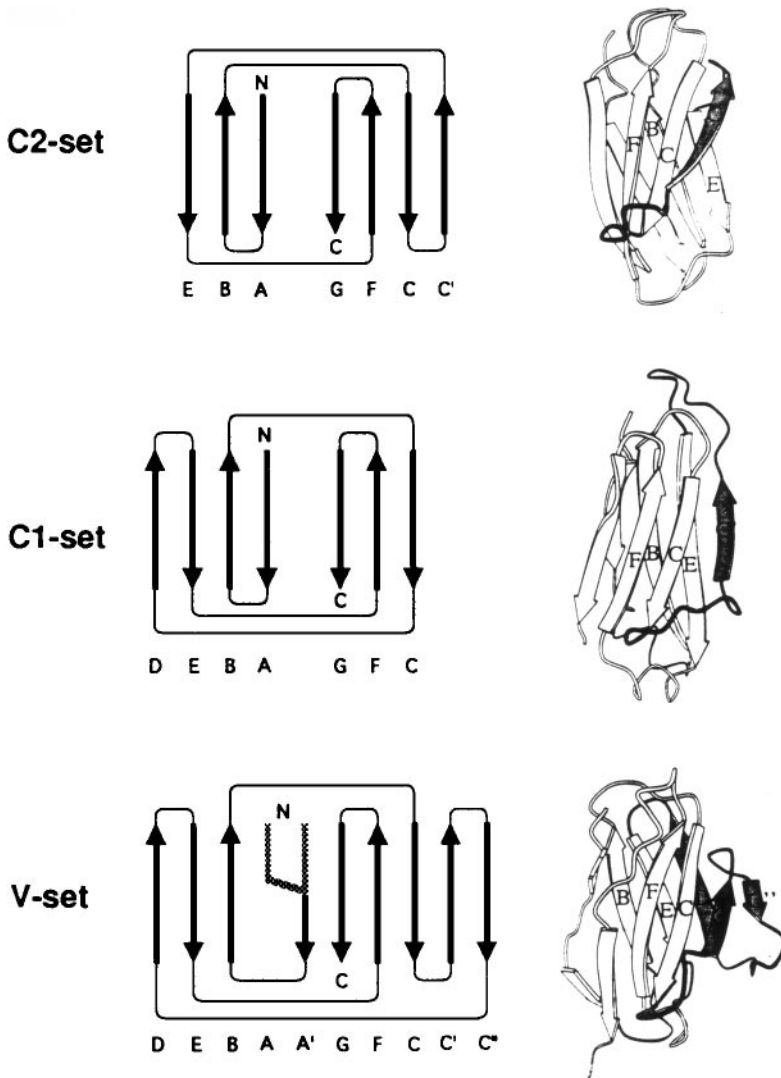


Figure 1 Schematic drawings of the strand topology of Ig C2-set, Ig C1-set, and Ig V-set domains are shown on the left. The two sheets of these domains are splayed as if they were a book lying face down on the page, and folding the sheets together as if closing this book will reproduce the proper three-dimensional relationship of the strands and sheets. On the right are MOLSCRIPT diagrams of representative structures of each type of domain. The core B, C, F, and E strands are labeled as are the variable C', C'', and D strands. These latter strands are also shaded to highlight the variable length and strand arrangement of this region. The Ig C2-set structure displayed is TNfn3 (Leahy et al 1992b), and the Ig C1-set and Ig V-set structures are the CL and VL domains from the AN02 Ig, respectively (Brunger et al 1991). The figures on the right were made with the program MOLSCRIPT (Kraulis 1991).

including several hydrophobic residues and an intersheet disulfide bond between cysteines on the B and F strands, is considered diagnostic of the Ig fold, but on rare occasions one or more of these invariant residues may be absent. Ig constant and variable domains are of different length (≈ 100 – 105 and ≈ 115 amino acids, respectively), and this difference involves a relative insertion or deletion of residues between the conserved cysteines on the B and F strands. Structures of Ig constant and variable domains reveal that this insertion corresponds to the addition of two antiparallel β strands on one edge of a β sheet. These strands are typically connected by a hairpin loop and form the second complementarity determining region (CDR2) of antibody variable domains. The two inserted strands in variable domains occur between the homologues of the C and D strands in constant domains and are labeled C' and C'' . Constant domains are thus made up of a three-stranded and a four-stranded β sheet comprising strands G-F-C and A-B-E-D, respectively, whereas variable domains are made up of a five- (or six)-stranded and a four- (or three)-stranded sheet comprising strands (A')-G-F-C- C' - C'' and (A)-B-E-D, respectively (Figure 1). In Ig variable domains, the A strand crosses from one sheet to the other to form a parallel arrangement with the G strand for all or a portion of its length, and the designation A' is used to indicate the region of the A strand in a parallel arrangement with the strand G.

As more Ig-like sequences were discovered, it became apparent that not all Ig-like sequences fit comfortably into the variable (V) or constant (C) domain subtypes. A characteristic of many of these domains is a smaller number of amino acids separating the conserved cysteines than was found even in constant domains. Williams proposed classifying these shorter sequences, along with other misfit sequences, as C2-set domains to distinguish them from sequences more closely related to those of antibody constant domains, which he termed C1-set domains (Williams 1987). The first atomic structure of a recognized C2-set Ig domain (the second domain of human CD4) revealed the expected strong structural homology to C1-set domains but with the unexpected sheet-switching of one of the β strands (Ryu et al 1990, Wang et al 1990). The shortened number of amino acids connecting the C and E strands resulted in the intervening strand (strand D in C1-set domains) remaining part of the G-F-C sheet and crossing to the other sheet only at the beginning of the E strand. To maintain a consistent nomenclature with V domains, the strand connecting the C and E strands in these C2-set domains is labeled the C' strand. These C2-set domains thus consist of a four-stranded and a three-stranded β sheet comprised of strands G-F-C- C' and A-B-E, respectively. This same strand organization is also observed in type III fibronectin repeat structures (Leahy et al 1992b, Main et al 1992). Not all members of the original sequences designated C2 will have this structure (Williams & Barclay 1988, Bork et al 1994), and the

C2 designation has evolved so that most authors now intend it to include only those domains with this structure.

The determination of additional β -sandwich structures, both in and out of the Ig superfamily, has revealed strand arrangements intermediate between the C2-, C1-, and V-set structures described above. As noted by Bork et al (1994), the structural type of Ig domains correlates well with the number of amino acids intervening between the C' and E strands. Domains with 7–10 residues in this connecting region have sheet-switched or C2-like structures, whereas C1-like structures typically have 20 or more intervening residues but too few residues to form complete C'' and D strands and a V-like structure. Bork et al (1994) also noted several hybrid domains with 12–13 residues in this intervening region that form a structure in between the C2- and C1-like structures. In these cases, the residues between strands C and E form part of a C' strand on one sheet but then cross over to the other sheet and finish as part of a D strand. These hybrid domains were termed H-type domains. Other authors have noted that the Ig-like domain of telokin, the C-terminal domain of myosin light chain kinase (Holden et al 1992), forms a structure intermediate between C1- and V-like domains (Harpaz & Chothia 1994). Although telokin is similar to C1-like domains in having a D strand and lacking C' and C'' strands (17 residues intervene between the C and E strands), the A-B and E-F connecting loops contain structural features conserved in V-like domains. The telokin domain has thus been recognized as a new type of Ig domain, termed the I or intermediate-type domain to indicate its intermediate status between C1- and V-like domains, and the sequences of several Ig-like domains in cell adhesion molecules suggest that these domains are likely to adopt I-like structures (Harpaz & Chothia 1994, Vaughn & Bjorkman 1996).

Ig domains consist of a β -sandwich structure with a common core and variable configurations of strands on one edge (Figure 1). In this scheme, a continuum of C2-H-C1-I-V-type structures exists that can be classified primarily by the increasing number of residues intervening between the C and E strands and the structure adopted by these residues. Sufficient sequence homology in the core structure exists between members of the C2-, C1-, I-, and V-type domains to establish a divergent evolutionary relationship between many of the domains in these classes. Domains with H-type structure and several domains that can be classified as C2-type domains do not show sequence homology to Ig domains, and the evolutionary relationship of these domains and Ig domains is less clear.

Interdomain interactions between Ig domains in multidomain proteins are important determinants of structure. Many Ig domain-containing proteins, such as immunoglobulins, T-cell receptors, and CD8, form intimately associated dimers with pseudo-twofold symmetry between pairs of Ig domains on different polypeptide chains. Other Ig-superfamily proteins, such as CD2, CD4,

the intercellular adhesion molecules (ICAMs), vascular adhesion molecules (VCAMs), and the muscle protein titin (Pfuhl & Pastore 1995), appear to consist of extended linear arrays of monomeric Ig domains as evidenced by electron micrographs (Staunton et al 1990) and high-resolution structures of multidomain fragments of CD2 (Jones et al 1992), CD4 (Ryu et al 1990, Wang et al 1990, Brady et al 1993) and VCAM (Jones et al 1995, Wang et al 1995). These extended arrays are facilitated by the emergence of the N and C termini from opposite ends of Ig domains (Figure 1). Ig dimer contacts are principally mediated by the F-G and C-C' loops, and these loops are often foreshortened in monomeric Ig domains (Ryu et al 1990, Wang et al 1990, Leahy et al 1992a).

Fibronectin Domains

Fibronectin (FN), a dimeric molecule of 220 to 250 kDa subunits, is a significant component of plasma and extracellular matrices and contains binding sites for collagen, heparin, fibrin, gelatin, and cell-surface receptors of the integrin family (Hynes 1990). When the primary sequence of fibronectin was determined, three types of homologous sequence repeats were discovered (Petersen et al 1983, Kornblihtt et al 1985). Fibronectin monomers are composed of 12 \approx 45 amino acid type I repeats (FN-I), 2 \approx 60 amino acid type II repeats (FN-II), and 15–17 \approx 90 amino acid type III repeats (FN-III). Sequences homologous to each of these repeats have been found in other proteins, with FN-III homologues among the most ubiquitous of protein domains, with homologues present in \approx 2% of animal sequences (Bork & Doolittle 1992). FN-III domains often occur in arrays of 10 or more repeats and have recently been discovered in yeast (Bateman & Chothia 1996). High-resolution structures have been determined for homologues of type I (Williams et al 1994), type II (Constantine et al 1992), and type III (DeVos et al 1992, Leahy et al 1992b, 1996, Main et al 1992, Dickinson et al 1994, Huber et al 1994) repeats, but this review considers primarily FN-III structures.

FN-III domains are topologically identical to the C2-set Ig domains described above; that is, a β -sandwich made up of a four-stranded and a three-stranded antiparallel β sheet consisting of strands G-F-C-C' and A-B-E, respectively (Figure 1). Although Ig domains and FN-III domains exhibit a high level of structural similarity (Leahy et al 1992b, Main et al 1992, Bork et al 1994), these domains exhibit distinctive hydrophobic core packing and possess different patterns of conserved residues. FN-III domains, for example, lack the highly conserved intersheet disulfide bond found in most Ig domains. FN-III domains do not appear to dimerize like some Ig domains, and the filament-like appearance of multidomain arrays of FN-III domains in electron micrographs of proteins such as fibronectin and tenascin suggests an extended structure for these arrays (Engel et al 1981, Erickson et al 1981). High-resolution structures

of a fragment of neuroglian containing two FN-III repeats and a fragment of fibronectin containing four such repeats reveal extended configurations and provide a glimpse of interdomain relationships and interactions between adjacent FN-III domains (Huber et al 1994, Leahy et al 1996).

Cadherin Domains

Cadherins are a large family of single-pass transmembrane proteins involved principally in calcium-dependent homotypic cell adhesion (Takeichi 1990, 1991). The prototypical cadherins consist of an extracellular region containing five tandem repeats of a characteristic ≈ 110 amino acid sequence, a transmembrane region, and a cytoplasmic domain that mediates interactions with the cytoskeleton. Cadherin-binding sites typically map to the N-terminal domain (Nose et al 1990). High-resolution structures have recently been determined for the N-terminal domain of N-cadherin (Shapiro et al 1995a), the N-terminal domain of E-cadherin (Overduin et al 1995, Overduin et al 1996), and the two N-terminal domains of E-cadherin (Nagar et al 1996). These structures reveal cadherin domains to be remarkably similar to Ig C1-set domains in overall fold and secondary-structure topology despite the absence of sequence homology (Figure 1). Notable differences between cadherin and Ig C1-set domains include the sheet-switching of the end of the A strand to form a parallel arrangement with the G strand and the existence of two turns of an unusual β -helix structure connecting strands C and D. The sheet-switching of the A strand is also seen in Ig V-type domains, so the cadherin structure may be considered to have the same topology as an Ig V-like domain with the C' and C'' strands replaced by the β -helix.

The structure of the two N-terminal domains of E-cadherin revealed three calcium ions bound at the interdomain interface (Nagar et al 1996). The residues involved in coordinating these ions are highly conserved in cadherin domains and indicate that not only is a similar arrangement of calcium ions bound at each of the interdomain interfaces in E-cadherin, but that such calcium binding is a shared feature of all members of the cadherin superfamily. The three bound calcium ions serve as a bridge between the two cadherin domains and appear to rigidify the relative orientation of these domains. A relatively small interface exists between the two domains (450 \AA^2 of buried surface area), and NMR linewidths indicate that the two domains tumble relatively freely in solution in the absence of calcium (Nagar et al 1996).

Divergent versus Convergent Evolutionary Relationships

Although it is increasingly clear that tertiary structure is more highly conserved than primary sequence, it is not always evident whether proteins with similar structures but unrelated amino acid sequences result from convergent or

divergent evolutionary paths. It could be argued that distinguishing between these alternatives represents primarily a semantic distinction, because all proteins can be considered as divergently related if one travels back far enough in evolutionary time. However, this view does not account for proteins (or protein domains) that arose through duplication and subsequent divergence of an existing fully formed protein. For the purposes of this discussion, a divergent evolutionary path is assumed to have resulted from duplication of a preexisting and fully formed domain.

The essentially identical folding topologies of FN-III, cadherin, and Ig superfamily domains coupled with their strikingly similar structures immediately stimulated the question of evolutionary relatedness of these domains. As yet, this question is not definitively answered—and may never be—but in the absence of convincing proof of a divergent relationship in the amino acid sequence relationships, most authors initially inclined toward a convergent relationship for these and other structurally similar domains (Richardson et al 1976, Leahy et al 1992b, Main et al 1992, Shapiro et al 1995a), although reasons for this choice seem mostly philosophical; other authors appear oppositely inclined (Overduin et al 1995, Wagner 1995). Distinctive hydrophobic core packing, lack of sequence similarity, and differing intron structures are cited in the most thorough study concluding a likely convergent relationship between these domains (Shapiro et al 1995b).

For other domain types, e.g. triose phosphate isomerase (TIM) barrels, it has been argued that conserved folding topology in the absence of amino acid sequence similarity represents a divergent evolutionary relationship (Reardon & Farber 1995). Similar enzymatic functions and an unvarying active site location are used to bolster this conclusion, but comparison of this system with the FN-III-Ig-cadherin comparisons highlights that perceived complexity of the folding topology influences notions about divergent versus convergent evolutionary paths. It is no doubt less difficult to believe that the relatively simple β -sandwich topology arose more frequently during evolution than the more complex TIM barrel. The observation of the Rossman-fold topology for integrin I domains (Figure 2) in the absence of sequence similarity to other proteins known to possess this fold is also a case in which the apparent complexity of the fold topology has led to the implicit assumption of a divergent evolutionary relationship (Lee et al 1995a, Qu & Leahy 1995). Without detailed knowledge of how new protein folds have been created, the issue of fold complexity may be more or less valid, and some authors have proposed independent origins for TIM barrels as well (Lesk et al 1989).

While a definitive answer to the evolutionary relatedness of Ig, cadherin, and FN-III domains is not yet possible, future studies could significantly influence this issue. Evidence that could establish a divergent relationship would be

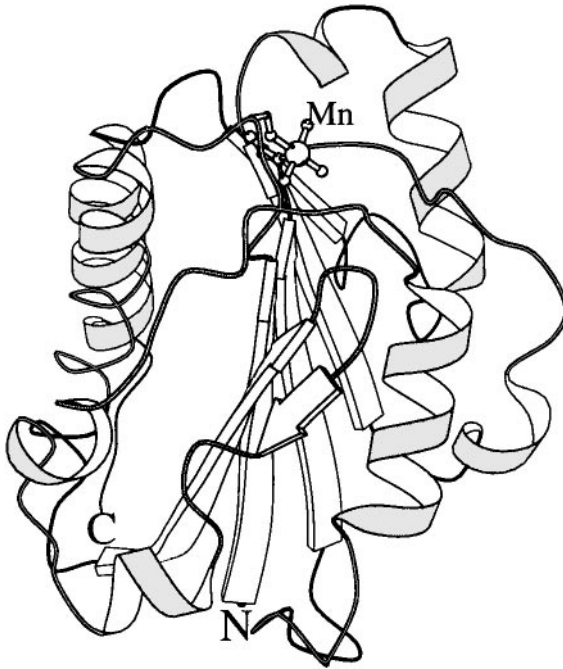


Figure 2 Secondary-structure diagram of the CD11a I domain with bound manganese ion (Qu & Leahy 1995). This figure was made with the program MOLSCRIPT (Kraulis 1991).

discovery of domains intermediate in amino acid sequence or domains that mix structural features previously characteristic of different domain subtypes. Also, phylogenetic or evolutionary clock analyses combined with whole genome sequences of anciently diverged organisms could pinpoint first occurrences of particular domain types and either undermine or strengthen the plausibility of a divergent relationship.

IMPLICATIONS OF ATOMIC-RESOLUTION STRUCTURES FOR CELL ADHESION

CD2

CD2, a cell-surface glycoprotein with two extracellular Ig domains, a membrane-spanning region, and a cytoplasmic region that is expressed primarily on T lymphocytes (T cells), plays a role in mediating interactions between T cells and target or accessory cells (Bierer et al 1989, Dustin & Springer 1991). Curiously, human CD2 interacts with CD58 (LFA-3), whereas rodent

CD2 interacts with CD48 (BCM1). CD2, CD48, and CD58 possess a similar domain organization, and their arrangement on the chromosome suggests that these molecules are derived from a common precursor (Wong et al 1990). What may once have been a homotypic cell adhesion interaction appears to have become heterotypic.

The structure of extracellular Ig domains of rat (Jones et al 1992) and human (Bodian et al 1994) CD2 (CD2D1D2) have been determined by X-ray crystallography, and the structure of the N-terminal Ig domain of CD2 (CD2D1) has also been determined by NMR for the rat (Driscoll et al 1991) and the human Igs (Withka et al 1993). A metastable dimer of rat CD2D1, in which CD2 monomers have interchanged β strands, has also been determined (Murray et al 1995). However, this form of CD2 appears to be an artifact of the expression system employed, and while providing an interesting example of how intimate dimers may arise during evolution (Bennett et al 1995), this dimeric structure does not appear directly relevant to CD2 biology. The CD2 structures reveal an N-terminal Ig V-like domain followed by an Ig C2-like domain. These domains are assembled in a rod-like fashion with a 6-amino acid linker between the domains that appears to be flexible. Interdomain flexibility is suggested by the relatively small amount of surface area buried between the two domains [$\approx 400 \text{ \AA}^2$ for CD2 versus $\approx 850 \text{ \AA}^2$ for the N-terminal two domains of CD4 (Ryu et al 1990, Wang et al 1990)] and a 7° flex in the interdomain relationship observed in independently determined structures of rat CD2D1D2 (Jones et al 1992).

An intriguing result obtained from the CD2 structures is the observation of virtually identical head-to-head packing between adjacent CD2D1 domains in the crystal lattices of both rat and human CD2. This similar interaction occurs despite considerable differences in the amino acid composition of these interfaces and of different crystallographic lattices. The independent occurrences of this interaction and the mapping of residues that are important for CD2-mediated cell adhesion to this interface suggest this arrangement may mimic the interaction of CD2 with its endogenous ligands. The C termini of the interacting CD2 molecules emerge from opposite ends of the lattice dimers, as would be expected for an interaction between molecules on opposite cell surfaces. Other dimers of Ig V-like domains (e.g. antibodies or CD8) involve the same surface employed by the CD2D1 lattice dimers, the GFCC'C'' sheet surface, but the relative orientation of the β sheets at the CD2 dimer interface differs by $\approx 60^\circ$ from orientation typically observed in immunoglobulins (Jones et al 1992).

Observation of the CD2 lattice dimer strengthens the notion that the heterotypic CD2-mediated interactions evolved from a homotypic antecedent and provides a basis for understanding and testing the molecular nature of these interactions. The CD2 dimer interface appears relatively flat and dominated by electrostatic interactions. These unusual features may be related to the dual

requirements for specificity and a high dissociation rate necessary to allow cells to eventually separate (Jones et al 1995, Wang et al 1995).

VCAM Domains

The major form of the cell-surface glycoprotein vascular cell adhesion molecule 1 (VCAM-1) consists of seven extracellular Ig domains, a single membrane-spanning region, and a cytoplasmic domain (Osborn et al 1989). VCAM-1 is inducibly expressed on endothelial cells and mediates interactions with integrins expressed by leukocytes during inflammatory and immune responses. Integrin-binding sites have been mapped to a conserved motif in domains one and four. Two groups recently reported very similar crystal structures of an integrin-binding fragment of VCAM-1 comprising the two N-terminal Ig-like domains (VCAMD1D2) (Jones et al 1995, Wang et al 1995). The N-terminal domain of VCAM-1 belongs to the I-set class of Ig domains (Harpaz & Chothia 1994), and the second domain belongs to the C2-set class of Ig domains. The overall structure of VCAM-1D1D2 is rod-like with 800–850 Å² of surface area buried between the domains. This value is comparable to the amount of surface area buried between domains one and two of CD4, whose interdomain orientation is thought to be rigid (Ryu et al 1994), but the tilt angle between the two VCAM domains differs by 12° in independent VCAM structures, indicating at least moderate flexibility at this junction (Wang et al 1995). Unlike CD2D1D2 (Jones et al 1992, Bodian et al 1994) and CD4D1D2 (Ryu et al 1990, Wang et al 1990), in which the adjacent domains are related by a near pseudo-twofold rotation axis ($\approx 160^\circ$), the rotation relating D1 and D2 of VCAM-1 is $\approx 130^\circ$.

Mutagenesis experiments have implicated two regions of the VCAM-1D1 sequence in integrin binding. The VCAM-1 structures reveal these two regions to be spatially contiguous on the C-D and E-F loops. The C-D loop protrudes from D1 and contains aspartic acid 40, a residue absolutely required for integrin binding. This exposed arrangement suggests that Asp 40 is likely to participate directly in integrin binding. The E-F loop is positioned somewhat behind the C-D loop, and it is unclear whether this region interacts directly with integrin or influences the conformation of the C-D loop, which is connected by several hydrogen bonds. Further evidence for a direct interaction of the C-D loop with the integrin receptor is provided by the ability of a cyclic peptide encompassing the C-D loop to inhibit the VCAM-1 integrin interaction (Wang et al 1995).

Peripheral Myelin Protein Zero

Peripheral myelin protein zero (P0) is the major structural component of peripheral nerve myelin (Lemke & Axel 1988). P0 consists of a single extracellular Ig domain, a membrane-spanning region, and a small cytoplasmic tail that is highly basic. Through homotypic adhesion of the extracellular domains, P0

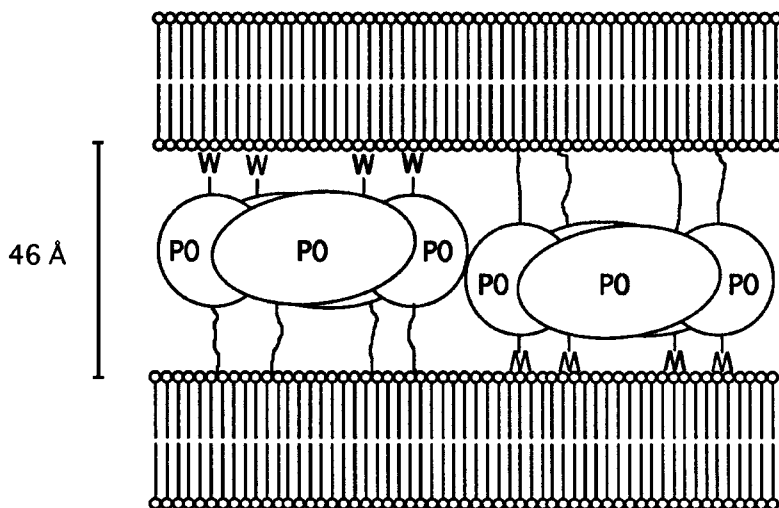


Figure 3 Schematic diagram of two tetramers of the P0 protein linking two membrane bilayers. On one side of the tetrameric disks, the conserved tryptophans emerge (labeled W), and on the other side of the tetramers, the C termini of the P0 protomers emerge. See Shapiro et al (1996) for further details.

mediates the adhesion of adjacent membranes in the spiral layers of the myelin sheath. Several mutations in P0 appear to be associated with a range of hereditary peripheral demyelinating neuropathies in humans (Warner et al 1996).

A high-resolution crystal structure of the extracellular Ig domain of rat P0 (P0ex) revealed the expected Ig V-like structure (Shapiro et al 1996). A unique structural feature of P0ex is an extended B-C loop that contains a solvent-exposed tryptophan at its apex. Also noteworthy is the manner in which P0ex protomers pack in the crystal lattice; the P0ex protomers assemble into disk-shaped tetramers with a fourfold symmetric axis running through the center of the disk perpendicular to the plane of the disk (Figure 3). The C termini of all four protomers emerge from one of the flat surfaces of the disk, whereas the exposed B-C loop tryptophans emerge from the opposite surface. These tetramers associate laterally in the crystals to form flat planes in which adjacent tetramers alternate between having their C termini (or exposed B-C loop tryptophans) emerge from one side of the sheet or the other (Figure 3).

The arrangement of P0ex into tetramers and the assembly of these tetramers into planar sheets immediately suggest that P0 in myelin membranes may be associated in a similar manner. Several lines of evidence support this model. First, P0ex crystals are grown in conditions of high protein concentration and

slightly alkaline pH, and the same conditions are required for maintenance of the native intermembrane spacing in peripheral nerve myelin (Inouye & Kirschner 1988). Second, ultracentrifugation studies confirm the formation of P0 tetramers in solution at conditions near physiological pH and ionic strength (and near the crystallization conditions). Finally, a model of adjacent myelin membranes constructed with the exposed B-C loop tryptophans from either side of the planar sheet of P0ex tetramers intercalating into apposed membranes results in an intermembrane spacing of $\approx 46 \text{ \AA}$ (Shapiro et al 1996). This value is in remarkable agreement with the value ($\approx 46 \text{ \AA}$) for the phosphate-to-phosphate distance found by X-ray diffraction of intact myelin (Kirschner et al 1989).

Thus the case for the interactions observed in the crystal lattice reflecting P0 interactions in peripheral nerve myelin is quite compelling and can form the basis for interpreting existing biological data, as well as the design of new experiments. P0 mutations found in demyelinating diseases do not map to specific hot spots and appear to result in a spectrum of phenotypes that complicate interpretation of the molecular basis of these diseases. Nonetheless, the lattice model of P0 interactions provides a starting point for evaluation and analysis of the likely molecular effects of disease-causing mutations (Warner et al 1996).

FN-III Domains

A primary interest of initial high-resolution structure determinations of FN-III domains was visualization of the integrin-binding RGD loop (Leahy et al 1992b, Main et al 1992, Dickinson et al 1994). In 1984, Pierschbacher & Ruoslahti reported the astonishing finding that peptides as small as six amino acids, derived from the tenth FN-III repeat of FN, could effectively block interactions of FN with its cell-surface receptor as long as the peptides included the sequence Arg-Gly-Asp (RGD) (Pierschbacher & Ruoslahti 1984). This result seemed paradoxical because these small peptides were unlikely to be adopting a native conformation, and these authors presciently concluded that the RGD sequence occurred on an extended loop of the FN protein (Pierschbacher & Ruoslahti 1984). The cell-surface receptor for FN has subsequently been identified and shown to be a member of the integrin family of cell-surface receptors, and recognition of RGD sequences in various ligands has proven a feature of many integrins (Hynes 1992). The involvement of RGD-mediated interactions in numerous cell adhesion and migration processes has stimulated great interest in the development of RGD peptides or RGD-peptide analogues as potential therapeutic agents modulating conditions such as thrombosis, inflammation, and cancer.

The first reported structure of a FN-III domain homologue was the human growth hormone receptor (HGHR) (DeVos et al 1992). Although the homology between cytokine receptors and FN-III domains had previously been noted (Bazan 1990, Patthy 1990), the implication of the HGHR structure for cell

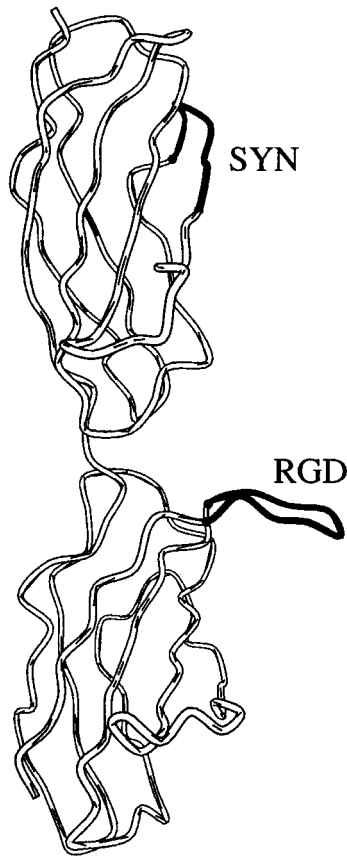


Figure 4 Diagram of an α -carbon trace of the ninth and tenth FN-III repeats of fibronectin (Leahy et al 1996). Shown in *black* are the synergy region on the ninth domain and the RGD loop on the tenth domain. This figure was made with the program MOLSCRIPT (Kraulis 1991).

adhesion was not widely appreciated. Subsequently, the structure of the tenth FN-III domain from FN (FNfn10) was determined by NMR (Main et al 1992), and the third FN-III repeat from tenascin (TNfn3) was determined by X-ray crystallography (Leahy et al 1992b). Both domains contain RGD sequences active in cell adhesion that occur in β -hairpin loops connecting strand F and G (Figure 4). The FNfn10 F-G loop has an insertion of two residues on either side of the RGD sequence relative to TNfn3 and most other FN-III domains, and these insertions extend the FNfn10 RGD loop an additional 5–7 Å away from the domain core. The RGD loop in the NMR structure and a subsequent

high-resolution crystal structure of FNfn10 (Dickinson et al 1994) both exhibited considerable disorder, consistent with the notion of a flexible and exposed RGD loop. The RGD sequence in TNfn3 occurred in a type II' β -hairpin turn, and structures of other RGD sequences active in cell adhesion invariably reveal either a type II'-like β turn or a poorly ordered loop (Adler et al 1991, Saudek et al 1991, Krezel et al 1994, Yamada et al 1995).

The RGD sequence is not the sole contributor to the interaction between FN and its integrin receptor. Mutations in a region in the C'-E loop of FNfn9 also interfered with FN-integrin interactions, and this region has become known as the synergy region (Aota et al 1991, 1994, Bowditch et al 1991, 1994). The interdomain orientation of FNfn9 and FNfn10 (i.e. the relative positions of the RGD loop and the synergy region) is clearly an important feature of the integrin-binding site on FN. Based principally on the interdomain relationship between the first and second Ig domains of CD4 (CD4D1 and CD4D2) (Ryu et al 1990, Wang et al 1990) and some rudimentary modeling, the contiguous arrays of FN-III domains were thought likely to adopt a zigzag configuration with adjacent domains related by an approximate pseudo-twofold axis (Leahy et al 1992b, Bowditch et al 1994). The structure of the first and second FN-III domains of *Drosophila* neuroglian (NGfn1 and NGfn2) revealed just such a pseudo-twofold relationship (Huber et al 1994). In addition, a Na⁺ ion was bound at the interdomain interface that must orient and rigidify the interdomain interface. NGfn1 and NGfn2 possess a functionality distinct from the remaining three FN-III domains in neuroglian (promotion of neurite outgrowth versus adherence of small cerebellar neurons), and the residues that coordinate the Na⁺ ion are not conserved at the three other interfaces between FN-III domains in NG. A vertebrate homologue of neuroglian, L1, is also likely to have a Na⁺ ion between the first and second FN-III domains. The presence of the Na⁺ ion may thus be important for the biological function of the first two FN-III domains of neuroglian and L1 (Huber et al 1994).

The crystal structure of a fragment of FN (FN-III domains seven through ten) revealed highly variable interdomain relationships (Leahy et al 1996). FNfn8 and FNfn9 were related by a near-pseudo-twofold relationship; FNfn7 and FNfn8 were related by a 112° rotation, whereas FNfn9 and FNfn10 were related by only a very small (43°) rotation. The atypically small rotation relating FNfn9 and FNfn10 places the RGD loop and the synergy region \approx 30–40 Å apart on the same surface of the FN molecule (Figure 4). The dimensions and layout of the binding site defined by these two regions could easily be spanned by a single integrin molecule. Most of the difference in rotation angles between the ninth and tenth FN-III domains of FN and the other FN-III interfaces can be attributed to a variant main-chain dihedral angle following Asp 1418 at the FNfn9/FNfn10 interface. This unusual Ψ angle results in the side chain of the

preceding residue, Ser 1417, becoming exposed to solvent. The amino acid at the position homologous to Ser 1417 in all other FN-III domains in FN (except FNfn1, which has no interface with an FN-III domain at its N terminus) is hydrophobic, and the residue at this position is buried in the FNfn7/FNfn8 and FNfn8/FNfn9 interfaces. The unusual interdomain relationship between FNfn9 and FNfn10 may thus be strongly influenced by this single amino acid side chain, and mutation of Ser 1417 to a hydrophobic residue may be an interesting way to alter the orientation between FNfn9 and FNfn10 and examine the functional consequences of such an rearrangement.

The proximity of the synergy region and the RGD loop in the FN7–10 structure suggests that the interdomain orientation between FNfn9 and FNfn10 observed in the crystal structure is similar to what is recognized by the FN receptor, but the small surface area buried between these domains also indicates that this interface is likely to be flexible. FN is known to associate in higher-order fibrillar complexes, and this association is promoted by association with integrins. An intriguing hypothesis is that the interactions between FN and the integrin stabilize the FNfn9/FNfn10 interface and that this stabilized structure is needed for fibril formation.

Cadherin Domains

The cadherin domain structures provide insight into two fundamental questions concerning the nature of cadherin function—what is the nature of homotypic cadherin interactions, and what is the role of calcium in cadherin function? The EcadD1D2 structure revealed three calcium ions bound primarily at the interface between the two cadherin domains (Nagar et al 1996). One calcium ion is coordinated primarily by EcadD1, another by EcadD2, and the third by residues from both domains. This third calcium ion establishes a network of calcium-mediated links between EcadD1 and EcadD2 through interactions with amino acid side chains that also coordinate each of the other two calcium ions. These calcium ions appear to rigidify and orient the interdomain interface between EcadD1 and EcadD2. This role for calcium had been suggested by electron micrographs of E-cadherin that reveal more rigid structures in the presence of Ca^{2+} (Pokutta et al 1994). A conserved histidine-alanine-valine (HAV) sequence in the F strand of cadherins, which has been implicated in cadherin adhesion interactions, is located on the opposite end of EcadD1 more than 30 Å away from the calcium-binding sites, suggesting that the calcium ions are unlikely to directly mediate cadherin adhesive interactions. The residues that coordinate all three calcium ions in EcadD1D2 are conserved at most cadherin domain interfaces, and the requirement of Ca^{2+} for cadherin-mediated adhesion thus appears to result from induced rigidification of the cadherin rather than from direct participation of the calcium ion at an interaction interface.

Crystal lattice interactions that may reflect physiological interactions of cadherins are observed in crystals of both NcadD1 and EcadD1D2. In each of three different crystal forms of NcadD1, the same two types of interactions between NcadD1 protomers occur (Shapiro et al 1995a). The first type of interaction, called the strand dimer, involves the reciprocal intercalation of the beginning of strand A of one molecule, including a highly conserved tryptophan, Trp 2, into the hydrophobic core of an adjacent molecule. This strand dimer buries 1800 Å² of surface area and results in an NcadD1 dimer pair with the C termini of both monomers emerging from the same side of the dimer. For this reason, the strand dimer may also be thought of as a lateral or parallel dimer. The second lattice interaction observed in the NcadD1 crystals involves the association of strand dimers into long arrays such that the C termini of adjacent strand dimers emerge from opposite sides of the array (Figure 5). This interaction has been termed the adhesion dimer, as it could represent interactions between cadherins from different cells. This adhesion dimer interface involves the F-strand HAV sequence but not the calcium-binding region of NcadD1. The pattern of conservation of the amino acids at the NcadD1 dimer interfaces is supportive of both dimer interactions being conserved in other cadherins.

A suggestive dimer interaction is also present between adjacent EcadD1D2 molecules in the crystal lattice, but this interaction differs substantially from the NcadD1 interactions (Nagar et al 1996). Like the NcadD1 strand dimer interaction, the EcadD1D2 dimer interaction is parallel, with the C termini of each monomer exiting from the same end of the dimer. The interaction between the EcadD1 domains in this dimer is much less extensive than the interaction between NcadD1 domains in the strand dimer (1000 versus 1800 Å² buried surface area), and the relative orientation of the monomers differs by $\approx 56^\circ$ in the two types of dimers. The EcadD1D2 dimer interface also involves regions associated with the calcium-binding sites, possibly explaining the requirement for calcium for EcadD1D2 dimers to be observed in solution. By contrast, NcadD1 forms dimers in solution even in the absence of calcium (Shapiro et al 1995a, Nagar et al 1996).

Participation of the calcium-binding site region in the dimer interface as observed for the EcadD1D2 dimer would not be possible for the strand dimer observed in the NcadD1 crystals. The C termini of the NcadD1 subunits, which are involved in coordinating calcium, are simply too far apart in the strand dimer. If cadherins employ a conserved lateral interaction, either one or both of the interactions observed for NcadD1 or EcadD1D2 are not correct. Subsequent to these structures, independent experiments have provided evidence that a lateral dimer (i.e. a dimer of similarly oriented cadherins emanating from the same surface) must form prior to the formation of adhesive or antiparallel dimers (Briehner et al 1996, Tomschy et al 1996). It is not yet possible to determine

whether one, both, or none of the lateral dimer models accurately represents physiological cadherin interactions. These models do, however, make specific predictions that can be assessed by mutagenesis and additional structural experiments, and this issue will no doubt be resolved shortly.

Integrin I Domains

Integrins are a large family of heterodimeric cell-surface receptors that mediate a wide variety of cell-cell and cell-matrix interactions (Hynes 1992). Integrin α and β chains are noncovalently linked, single-pass transmembrane proteins of 150–180 and 90–110 kDa molecular mass, respectively, with the majority of their polypeptide chains on the extracellular surface of the cell. Integrin ligands include extracellular matrix proteins such as fibronectin, laminin, collagen, and fibrinogen, as well as cell-surface molecules such as intercellular adhesion molecules (ICAMs) and vascular cell adhesion molecules (VCAMs). Many, if not all, integrins possess the ability to signal in both directions across the cell membrane. Cells not only respond to interaction of integrins with ligand but also appear to up-regulate the affinity of cell-surface integrins for ligands in response to cellular activation (Diamond & Springer 1994, Schwartz et al 1995). A subset of integrin α chains contain a ≈ 200 amino acid inserted I domain that is homologous to the von Willebrand Factor A domains (Colombatti & Bonaldo 1991). I domains contain a conserved divalent cation-binding site and invariably possess a major component of ligand-binding activity of I domain-containing integrins (Diamond et al 1993, Michishita et al 1993, Bilsland et al 1994, Kamata & Takada 1994, Kern et al 1994, Landis et al 1994, Randi & Hogg 1994).

Crystal structures have now been reported for the CD11b/CD18 ($\alpha_M \beta_2$, Mac-1) I domain in the presence of magnesium (Lee et al 1995a) and manganese (Lee et al 1995b), as well as the CD11a/CD18 ($\alpha_L \beta_2$, LFA-1) I domain in the presence of manganese (Qu & Leahy 1995), magnesium (Qu & Leahy 1996), and in the absence of divalent cations (Qu & Leahy 1996). These structures reveal a familiar fold that has a core parallel β sheet surrounded by α helices and is termed the nucleotide-binding or Rossman fold (Rossman et al 1974) (Figure 1). Unlike Ig-like domains, the N and C termini of I domains emerge from the same surface, and thus this domain can be inserted into an existing structure with minimal disruption. The bound cation is coordinated directly by the side chains of three residues and makes water-mediated contacts with the side chains of two other residues that include an aspartic acid and two serines in a conserved DXSXS motif, a threonine, and another aspartic acid. These residues have been collectively referred to as the MIDAS motif (Lee et al 1995a). The metal coordination geometry of this site is octahedral, which leaves one metal coordination site unligated by protein. In the structure of the CD11b I domain with bound Mg^{2+} , a glutamic acid side chain from an adjacent molecule in the

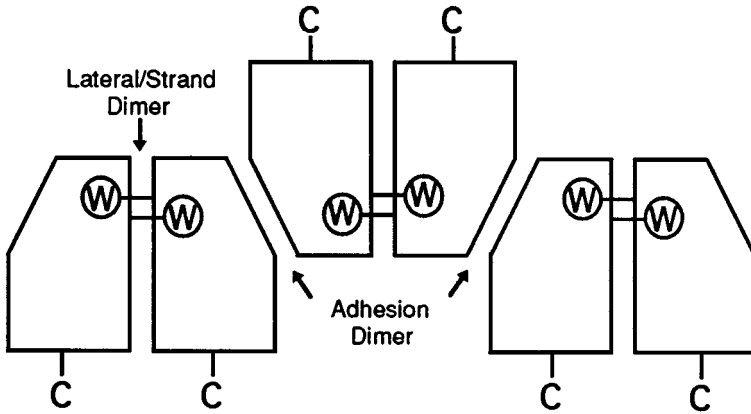


Figure 5 Schematic diagram showing the two types of dimers observed in NcadD1 crystals (Shapiro et al 1995a). The lateral/strand dimer involves intercalation of part of strand A, including a conserved tryptophan residue into the hydrophobic core of an adjacent molecule and these tryptophans are indicated by W. The adhesion dimer interface links up lateral dimers such that the C termini of alternating lateral dimers emerge from opposite ends of the assembly. A different lateral dimer interface and no adhesion dimer interface were observed in crystals of EcadD1D2 (Nagar et al 1996).

crystal lattice occupies this sixth metal coordination site, and the nature of the coordination of Mg^{2+} by the protein differs subtly in this structure from all other structures of I domains with a bound metal ion. In this case, the threonine directly coordinates the metal, whereas the aspartic acid outside the DXSXS motif hydrogen bonds to a metal-bound water molecule. In all other structures this coordination mode is reversed, with the threonine interacting indirectly and the aspartic acid directly with the bound metal ion. The structural differences observed in the CD11b I domain with bound Mg^{2+} are not a consequence of the identity of the bound cation because structures of the CD11a I domain with either manganese or magnesium ion bound show no significant differences (Lee et al 1995b, Qu & Leahy 1996), suggesting that it is the presence of the external glutamic acid side chain that is responsible for the altered metal coordination.

The coordination of the Mg^{2+} in crystals of CD11b I(Mg) by a glutamic acid side chain from an adjacent molecule in the crystal lattice led to the suggestion that a metal-carboxylate side chain interaction reflects physiological interactions between I domains and ligands (Lee et al 1995a). Several additional lines of evidence support this contention. Metals are required for ligand binding by integrins, and a conserved acidic amino acid (e.g. the Asp in RGD or Asp 40 in the C-D loop of VCAM-1) is an almost universal feature of integrin-binding sites

in integrin ligands (Smith & Cheresch 1991, Hynes 1992). Mutations found in the CD11a I domain that interfere with ligand binding map to the metal-binding surface (Huang & Springer 1995), and the observation of no significant differences between I domain structures with or without bound cation indicate that the cation is not required for induction of a distant binding site on the I domain (Qu & Leahy 1996).

The CD11b I domain structure with bound Mg^{2+} possesses other unique structural features when compared with all other I domain structures. Most notably, the $\alpha 7$ helix and preceding loop and strand shift significantly in position relative to the position of these regions in other I domain structures. This shift causes two phenylalanine residues (Phe 275 and 302) to become exposed to solvent in the Mg^{2+} -bound CD11b I domain structure, which is great enough to affect any interactions made by the surface of the I domain. The structural changes in the metal-binding site appear correlated with the changes in the $\alpha 7$ region and thus with the presence in the metal-coordination sphere of the side chain of the glutamic acid from an adjacent molecule in the crystal. The shift in the position of the $\alpha 7$ helix and preceding residues observed in the CD11b I(Mg) structure relative to the CD11b I(Mn) structure (and all CD11a I structures) has been proposed as a structural basis for affinity modulation of I domain-containing integrins (Lee et al 1995b). At the moment, there is little evidence beyond the existence of the distinct CD11b I domain structures to judge the merits of this intriguing conjecture, but it is unlikely that lability of the I domain structure is random.

The structure of the CD11a I domain determined in the absence of bound cation also does not differ significantly from the structure of this domain with bound manganese or magnesium ion (Qu & Leahy 1996). The absence of bound cation results in a deeper and more negatively charged cavity in the metal-binding site region, but shifts in protein atoms relative to metal-bound structures are remarkably slight in most regions of the protein including the metal-binding site. Large shifts (up to 12 Å) in the structure of the metal-free CD11a I domain structure relative to the metal-bound structures do occur in the C-terminal region of the $\alpha 7$ helix. These shifts do not involve the N-terminal turns of the $\alpha 7$ helix or preceding loops or strands and occur in a direction different than the shift observed in the $\alpha 7$ helix of the CD11b I domain. A reciprocal interaction between adjacent molecules in the crystal lattices of the metal-bound forms of the CD11a I domain in which the C termini intercalate between the $\alpha 7$ helix and the body of the I domain appears to account for the shifts in this region (Qu & Leahy 1996). The shift in the $\alpha 7$ helix confirms the lability of this region seen in the CD11b I domain structures and possibly indicates that at least portions of this helix are stabilized or influenced in intact integrins by interactions with other regions of the integrin.

The discovery of an I domain-like structure in all integrin β chains raises the possibility that features of interactions between I domains and ligands may be present in all integrins. Utilizing the conserved DXSXS motifs found in both I domains and integrin β chains to align their sequences, it was discovered that a hydrophathy plot of the region of integrin β chains aligned with the I domain strongly implied an I domain-like structure in integrin β chains (Lee et al 1995a). The buried parallel β strands and the amphipathic α helices in the I domain structure provide a characteristic signature in a hydrophathy plot that allows a high level of confidence in this alignment. In addition, this alignment places a conserved threonine and aspartic acid downstream of the DXSXS sequence in proper positions to form an I domain-like metal-binding site, which indicates that an I domain-like divalent cation-binding site also exists in all integrin β chains. Both mutagenesis and ligand cross-linking experiments implicate this I domain-like region of integrin β chains in ligand binding and are thus consistent with the possibility of similar ligand-binding mechanisms being employed by this region of integrin β chains and I domains (D'Souza et al 1988, Smith & Cheresch 1988, Takada et al 1992, Bajt & Loftus 1994).

The homology of I domains to A domains from von Willebrand Factor (vWF) allowed mapping of several disease-causing mutants on I domain-based models of vWF A domain structure. Type IIa and IIb von Willebrand Disease (vWD) result from mutations in the first A domain of vWF and exhibit different biochemical defects, as well as variable degrees of disease severity (Ginsburg & Sadler 1993). Alignment of the I domain and vWF A domain amino acid sequences allowed identification of the sites and probable molecular consequences of these vWD-causing mutations. Correlation of the nature of the mutations with the known molecular defects in both type IIa and IIb vWD led to hypotheses concerning the mechanisms underlying these forms of vWD (Qu & Leahy 1995).

SUMMARY AND PERSPECTIVE

The rapid accumulation of high-resolution structures of biomolecules has contributed to a general framework for thinking about proteins that has touched virtually all fields of biology, including cell adhesion. Sequence-structure relationships are much better understood, which allows even remotely related amino acid sequences to be placed in the same structural class. This knowledge can provide information about the protein as to overall dimensions, what regions distant in the primary sequence are likely to be adjacent in the tertiary structure, and either rule out or suggest possible molecular bases for biological function. The modeling of multidomain proteins based on subunit structure, as well as the mapping of vWD disease-causing mutations onto the homologous

integrin I domain structures, provides examples of direct benefit from cell adhesion processes (Huber et al 1994, Qu & Leahy 1995, Leahy et al 1996, Nagar et al 1996). Furthermore, the accumulation of multiple structures from a class of homologous domains, such as the Ig superfamily, provides a framework for amino acid sequence alignment and subclassification of these domains (Bork et al 1994, Harpaz & Chothia 1994, Vaughn & Bjorkman 1996). Examination of the distribution and function of more highly related IgSF subtypes will no doubt illuminate possible evolutionary histories of these domains. A less well-appreciated implication of the large database of protein structures is the increasing frequency with which amino acid sequences can be identified as homologous to proteins of known structure despite the absence of detectable sequence homology. A seven-bladed β -propeller structure was recently proposed for the N-terminal region of integrin α chains based on arguments of compatibility of the integrin α -chain sequence with this known structure at several levels (Springer 1997). This structure prediction seems likely to be correct in principle and immediately aids interpretation and design of experiments examining integrin function. Such predictions will be made with greater frequency and confidence as the number of existing protein folds for which no structure has been determined shrinks and the ability to correlate sequences with known structures improves.

Perhaps the most direct way in which high-resolution structure determinations have advanced knowledge of cell adhesion processes has been the identification and visualization of binding sites. Identification or characterization of a ligand-interacting surface has been an issue for virtually all the structures discussed in this review. In many cases, such as RGD sequences, the HAV sequence in cadherins, or Asp 40 in VCAM-1, a component of the binding site had been identified prior to the structure determination, and it is the structure and environment of this component that is of interest. In other cases such as the integrin I domains, the ligand-binding site is less certain, and the structure can serve as a basis for selecting appropriate residues for mutagenesis designed to map out the ligand-binding surface. In addition to ligand-binding sites, visualization of metal-binding sites has also proven valuable. Knowledge of the residues involved in coordinating a bound metal allows identification of homologues that also contain a metal-binding site, and the environment of the bound metal ion can provide clues to its function. The calcium ions in cadherins and the sodium ion in neuroglian certainly serve to rigidify these molecules and seem unlikely to participate directly in ligand binding, while the divalent cation in integrin I domains seems likely to interact directly with ligand.

Molecular characterization of cell adhesion-binding sites also produces practical benefits by aiding drug design. Cell adhesion molecules are potential targets for drugs modulating such processes as inflammation, thrombosis, viral

infection, and cancer. Drug design strategies typically employ one of two approaches. In the first, molecules are designed to mimic the salient chemical features of regions of a protein involved in ligand or receptor binding and thus competitively inhibit the physiological interaction. Drug design efforts utilizing this strategy include synthesis of RGD loop mimics to inhibit thrombosis, VCAM C-D loop mimics to inhibit inflammatory responses, and CD4 HIV-binding loop mimics as possible treatments for AIDS. A second drug design strategy involves the identification of pockets or cavities on the protein surface at binding interfaces and the design of small molecules to occupy these pockets (Kuntz 1992). A small molecule that binds to CD4 and disrupts binding of HIV to CD4 was recently produced by this approach (Li et al 1997).

Implications of crystal structures often transcend the information from the molecular structure alone. Interactions between neighboring molecules in a crystal lattice may reflect biological interactions, a fact that is especially likely in the case of homotypic cell adhesion molecules. In the case of P0, correlation with other lines of evidence provides a compelling case for an observed lattice interaction reflecting a physiological interaction, whereas in the case of the cadherin structures, conflicting results cloud the issue. The observed lattice interactions in the cadherin structures do, however, allow specific predictions that can be tested by experiment, and the nature of homotypic cadherin interactions is likely to be resolved in the near future. An interesting twist on the information contained in lattice interactions is provided by the interaction observed between symmetry-related CD2 molecules in several different crystal lattices. CD2 does not normally self-associate, but CD2 target ligands are homologous to CD2, and CD2 interactions are thought to arise from duplication and divergence of an ancestral homotypic adhesion molecule. The presence of an adhesion-like interaction between CD2 molecules not only provides a model testable by experiment but also appears to support the notion that CD2 is derived from a precursor involved in homotypic adhesion.

As noted above, the odds are currently strongly biased that the determination of the structure of a protein with no known sequence homologies will result in the discovery of an unexpected structural homology rather than a new protein fold (Murzin 1996, Holm & Sander 1997). This fact has been demonstrated for the Ig, FN-III, and cadherin folds, as well as for I domain and Rossman folds. In some instances, a divergent evolutionary relationship is obvious, but in cases such as the Ig, cadherin, and FN-III domains, the evolutionary relationship is less clear. Did these similar protein folds arise from a common precursor or does the β -sandwich fold have special properties that cause it to have been independently developed for what are ultimately similar purposes? Although high-resolution structures often do not readily answer this question, structural results are certainly responsible for the question being posed. Whether the

problem of the origins of these domains is intractable or awaiting either a new piece of information or a savant who can compose a compelling argument from existing information remains to be seen.

What can we expect in the future? The structures of many more cell adhesion domains and multidomain arrays will certainly be solved and further refine our understanding of these proteins. Co-crystals of ligands and receptors will also become more commonplace and allow direct visualization of cell adhesion interactions. A more distant goal, perhaps the holy grail of structural studies of cell adhesion, would be to provide a structural basis for understanding affinity modulation in integrins. These large molecules ultimately may be too flexible to yield diffraction-quality crystals, but characterization of ligand-binding fragments may prove sufficient to provide some answers.

ACKNOWLEDGMENTS

I gratefully acknowledge support from the American Cancer Society, the Searle Scholars Program, and the National Institutes of Health.

Visit the *Annual Reviews* home page at
<http://www.annurev.org>.

Literature Cited

- Adler M, Lazarus RA, Dennis MS, Wagner G. 1991. Solution structure of kistrin, a potent platelet aggregation inhibitor and GP IIb-IIIa antagonist. *Science* 253:445-48
- Amzel LM, Poljak RJ. 1979. Three-dimensional structure of immunoglobulins. *Annu. Rev. Biochem.* 48:961-97
- Aota S, Nagai T, Yamada KM. 1991. 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.* 266:15938-43
- Aota S, Nomizu M, Yamada KM. 1994. The short amino acid sequence Pro-His-Ser-Arg-Asn in human fibronectin enhances cell-adhesive function. *J. Biol. Chem.* 269:24756-61
- Bajt ML, Loftus JC. 1994. Mutation of a ligand binding domain of $\beta 3$ integrin: integral role of oxygenated residues in $\alpha \text{IIb}\beta 3$ (GPIIb-IIIa) receptor function. *J. Biol. Chem.* 269:20913-19
- Bateman A, Chothia C. 1996. Fibronectin type III domains in yeast detected by a hidden Markov model. *Curr. Biol.* 6:1544-47
- Bazan JF. 1990. Structural design and molecular evolution of a cytokine receptor superfamily. *Proc. Natl. Acad. Sci. USA* 87:6934-38
- Bennett MJ, Schlunegger MP, Eisenberg D. 1995. 3D domain swapping: a mechanism for oligomer assembly. *Protein Sci.* 4:2455-68
- Bentley GA, Boulot G, Karjalainen K, Mariuzza RA. 1995. Crystal structure of the β chain of a T cell antigen receptor. *Science* 267:1984-87
- Bernstein FC, Koetzle TF, Williams GJ, Meyer EE, Brice MD, et al. 1978. The protein data bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* 112:535-42
- Bierer BE, Sleckman BP, Ratnofsky SE, Burakoff SJ. 1989. The biological roles of CD2, CD4, and CD8 in T-cell activation. *Annu. Rev. Immunol.* 7:579-99
- Bisland CAG, Diamond MS, Springer TA. 1994. The leukocyte integrin p150,95 (CD11c/CD18) as a receptor for iC3b. *J. Immunol.* 152:4582-89
- Bjorkman PJ, Saper MA, Samraoui B, Bennett WS, Strominger JL, Wiley DC. 1987. Structure of the human class I histocompatibility antigen, HLA-A2. *Nature* 329:506-12
- Bodian DL, Jones EY, Harlos K, Stuart DI, Davis SJ. 1994. Crystal structure of the

- extracellular region of human cell adhesion molecule CD2 at 2.5 Å resolution. *Structure* 2:755–66
- Bork P, Doolittle R. 1992. Proposed acquisition of an animal protein domain by bacteria. *Proc. Natl. Acad. Sci. USA* 89:8990–94
- Bork P, Downing AK, Kieffer B, Campbell ID. 1996. Structure and distribution of modules in extracellular proteins. *Q. Rev. Biophys.* 29:119–67
- Bork P, Holm L, Sander C. 1994. The immunoglobulin fold: structural classification, sequence patterns and common core. *J. Mol. Biol.* 242:309–20
- Bowditch RD, Halloran CE, Aota S, Obara M, Plow EF, et al. 1991. Integrin $\alpha_{IIb}\beta_{IIIa}$ (platelet GPIIb-IIIa) recognizes multiple sites in fibronectin. *J. Biol. Chem.* 266:23323–28
- Bowditch RD, Hariharan M, Tominna EF, Smith JW, Yamada KM, et al. 1994. Identification of a novel integrin binding site in fibronectin. Differential utilization by $\beta 3$ integrins. *J. Biol. Chem.* 269:10856–63
- Brady RL, Dodson EJ, Dodson GG, Lange G, Davis SJ, et al. 1993. Crystal structure of domains 3 and 4 of rat CD4: relation to the NH_2 -terminal domains. *Science* 260:979–83
- Briher WM, Yap AS, Gumbiner BM. 1996. Lateral dimerization is required for the homophilic binding activity of C-cadherin. *J. Cell Biol.* 135:487–96
- Brunger AT, Leahy DJ, Hynes TR, Fox RO. 1991. The 2.9 Å resolution structure of an anti-dinitrophenyl-spin-label monoclonal antibody Fab fragment with bound hapten. *J. Mol. Biol.* 221:257–70
- Colombatti A, Bonaldo P. 1991. The superfamily of proteins with von Willebrand factor type A-like domains: one theme common to components of extracellular matrix, hemostasis, cellular adhesion, and defense mechanisms. *Blood* 77:2305–25
- Constantine KL, Madrid M, Banyai L, Trexler M, Patthy L, Llinas M. 1992. Refined solution structure and ligand-binding properties of PDC-109 domain b. A collagen-binding type II domain. *J. Mol. Biol.* 223:281–98
- DeVos AM, Ultsch M, Kossiakoff AA. 1992. Human growth hormone and extracellular domain of its receptor: crystal structure of the complex. *Science* 255:306–12
- Diamond MS, Garcia-Aguilar J, Bickford JK, Corbi AL, Springer TA. 1993. The I-domain is a major recognition site on the leukocyte integrin Mac-1 (CD11b/CD18) for four distinct ligands. *J. Cell Biol.* 120:1031–43
- Diamond MS, Springer TA. 1994. The dynamic regulation of integrin adhesiveness. *Curr. Biol.* 4:506–17
- Dickinson CD, Veerapandian B, Dai XP, Hamlin RC, Xuong NH, et al. 1994. Crystal structure of the tenth type III cell adhesion module of human fibronectin. *J. Mol. Biol.* 236:1079–92
- Driscoll PC, Cyster JG, Campbell ID, Williams AF. 1991. Structure of domain 1 of rat T lymphocyte CD2 antigen. *Nature* 353:762–65
- D'Souza SE, Ginsberg MH, Burke TA, Lam SC-T, Plow EF. 1988. Localization of an Arg-Gly-Asp recognition site within an integrin adhesion receptor. *Science* 242:91–93
- Dustin ML, Springer TA. 1991. Role of lymphocyte adhesion receptors in transient interactions and cell locomotion. *Annu. Rev. Immunol.* 9:27–66
- Engel J, Odermatt E, Engel A, Madri JA, Furthmayr H, et al. 1981. Shapes, domain organizations and flexibility of laminin and fibronectin, two multifunctional proteins of the extracellular matrix. *J. Mol. Biol.* 150:97–120
- Erickson HP, Carrell N, McDonagh J. 1981. Fibronectin molecule visualized in electron microscopy: a long, thin flexible strand. *J. Cell Biol.* 91:673–78
- Fields BA, Ober B, Malchiodi EL, Levedeva MI, Braden BC, et al. 1995. Crystal structure of the $V\alpha$ domain of a T cell antigen receptor. *Science* 270:1821–24
- Garboczi DN, Ghosh P, Utz U, Fan QR, Bid-dison WE, Wiley DC. 1996. Structure of the complex between human T-cell receptor, viral peptide and HLA-A2. *Nature* 384:134–41
- Garcia KC, Degano M, Stanfield RL, Brunmark A, Jackson MR, et al. 1996. An $\alpha\beta$ T cell receptor structure at 2.5 Å and its orientation in the TCR-MHC complex. *Science* 274:209–19
- Ginsburg D, Sadler JE. 1993. von Willebrand disease: a database of point mutations, insertions, and deletions. *Thromb. Hemost.* 69:177–84
- Harpaz Y, Chothia C. 1994. Many of the immunoglobulin superfamily domains in cell adhesion molecules and surface receptors belong to a new structural set which is close to that containing variable domains. *J. Mol. Biol.* 238:528–39
- Hill RL, Delaney R, Fellows RE, Lebovitz HE. 1966. The evolutionary origin of the immunoglobulins. *Proc. Natl. Acad. Sci. USA* 56:1762–69
- Holden HM, Ito M, Hartshorne DJ, Rayment I. 1992. X-ray structure of telokin, the C-terminal domain of myosin light chain kinase, at 2.8 Å resolution. *J. Mol. Biol.* 227:840–51
- Holm L, Sander C. 1997. New structure—novel fold? *Structure* 5:165–71
- Holmgren A, Branden C-I. 1989. Crystal structure of chaperone protein PapD reveals an immunoglobulin fold. *Nature* 342:248–51

- Huang C, Springer TA. 1995. A binding interface on the I domain of lymphocyte function-associated antigen-1 (LFA-1) required for specific interaction with intercellular adhesion molecule 1 (ICAM-1). *J. Biol. Chem.* 270:19008–16
- Huber AH, Wang YM, Bieber AJ, Bjorkman PJ. 1994. Crystal structure of tandem type III fibronectin domains from *Drosophila* neuroglian at 2.0 Å. *Neuron* 12:717–31
- Hynes RO. 1990. *Fibronectins*. New York: Springer-Verlag
- Hynes RO. 1992. Integrins: versatility, modulation, and signaling in cell adhesion. *Cell* 69:11–25
- Inouye H, Kirschner DA. 1988. Membrane interaction in nerve myelin. I. Determination of surface charge from effects of pH and ionic strength. *Biophys. J.* 53:235–46
- Jones EY, Davis SJ, Williams AF, Harlos K, Stuart DI. 1992. Crystal structure at 2.8 Å resolution of a soluble form of the cell adhesion molecule CD2. *Nature* 360:232–39
- Jones EY, Harlos K, Bottomley MJ, Robinson RC, Driscoll PC, et al. 1995. Crystal structure of an integrin-binding fragment of vascular cell adhesion molecule-1 at 1.8 Å resolution. *Nature* 373:539–44
- Kamata T, Takada Y. 1994. Direct binding of collagen to the I domain of integrin $\alpha 2\beta 1$ (VLA-2, CD49b/CD29) in a divalent cation-dependent manner. *J. Biol. Chem.* 269:26006–10
- Kern A, Briesewitz R, Bank I, Marcantonio EE. 1994. The role of the I domain in ligand binding of the human integrin $\alpha 1\beta 1$. *J. Biol. Chem.* 269:22811–16
- Kirschner DA, Inouye H, Ganser AL, Mann V. 1989. Myelin membrane structure and composition correlated: a phylogenetic study. *J. Neurochem.* 53:1599–609
- Kornbliht AR, Umezawa K, Vibe-Pedersen K, Baralle FE. 1985. Primary structure of human fibronectin: differential splicing may generate at least 10 polypeptides from a single gene. *EMBO J.* 1755–59
- Kraulis PJ. 1991. MOLSCRIPT: a program to produce both detailed and schematic plots of proteins. *J. Appl. Crystallogr.* 24:946–50
- Krezel AM, Wagner G, Seymour-Ulmer J, Lazarus RA. 1994. Structure of the RGD protein decorsin: conserved motif and distinct function in leech proteins that affect blood clotting. *Science* 264:1944–47
- Kuntz ID. 1992. Structure-based strategies for drug design and discovery. *Science* 257:1078–82
- Landis RC, McDowall A, Holness CLL, Littler AJ, Simmons DL, Hogg N. 1994. Involvement of the “I” domain of LFA-1 in selective binding to ligands ICAM-1 and ICAM-3. *J. Cell Biol.* 126:529–37
- Leahy DJ, Aukhil I, Erickson HP. 1996. 2.0 Å crystal structure of a four-domain segment of human fibronectin encompassing the RGD loop and synergy region. *Cell* 84:155–64
- Leahy DJ, Axel R, Hendrickson WA. 1992a. Crystal structure of a soluble form of the human T cell coreceptor CD8 at 2.6 Å resolution. *Cell* 1145–62
- Leahy DJ, Hendrickson WA, Aukhil I, Erickson HP. 1992b. Structure of a fibronectin type III domain from tenascin phased by MAD analysis of the selenomethionyl protein. *Science* 258:987–91
- Lee J-O, Rieu P, Arnaout MA, Liddington R. 1995a. Crystal structure of the A domain from the α subunit of integrin CR3 (CD11b/CD18). *Cell* 80:631–38
- Lee J-O, Bankston LA, Arnaout MA, Liddington RC. 1995b. Two conformations of the integrin A-domain (I-domain): a pathway for activation. *Structure* 3:1333–40
- Lemke G, Axel R. 1988. Isolation and sequence of a cDNA encoding the major structural protein of peripheral myelin. *Cell* 40:501–8
- Lesk AM, Branden CI, Chothia C. 1989. Structural principles of α/β proteins: the packing of the interior of the sheet. *Proteins* 5:139–48
- Li S, Gao J, Satoh T, Friedman TM, Edling AE, et al. 1997. A computer screening approach to immunoglobulin superfamily structures and interactions: discovery of small non-peptidic CD4 inhibitors as novel immunotherapeutics. *Proc. Natl. Acad. Sci USA* 94:73–78
- Main AL, Harvey TS, Baron M, Boyd J, Campbell ID. 1992. The three-dimensional structure of the tenth type III molecule of fibronectin: an insight into RGD-mediated interactions. *Cell* 71:671–78
- Michishita M, Videm V, Arnaout MA. 1993. A novel divalent cation-binding site in the A domain of the $\beta 2$ integrin CR3 (CD11b/CD18) is essential for ligand binding. *Cell* 72:857–67
- Murray AJ, Lewis SJ, Barclay AN, Brady RL. 1995. One sequence, two folds: a metastable structure of CD2. *Proc. Natl. Acad. Sci. USA* 92:7337–41
- Murzin AG. 1996. Structural classification of proteins: new superfamilies. *Curr. Opin. Struct. Biol.* 6:386–94
- Nagar B, Overduin M, Ikura M, Rini JM. 1996. Structural basis of calcium-induced E-cadherin rigidification and dimerization. *Nature* 380:360–64
- Nose A, Tsuji K, Takeichi M. 1990. Localization of specificity determining sites in cad-

- herin cell adhesion molecules. *Cell* 61:147–55
- Osborn L, Hession C, Tizard R, Vasallo C, Luhowskyj S, et al. 1989. Direct expression cloning of vascular cell adhesion molecule 1, a cytokine-induced endothelial protein that binds to lymphocytes. *Cell* 59:1202–11
- Overduin M, Harvey TS, Bagby S, Tong KI, Yau P, et al. 1995. Solution structure of the epithelial cadherin domain responsible for selective cell adhesion. *Science* 267:386–89
- Overduin M, Tong KI, Kay CM, Ikura M. 1996. ¹H, ¹⁵N, ¹³C resonance assignments and monomeric structure of the amino-terminal extracellular domain of epithelial cadherin. *J. Biomol. NMR* 7:173–89
- Patthy L. 1990. Homology of a domain of the growth hormone/prolactin receptor family with type III modules of fibronectin. *Cell* 61:13–14
- Petersen TE, Thogersen HC, Skorstengaard K, Vibe-Pedersen K, Sahl P, et al. 1983. Partial primary structure of bovine plasma fibronectin. Three types of internal homology. *Proc. Natl. Acad. Sci. USA* 80:137–41
- Pfuhl M, Pastore A. 1995. Tertiary structure of an immunoglobulin-like domain from the giant muscle protein titin: a new member of the I set. *Structure* 3:391–401
- Pierschbacher MD, Ruoslahti E. 1984. Cell attachment activity of fibronectin can be duplicated by small synthetic fragments of the molecule. *Nature* 309:30–33
- Pokutta S, Herrenknecht K, Kemler R, Engel J. 1994. Conformational changes of the recombinant extracellular domain of E-cadherin upon calcium binding. *Eur. J. Biochem.* 223:1019–26
- Poljak RL, Amzel LM, Avey HP, Chen BL, Phizackerley RP, Saul F. 1973. Three-dimensional structure of the Fab fragment of a human immunoglobulin at 2.8-Å resolution. *Proc. Natl. Acad. Sci. USA* 70:3305–10
- Potts JR, Campbell ID. 1994. Fibronectin structure and assembly. *Curr. Opin. Cell Biol.* 6:648–55
- Qu A, Leahy DJ. 1995. Crystal structure of the I-domain for the CD11a/CD18 (LFA-1, $\alpha_L\beta_2$) integrin. *Proc. Natl. Acad. Sci. USA* 92:10277–81
- Qu A, Leahy DJ. 1996. The role of the divalent cation in the structure of the I domain from the CD11a/CD18 integrin. *Structure* 4:931–42
- Randi AM, Hogg N. 1994. I domain of β_2 integrin lymphocyte function-associated antigen-1 contains a binding site for ligand intercellular adhesion molecule-1. *J. Biol. Chem.* 269:12395–98
- Reardon D, Farber GK. 1995. The structure and evolution of α/β barrel proteins. *FASEB J.* 9:497–503
- Richardson JS, Richardson DC, Thomas KA. 1976. Similarity of three-dimensional structure between the immunoglobulin domain and the copper, zinc superoxide dismutase subunit. *J. Mol. Biol.* 102:221–35
- Rossmann MG, Moras D, Olsen KW. 1974. Chemical and biological evolution of a nucleotide-binding protein. *Nature* 250:194–99
- Ryu S-E, Kwong PD, Truneh A, Porter TG, Arthos J, et al. 1990. Crystal structure of an HIV-binding recombinant fragment of human CD4. *Nature* 348:419–26
- Ryu S-E, Truneh A, Sweet RW, Hendrickson WA. 1994. Structures of an HIV and MHC binding fragment from human CD4 as refined in two crystal lattices. *Structure* 2:59–74
- Saudek V, Atkinson RA, Pelton JT. 1991. Three-dimensional structure of echistatin, the smallest active RGD protein. *Biochemistry* 30:7369–72
- Schiffer M, Girling RL, Ely KR, Edmundson AB. 1973. Structure of a lambda-type Bence-Jones protein at 3.5 Å resolution. *Biochemistry* 12:4620–31
- Schwartz MA, Schaller MD, Ginsberg MH. 1995. Integrins: emerging paradigms of signal transduction. *Annu. Rev. Cell Dev. Biol.* 11:549–99
- Shapiro L, Doyle JP, Hensley P, Colman DR, Hendrickson WA. 1996. Crystal structure of the extracellular domain from P0, the major structural protein of peripheral nerve myelin. *Neuron* 17:435–49
- Shapiro L, Fannon AM, Kwong PD, Thompson A, Lehmann MS, et al. 1995a. Structural basis of cell-cell adhesion by cadherins. *Nature* 374:327–37
- Shapiro L, Kwong PD, Fannon AM, Colman DR, Hendrickson WA. 1995b. Considerations on the folding topology and evolutionary origin of cadherin domains. *Proc. Natl. Acad. Sci. USA* 92:6793–97
- Singer SJ, Doolittle RF. 1966. Antibody active sites and immunoglobulin molecules. *Science* 153:13–25
- Smith JW, Cheresch DA. 1988. The Arg-Gly-Asp binding domain of the vitronectin receptor. Photoaffinity cross-linking implicates amino acid residues 61–203 of the β subunit. *J. Biol. Chem.* 263:18726–31
- Smith JW, Cheresch DA. 1991. Labeling of integrin $\alpha_V\beta_3$ with ⁵⁸Co(III). *J. Biol. Chem.* 266:11429–32
- Springer TA. 1997. Folding of the N-terminal, ligand-binding region of integrin α -subunits into a β -propeller domain. *Proc. Natl. Acad. Sci. USA* 94:65–72

- Staunton DE, Dustin ML, Erickson HP, Springer TA. 1990. The arrangement of the immunoglobulin-like domains of ICAM-1 and the binding sites for LFA-1 and rhinovirus. *Cell* 61:243–54
- Stern LJ, Brown JH, Jardetzky TS, Gorga JC, Urban RG, et al. 1994. Crystal structure of the human class II MHC protein HLA-DR1 complexed with an influenza virus peptide. *Nature* 368:215–21
- Takada Y, Ylanne J, Mandelman D, Puzon W, Ginsberg MH. 1992. A point mutation of integrin $\beta 1$ subunit blocks binding of $\alpha 5\beta 1$ to fibronectin and invasin but not recruitment to adhesion plaques. *J. Cell Biol.* 119:913–24
- Takeichi M. 1990. Cadherins: a molecular family important in selective cell-cell adhesion. *Annu. Rev. Biochem.* 59:237–52
- Takeichi M. 1991. Cadherin cell adhesion receptors as a morphogenetic regulator. *Science* 251:1451–55
- Thomsen NK, Soroka V, Jensen PH, Berezin V, Kiselyov VV, et al. 1996. The three-dimensional structure of the first domain of neural cell adhesion molecule. *Nat. Struct. Biol.* 3:581–85
- Tomschy A, Fauser C, Landwehr R, Engel J. 1996. Homophilic adhesion of E-cadherin occurs by a co-operative two-step interaction of N-terminal domains. *EMBO J.* 15:3507–14
- Vaughn DE, Bjorkman PJ. 1996. The (Greek) key to structures of neural adhesion molecules. *Neuron* 16:261–73
- Wagner G. 1995. E-cadherin: a distant member of the immunoglobulin superfamily. *Science* 267:342
- Wang J, Yan Y, Garrett TPJ, Liu J, Rodgers DW, et al. 1990. Atomic structure of a fragment of human CD4 containing two immunoglobulin-like domains. *Nature* 348:411–18
- Wang J-H, Pepinsky RB, Stehle T, Liu J-H, Karpusas M, et al. 1995. The crystal structure of an N-terminal two-domain fragment of vascular cell adhesion molecule 1 (VCAM-1): a cyclic peptide based on the domain 1 C-D loop can inhibit VCAM-1- $\alpha 4$ integrin interaction. *Proc. Natl. Acad. Sci. USA* 92:5714–18
- Warner LE, Hilz MJ, Appel SH, Killian JM, Kolodny EH, et al. 1996. Clinical phenotypes of different MPZ (P0) mutants may include Charcot-Marie-Tooth Type 1B, Dejerine-Sottas, and congenital hypomyelination. *Neuron* 17:451–60
- Williams AF. 1987. A year in the life of the immunoglobulin superfamily. *Immunol. Today* 8:298–303
- Williams AF, Barclay AN. 1988. The immunoglobulin superfamily—domains for cell surface recognition. *Annu. Rev. Immunol.* 6:381–405
- Williams MJ, Phan I, Harvey TS, Rostagno A, Gold LI, Campbell ID. 1994. Solution structure of a pair of fibronectin type 1 modules with fibrin binding activity. *J. Mol. Biol.* 235:1302–11
- Withka JM, Wyss DF, Wagner G, Arulanandam ARN, Reinherz EL, Recny MA. 1993. Structure of the glycosylated adhesion domain of human T lymphocyte glycoprotein CD2. *Structure* 1:69–81
- Wong YW, Williams AF, Kingsmore SF, Seldin MF. 1990. Structure, expression, and genetic linkage of the mouse BCM1 (OX45 or Blast-1) antigen. *J. Exp. Med.* 171:2115–30
- Yamada T, Song H, Inaka K, Shimada Y, Kikuchi M, Matsushima M. 1995. Structure of a conformationally constrained Arg-Gly-Asp sequence inserted into human lysozyme. *J. Biol. Chem.* 270:5687–90



CONTENTS

Genetics of Transcriptional Regulation in Yeast: Connections to the RNA Polymerase II CTD, <i>Marian Carlson</i>	1
Mitochondrial Preprotein Translocase, <i>Nikolaus Pfanner, Elizabeth A. Craig, Angelika Hönlinger</i>	25
Left-Right Asymmetry in Animal Development, <i>William B. Wood</i>	53
Microtubule Polymerization Dynamics, <i>Arshad Desai, Timothy J. Mitchison</i>	83
Molecular and Functional Analysis of Cadherin-Based Adherens Junctions, <i>Alpha S. Yap, William M. Brieher, Barry M. Gumbiner</i>	119
Genetic Analysis of the Actin Cytoskeleton in the Drosophila Ovary, <i>Douglas N. Robinson, Lynn Cooley</i>	147
Assembly and Enlargement of the Primary Cell Wall in Plants, <i>Daniel J. Cosgrove</i>	171
Light Control of Plant Development, <i>Christian Fankhauser, Joanne Chory</i>	203
Adipocyte Differentiation and Leptin Expression, <i>Cheng-Shine Hwang, Thomas M. Loftus, Susanne Mandrup, M. Daniel Lane</i>	231
Cyclin-Dependent Kinases: Engines, Clocks, and Microprocessors, <i>David O. Morgan</i>	261
Initiation of DNA Replication in Eukaryotic Cells, <i>Anindya Dutta, Stephen P. Bell</i>	293
The LIN-12/Notch Signaling Pathway and Its Regulation, <i>Judith Kimble, Pat Simpson</i>	333
Implications of Atomic-Resolution Structures for Cell Adhesion, <i>Daniel J. Leahy</i>	363
Bacterial Cell Division, <i>David Bramhill</i>	395
Neural Cell Adhesion Molecules of the Immunoglobulin Superfamily: Role in Axon Growth and Guidance, <i>Frank S. Walsh, Patrick Doherty</i>	425
The Two-Component Signaling Pathway of Bacterial Chemotaxis: A Molecular View of Signal Transduction by Receptors, Kinases, <i>Joseph J. Falke, Randal B. Bass, Scott L. Butler, Stephen A. Chervitz, Mark A. Danielson</i>	457
Cellular Functions Regulated by SRC Family Kinases, <i>Sheila M. Thomas, Joan S. Brugge</i>	513
Formation and Function of Spemann's Organizer, <i>Richard Harland, John Gerhart</i>	611
Nuclear Assembly, <i>Tracey Michele Gant, Katherine L. Wilson</i>	669

Plant Cell Morphogenesis: Plasma Membrane Interactions with the Cytoskeleton and Cell Wall, <i>John E. Fowler, Ralph S. Quatrano</i>	697
The Design Plan of Kinesin Motors, <i>Ronald D. Vale, Robert J. Fletterick</i>	745
Structure, Function, and Regulation of the Vacuolar (H ⁺)-ATPase, <i>Tom H. Stevens, Michael Forgac</i>	779