Jia-huai Wang
Timothy A. Springer

Structural specializations of immunoglobulin superfamily members for adhesion to integrins and viruses

Authors' addresses
Jia-huai Wang[1], *Timothy A. Springer*[2],
[1]Dana-Farber Cancer Institute and Department of Pediatrics, Harvard Medical School, Boston, Massachussetts, USA.
[2]Center for Blood Research, Department of Pathology, Harvard Medical School, Boston, Massachussetts, USA.

Correspondence to:
Jia-huai Wang
Dana-Farber Cancer Institute and Department of Pediatrics
Harvard Medical School
44 Binney Street
Boston MA 02115
USA

Acknowledgements
We thank our colleagues for their important contributions to the X-ray crystal structures that are reviewed here, and in particular, Jose Casasnovas and Kemin Tan. We also thank Kemin Tan for preparing Figs. 2, 3, 5C & 7. This work was supported by NIH grants HL48675, AI31921, and CA31798.

Immunological Reviews 1998
Vol. 163: 197–215
Printed in Denmark.
Immunological Reviews
ISSN 0105-2896

Summary: The circulation and migration of leukocytes are critical for immune surveillance and immune response to infection or injury. The key step of leukocyte recruitment involves the adhesion between immunoglobulin superfamily (IgSF) proteins on endothelium and integrin molecules on leukocyte surfaces. Some of the IgSF members are subverted as virus receptors. Four crystal structures of N-terminal two-domain fragments of these IgSF proteins have been determined: intercellular adhesion molecule-1 (ICAM-1), ICAM-2, vascular adhesion molecule-1 (VCAM-1), and mucosal addressin cell adhesion molecule-1 (MAdCAM-1). An acidic residue near the bottom of domain 1 plays a key role in integrin binding. For ICAM-1 and ICAM-2, this glutamic acid residue is located on a flat surface, complementary to the flat surface of the I domain of the integrin to which they bind, lymphocyte function-associated antigen-1 (LFA-1). For VCAM-1 and MAdCAM-1, the acidic residue is aspartic acid, and it resides on a protruded CD loop which may be complementary to a more pocket-like structure in the α4 integrins to which they bind, which lack I domains. A number of unique structural features of this subclass of IgSF have been identified which are proposed to consolidate the domain structure to resist force during adhesion to integrins. Different mechanisms are proposed for the different CAMs to present the integrin-binding surface toward the opposing cell for adhesion, and prevent *cis* interaction with integrins on the same cell. Finally, CD4 and ICAM-1 are compared in the context of ligand binding and virus binding, which shows how human immunodeficiency virus and rhinovirus fit well with the distinct structural feature of their cognate receptors.

Introduction

Cell adhesion molecules of the immunoglobulin superfamily (IgSF) have important functions in cell-cell recognition in inflammation and immune responses. A subset of the IgSF proteins is specialized for binding to integrins, and can also be recognized by greater amino acid sequence homology to one another than to other IgSF members. These are the intercellular adhesion molecules (ICAM-1, ICAM-2, ICAM-3, ICAM-4, and ICAM-5), vascular adhesion molecule-1 (VCAM-1), and mucosal addressin cell adhesion molecule-1 (MAdCAM-1) (1–4). Recently, high-resolution X-ray crystal structures have been determined for the N-terminal two-domain fragments of four of these molecules, ICAM-1 (5, 6), ICAM-2 (7), VCAM-1

(8, 9), and MAdCAM-1 (10). All four are expressed on endothelium and function in leukocyte binding to the blood vessel wall and transendothelial emigration at inflammatory sites. ICAM-1 and VCAM-1 are greatly increased on endothelium by cytokines at inflammatory sites, and act as traffic signal molecules to promote selective leukocyte emigration at these sites. ICAM-2 is expressed widely and constitutively on endothelium. MAdCAM-1 is constitutively expressed on a highly specialized subset of blood vessels found in mucosal tissues, including high endothelial venules of Peyer's patches and venules in the lamina propria. MAdCAM-1 facilitates preferential trafficking to these tissues of a subset of lymphocytes that appears to have been selected by antigens in mucosal sites. All four IgSF members are expressed on non-vascular cells to varying extents. ICAM-2 is expressed on lymphocytes, monocytes, and platelets. ICAM-1 is inducible by cytokines on many cell types, including epithelial cells, connective tissue cells, monocytes, and antigen-presenting dendritic cells, and is upregulated on B lymphocytes in germinal centers. ICAM-1 and ICAM-2 are important in antigen-specific recognition by lymphocytes and in effector functions of neutrophils, monocytes, and natural killer cells. Very few epithelial cells express ICAM-1 constitutively, but among these are certain epithelial cells of the nasal mucosa (11). These cells are infected by human rhinovirus (HRV) serotypes that use ICAM-1 as their receptor.

The X-ray crystal structures of functionally critical two-domain fragments of ICAM-1, ICAM-2, VCAM-1, and MAdCAM-1 yield important insights into IgSF: integrin interactions, and how the IgSF members can selectively recognize different integrins. Integrins are formed from α and β subunits that can associate into at least 21 unique αβ heterodimers. Both the α and β subunits contribute to ligand specificity, and influence two-way signaling through integrins. "Outside-in" signaling regulates the cell in response to the external environment; "inside-out" signaling allows integrin conformation, affinity for ligand, and lateral associations of the extracellular domain to be rapidly regulated in response to signals from within the cell (12–14). The integrin lymphocyte function-associated antigen-1 (LFA-1) contains the αL and β2 integrin subunits, and binds ICAM-1 and ICAM-2. The integrins α4β1 (very late antigen-4, VLA-4) and α4β7 preferentially bind VCAM-1 and MAdCAM-1, respectively (*Table* 1).

ICAM-1 and VCAM-1 contain five and usually seven IgSF domains, respectively, whereas ICAM-2 and MAdCAM-1 each contain two IgSF domains. The crystal structures are all of the two most N-terminal IgSF domains. The first domain always has an important integrin-binding surface, and the second domain of VCAM-1 and MAdCAM-1 directly contributes to integrin recognition as well. The second domain also appears to be important for orienting the recognition surface on the first domain.

Table 1. Integrins and IgSF molecules

Integrin	Subunits	I domain	Preferential IgSF ligand
LFA-1	αLβ2	+	ICAM-1,2,3,4,5
Mac-1	αMβ2	+	ICAM-1
VLA-4	α4β1	–	VCAM-1
LPAM	α4β7	–	MAdCAM-1

The recognition surface on the IgSF domain to which integrins bind is unique compared to other known IgSF recognition modalities, and we will compare the integrin recognition sites to those utilized by other adhesion molecules and by antibodies. Furthermore, we will discuss how the recognition surfaces of ICAMs differ from VCAM-1 and MAdCAM-1 and how this correlates with structural differences between the integrins that bind to these IgSF members.

Orientation of adhesion molecules on the cell surface is important for optimal exposure of the binding sites for recognition in *trans* by counter-receptors on other cells (7). Orientation is also important to prevent recognition in *cis* by counter-receptors expressed on the same cell. ICAM-1 and ICAM-2 are frequently co-expressed on leukocytes with the integrin to which they bind, LFA-1, and all evidence to date suggests that *cis* interactions cannot occur. We will discuss architectural features important to orient IgSF molecules so that binding can occur in *trans* but not in *cis*.

Adhesion molecules differ from receptors for soluble ligands in that the receptor:ligand interaction surface (bond) is often subjected to force (15). We therefore discuss features of the recognition surface and of the linkages between domains that may suit the integrin-binding IgSF members to resist the forces applied to them. These forces would otherwise tend to break apart the receptor-ligand interface and denature the domains that link this interface to the cell.

Finally, the IgSF molecules ICAM-1 and CD4 are recognized by viruses, and we will contrast the recognition sites utilized by viruses to those utilized in cell adhesion. Interactions with these molecules have important consequences for the virus above and beyond the receptor function. Binding of ICAM-1 to rhinovirus triggers release of the infectious RNA (16–18); binding of CD4 causes a conformational change in gp120 of human immunodeficiency virus (HIV) that enables binding to co-receptors, and then fusion of the viral and cellular membranes (19).

The IgSF domain as a modular component of cell-surface proteins

A high proportion of cell-surface molecules are modular, i.e. they are built of multiple domains arranged in tandem along the polypeptide chain. Truly modular domains have been defined as those that can be found adjacent to or interspersed between other types of domains (20); the IgSF domain fulfils this definition. IgSF domains, and the majority of the other widespread modules, have phase 1 introns near the domain boundary, i.e. the codon is split between nucleotides 1 and 2. This facilitates tandem duplication and exchange of exons, so that cell-surface molecules with widely varying functions can be built from a relatively small number of modules (20). It is hypothesized that the number of modular extracellular proteins greatly expanded in the evolution of metazoa, and fulfilled the need for cell-cell recognition in multicellular organisms. From an evolutionary point of view, the IgSF module must have particularly attractive features, because it is the most abundant module known, with at least 1,000 examples not including species homologs (20). On the human leukocyte surface, about 40% of proteins are predicted to contain IgSF domains (21). Furthermore, the IgSF domain is found as a module within intracellular proteins such as titin and twitchin in muscle (2).

The IgSF domain is evolutionarily ancient, as exemplified by both intracellular and extracellular IgSF domains in C. elegans (2); the function of the IgSF domain in antigen-specific recognition by antibodies and T-cell receptors is an evolutionarily recent adaptation of vertebrates. Nonetheless, the three-dimensional structure of the IgSF domain was first characterized in antibodies. Since the nomenclature of IgSF domains is with reference to antibodies, and the types of interactions seen between domains in antibodies provide important comparisons to those seen with adhesion molecules, it is worthwhile to begin with a brief description of Ig domains.

The structural unit of the Ig domain consists of two antiparallel β-sheets that are packed or "sandwiched" together. Each sheet is amphipathic, with the hydrophobic side facing inward and acting as the butter that holds the sandwich together, and the hydrophilic side facing outward. Usually a conserved disulfide bond links together the two sheets through cysteines present near the center of each sheet. The β-strands in the IgSF domain are lettered A through G, in the order in which they appear in the amino acid sequence (Fig. 1). Three of the connecting peptides (loops) between these strands link strands located in opposite sheets. These loops provide another mechanism to link together the two β-sheets and help make the IgSF domain a stable module. The order of the strands within the two sheets defines the topology of the IgSF domain. Despite the common fold, IgSF members display tremendous variation in edge strands and loop conformation (Fig. 1) (2). The V domain is the largest, and all other domains can be thought of as topologically, although not necessarily evolutionarily, derived from the V domain by removal of edge strands (Fig. 1). The BE strands in one sheet and GFC strands in the other are always present. These strands form the core of the Ig fold, and are linked at one end by the BC and at the other by the EF loops. The B, C, E, and F strands are thus "crossover strands". *In vitro* folding experiments have shown that several residues in these strands nucleate folding of the Ig domain, i.e. form a "crossover nucleus". These are the most highly conserved residues in IgSF domains (22). The only other loops connect neighboring strands in sheets, and connect the edge of one sheet to the edge of the other, which allows topology to be preserved despite variation in edge strands.

Ig domains have dimensions of approximately 2 × 2.5 × 4 nm, and for the sake of description can be considered to have six sides termed faces, edges, and ends. The surface of each β-sheet that faces outward is known as a face. Thus, the V domain has an ABED face and an A'GFCC'C" face. The side of the domain formed by the edge strands will be termed an edge. Thus, in C2 domains the C strand of the GFC sheet and the D strand of the ABED sheet form the CD edge (Fig. 1). The loops in IgSF domains are named after the strands that they connect, and the ends of the domain formed by the loops will be termed the top and the bottom, for the ends that contain the N-terminus and C-terminus of the domain sequence, respectively. The loops at the top of the V domain are the BC, C'C", DE, and FG loops. The complementarity-determining loops of antibodies and T-cell antigen receptors are formed by the BC, C'C", and FG loops. Binding of antibodies to antigens was the first example of recognition by IgSF domains. It was therefore widely hypothesized that recognition by other IgSF family members would also involve the loops at the top of domains. However, one of our points will be that IgSF domains are highly versatile, and can use almost any surface for recognition.

Chothia & Janin (23) proposed that β-sheets pack in an "aligned" mode. In this mode, the strands from two sheets run in parallel, making an angle of 30°, and the side chains from one strand meet their counterparts from the opposite sheet in an aligned way throughout the whole interface. This is usually true for intradomain β-sheet packing as in the Ig β-sandwich. When two Ig domains pack against each other, however, their strands are misaligned due to side chain intercalation. The two interacting β-sheet faces usually make an angle of 60° with respect to each other. This has been observed for both variable

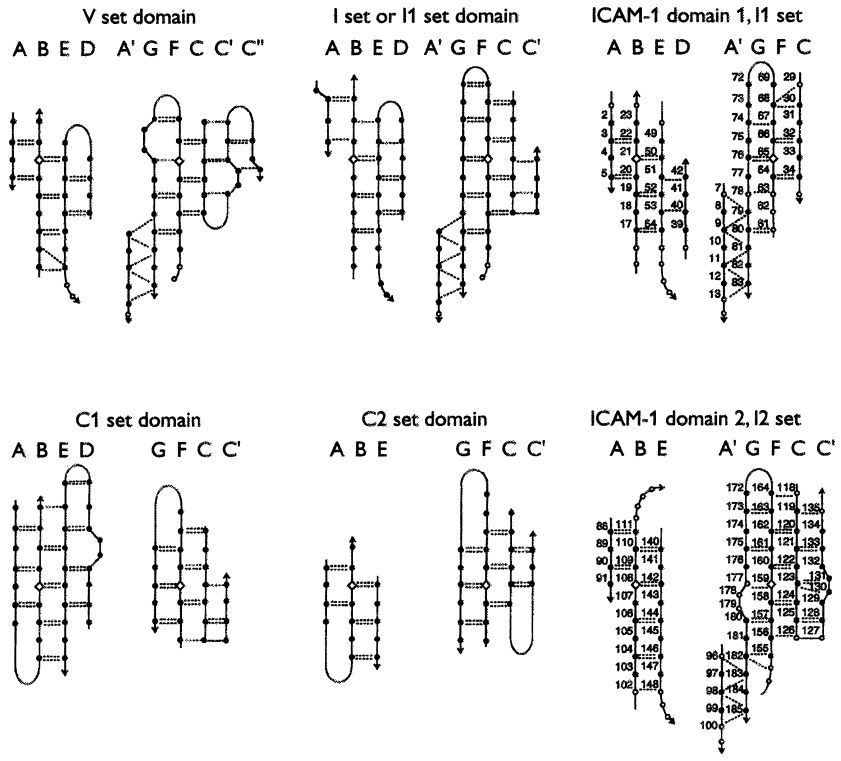

Fig. 1. IgSF domain frameworks. The frameworks for I set, V set, C1 set, and C2 set domains are adapted from Chothia & Jones (2), and represent the structural features usually conserved in each of these sets. Solid circles represent β-sheet residues, and open circles conserved loop residues. The frameworks for ICAM-1 domain 1 (I1 set) and ICAM-1 domain 2 (I2 set) are from (5). Here, solid circles represent residues defined as β-strand by DSSP (100), and open circles represent other residues that are largely conserved in domain 1 of ICAM-2, VCAM-1, and MAdCAM-1, or domain 2 of ICAM-2 and VCAM-1. Dashed lines represent backbone hydrogen bonds. The lines represent bonds to backbone NH groups (before each circle in the amino acid sequence) and CO groups (after each circle). The directionality of the polypeptide chain is indicated by arrows. The A' and G strands are parallel β-strands, and their hydrogen bond network therefore appears different from the others. The positions usually occupied by the cysteines that form the intersheet disulfide are shown as diamonds. This allows comparison between the different frameworks. The frameworks differ not only in strand content, but also in how far strands extend "above" and "below" these cysteines in sequence. Each domain is shown as two sheets. To imagine the two sheets aligned as in the three-dimensional structure, the left sheet should be moved over the right sheet, so that the cysteines marked with diamonds in the B and F strands are aligned.

domain interdomain pairing (24) and constant domain interdomain pairing (25) in antibodies.

In all IgSF members studied to date, there is a rotation of approximately 130° to 170° between successive domains, so that the faces exposed to one side alternate. Therefore, in antibodies, whereas the VH and VL domains pair along the A'GFCC'C" face, the succeeding CH1 and CL domains pair along the opposite ABED face. The CH2 domains face each other along their GFC faces, but there is solvent in between and these domains do not bind one another. The CH3 domains bind one another through their ABED faces, and therefore pairing along this face is considered characteristic of antibody C domains. These types of pairings form a reference point for discussing both recognition and dimerization interactions among adhesive IgSF members.

Structural features of the IgSF domains of the integrin ligands

Domain structure

At first glance, the four structures for the N-terminal two-domain fragments of ICAM-1, ICAM-2, VCAM-1, and MAdCAM-1 look very similar (Fig. 2). The sequences of these four proteins align well structurally in many important regions that are unique to them (Fig. 3). They each consist of two concatenated Ig-like domains. Domain 1 is smaller (83 to 90 residues) than domain 2 (102 to 112 residues). Part of the extra length of domain 2 is accounted for by two long loops that wrap around the bottom of domain 1. In contrast, in other known structures of IgSF proteins, such as CD4, CD2, antibodies, and T-cell receptors, the functionally important N-terminal domain 1 is larger than domain 2, and these domains have about 110 and 100 residues, respectively. Compared to these IgSF molecules, the tops of the N-terminal domains are com-

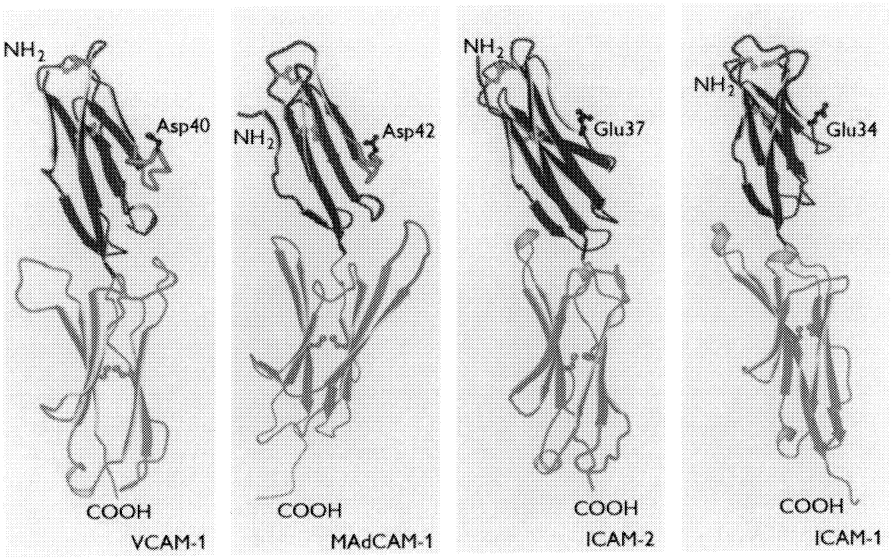

Fig. 2. Crystal structures of N-terminal two-domain fragments of VCAM-1 (8, 9), MAdCAM-1 (10), ICAM-2 (7), and ICAM-1 (5). Domains 1 and 2 are colored red and green, respectively, while the disulfide bonds are in yellow. The pivot point between the two domains, an invariant tyrosine, is colored purple. The "extra" disulfide bond is near the tip of the molecules. The acidic side chains for the key integrin-binding residue are shown as balls and sticks. Note that these key residues are all on the CD edge. The aspartic residue of VCAM-1 and MAdCAM-1 resides on the protruded CD loop, whereas the glutamic residue of ICAM-2 and ICAM-1 is located at the end of strand C on a flat surface. The DE β-ribbon extension on domain 2 of MAdCAM-1 is drawn in yellow. It is on the same face as Asp42 in domain 1. For clarity, the glycans are not shown. The figure was prepared with MOLSCRIPT (101).

pact and slender, because they lack C' and C" strands. The BC and FG loops at the top of domain 1 are consolidated by an "extra" disulfide that is characteristic of domain 1 in IgSF molecules that are integrin ligands. The surface area buried in the interface between domain 1 and domain 2 is relatively small, for example, 300 Å^2 per domain in ICAM-1 (5). All of these molecules contain an acidic residue that is of crucial importance for binding to integrins. This residue, a Glu in ICAM-1 and 2 and an Asp in VCAM-1 and MAdCAM-1, is located near the bottom of domain 1 (*Fig.* 2).

In ICAM-1, ICAM-2, and VCAM-1, domains 1 and 2 differ from one another in a manner that is analogous to the difference between the C1 and the C2 sets. The V set and C1 set of IgSF domains are named based on similarity to antibody V and C domains in content of edge strands (*Fig.* 1). There are also differences in how far strands extend "above" and "below" the B and F strand cysteines, and in certain framework residues (2). The C set has been subdivided into the C1 and C2 sets. C1 set domains contain a D strand and are found in antibody, T-cell receptor, and MHC molecules. The C1 set has been variously defined as lacking (21) or containing (2) a C' strand. The C' strand is short in domains that contain D strands and lack C" strands, since it runs along the bottom of the domain rather than all the way from the bottom to the top of the domain. C2 set domains lack a D strand and are found in adhesion molecules. The I set domain has a topology intermediate between the V set and the C set, and is largely built from the V set framework (*Fig.* 1) (26). I set domains are found in muscle proteins and in *Caenorhabolitis elegans*, and are hypothesized to be the type of IgSF domain found earliest in evolution (2).

The structures of the IgSF integrin ligands define a new subdivision of the I set into the I1 and I2 sets. Careful examination of the structures of the IgSF integrin ligands suggests that both domains 1 and 2 of these molecules belong to the I set (*Fig.* 1) (5). The key feature is the presence of the A-A' kink and the A' strand on one edge of the β-sandwich. An important feature of the V set and I set frameworks that distinguishes them from the C set frameworks is that the A strand runs about halfway down the domain in the ABED sheet, then crosses over the β-sandwich at a kink which usually involves a *cis* proline, and then joins the A'GFCC' sheet. The A' and G strands run in the same direction, and thus are the only example of parallel β-strands in the Ig domain (*Fig.* 1). We have proposed the division of I set domains into the I1 and I2 sets, based on the presence of a D strand in the I1 set and not in the I2 set (7). This division is strictly analogous to the division of the C set into the C1 and C2 sets, which differ in presence and absence of D strands, respectively. The I1 set also resembles the C1 set in that a short C' strand is found in some but not all members. Domain 1 of all four CAMs belongs to the I1 set, as do previously defined I set domains. The I2 set is defined by domain 2 of ICAM-1, ICAM-2, and VCAM-1. Their A-A' kink and the position of the β-sheet framework positions relative to the cysteines in the B and F strands distinguish them from C2 set domains (*Fig.* 1). Domain 2 of MAdCAM-1 differs from domain 2 of ICAM-1, ICAM-2, and VCAM-1 since it contains a D strand and belongs to the I1 set. MAdCAM-1 is more homologous to VCAM-1 than to any other IgSF molecule, and therefore domain 2 may represent interconversion in evolution between the I1 and I2 sets.

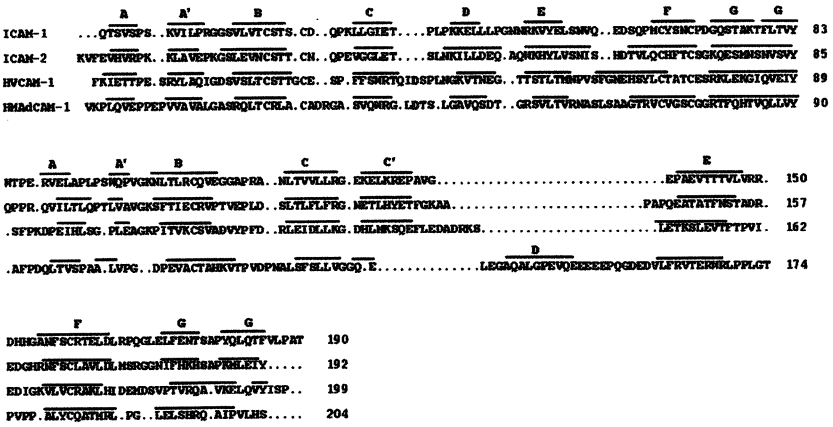

Fig. 3. Structural alignment of the four sequences. β-strands defined by DSSP (100) are overlined. The α-carbons of the four structures were aligned to a common framework with 3DMALIGN of MODELLER (102) (http://guitar.rockefeller.edu/modeller/modeller.html). The key integrin-binding residues on domain 1 are colored red. They are not aligned between ICAMs and VCAM-1/MAdCAM-1, reflecting the subdivision of these IgSF molecules into ligands for I-domain-containing and I-domain-lacking integrins, respectively. In domain 2, MAdCAM-1, but not the other IgSF integrin ligands, contains a D strand. The extended DE loop is shown in purple. The invariant tyrosine that is the last residue of domain 1, the invariant *cis* proline in the BC loop of domain 2, and three other conserved residues that participate in the conserved hydrophobic patch in the interdomain interface are shown in magenta.

The I set classification has an important functional significance for integrin-binding IgSF domains. Stress forces acting on IgSF and integrin molecules as they function in adhesion in the vasculature or to provide traction during cell migration will strain, i.e. lengthen, distort, and tend to pull apart the receptor-ligand-binding interface and the domains that lie between this interface and anchorage points in the cell membrane and cytoskeleton. Distortion of the domain bearing the ligand-binding interface, or of domains that are interdependent in conformation with the ligand-binding domain, will distort the ligand-binding site, and increase the rate constant for receptor-ligand dissociation (15). The I set structure seen in domains 1 and 2 of integrin ligands may be important in ability to resist tensile forces. The A-A' kink, along with many other structural features, helps to consolidate the domain structure. The movement of the polypeptide chain from one sheet to the other at the kink serves to knit together the two β-sheets at the middle of one edge of the domain. Within a domain, tensile forces will be distributed over multiple β-strands; however, at domain junctions, forces will be borne primarily by the G strand of the domain on the N-terminal side and the A strand of the domain on the C-terminal side. The presence of the A' strand means that the G strand is not an edge strand, as it would be in a C2 domain, and thus the G strand should be kinetically less susceptible to being peeled by applied force from the β-sheet. Not all IgSF adhesion molecules may be designed to withstand tensile forces; some may be designed to signal cell-cell contact but not to withstand force. Integrins in general are associated with cell migration, which requires "traction" or resistance to slippage at adhesive contacts as cells migrate through tissues; integrins are also associated with focal contacts where the actin cytoskeleton terminates and that are found at stressed sites such as myotendinous junctions. Thus, it may be particularly important for IgSF molecules that bind integrins to have a structure that is resistant to mechanical stress. It is interesting that IgSF domains in muscle proteins that resist force, such as titin, also belong to the I set (2).

Domain junction and the pivot

The junction between domain 1 and 2 has a conserved structural pattern among all of the four CAMs. The last residue of domain 1 is an invariant tyrosine, which is an integrated member of the A'GFCC' β-sheet hydrogen-bonding network of domain 1 (*Figs.* 1 & 3). The residue immediately following forms a "bridge" to the BC loop of domain 2 via two main chain hydrogen bonds (*Fig.* 4). An invariant proline on the BC loop lies just after the "bridge". This proline is in *cis* configuration such that its pyrrolidine ring facilitates non-polar interaction with a group of conserved hydrophobic residues from both domain 1 and 2 to form a hydrophobic patch in the interdomain interface. The hydrophobic patch brings together the bottom part of the A strand from domain 1 and the FG loop of domain 2. Except for ICAM-2, there is another hydrophobic patch on the other side of the interface, in which the EF loop of domain 1 and the C'E loop (or DE loop in MAdCAM-1) of domain 2 interact. The two hydrophobic patches define the nature of domain 1-domain 2 interface for ICAM-1, VCAM-1, and MAdCAM-1.

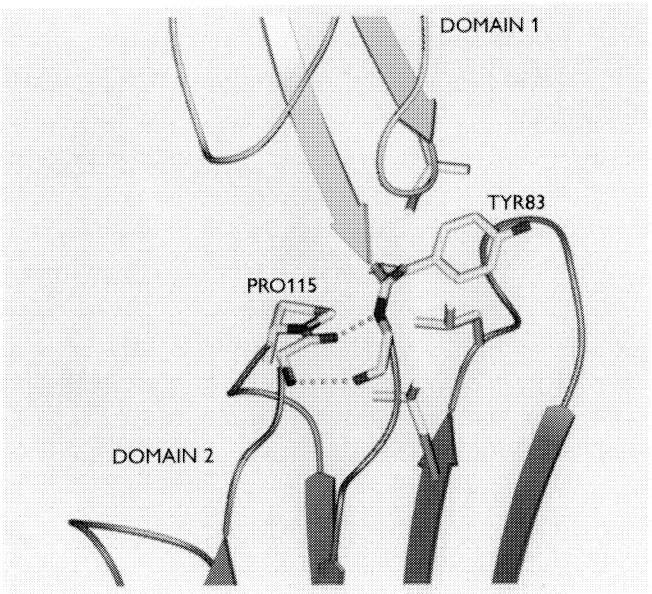

Fig. 4. The interdomain junction in ICAM-1. Side chains are shown only for conserved residues that make up the hydrophobic patch. Backbone amide and carbonyl groups only are shown for Trp84 and Ala114, which participate in the conserved hydrogen bond between the first residue in domain 2 and the residue preceding the *cis* proline in the BC loop of domain 2. The rest of the structure is shown as a ribbon drawing. The figure was prepared with SETOR (103).

Because the very first residue of domain 2 forms two hydrogen bonds to the BC loop of domain 2, interdomain flexibility comes mostly from rotation around a "pivot" point (9), between the invariant tyrosine at the end of domain 1 and the first residue of domain 2. Interdomain flexibility has been observed for ICAM-1, where two independent molecules are present per asymmetric unit (5), and in VCAM-1, where two independent molecules are present in each of two different crystal forms (9, 27). The variation in angle between domains is 17° for ICAM-1, and up to 35° for VCAM-1. The two hydrophobic patches act as "cushions" on which the domains seem to be able to rock back and forth in a cradle-like structure formed by the two long loops from the top of domain 2 (28). The pivot may enable multidomain CAMs like VCAM-1 and ICAM-1 to function in a dynamic environment. A similar junctional variability of at least 10° has been observed in the hydrophobic interface between domain 2 and 3 of CD4 (29). In contrast, CD4 has a fairly rigid linkage between domains 1 and 2 and between domains 3 and 4 (30–32). In CD4, β-strand G of domain 1 is continuous with β-strand A of domain 2. Furthermore, domain 1 inserts deeply into the top of domain 2. There is no pivot-like junction. The same is true in the CD4 domain 3-4 structure. It is conceivable that, for CD4, the two relatively rigid two-domain units (domain 1-2 and domain 3-4) separately interact with the T-cell receptor and MHC molecule, respectively, and the flexible domain 2-3 junction allows for minor adjustments to enable a good fit with T-cell receptor:MHC-binding interactions that differ in orientation. On the other extreme is the structure of CD2, where a linker loosely connects domain 1 and domain 2 (27).

For human MAdCAM-1, the two Ig-like domains are followed by a mucin-like domain that forms a long extension to the cell surface. Domain 2 of MAdCAM-1 has an unusually high proline content of 12.3%, compared to 5.1% for average proteins (33). In particular, these prolines are concentrated at the bottom of domain 2, in the AB loop and EF loops, near the junction to the mucin-like domain. They are predicted to rigidify this part of the structure against the force applied on the end of the G strand at the junction to the mucin-like domain during adhesive interactions (10).

The "extra" disulfide bond

A hallmark of IgSF members that function as integrin ligands is the presence of two disulfide bonds in the integrin-binding domain 1. The "extra" disulfide bond is located near the top of domain 1, and joins the BC loop to the end of the F strand. Domains 4–6 of VCAM-1 are highly homologous to domains 1–3. Like domain 1, domain 4 has integrin-binding activity, and also contains cysteines that are predicted by sequence homology to form an "extra" disulfide bond at the top of domain 4. In antibodies, the three complementarity-determining loops at the top of the V domain can vary widely in sequence, and adopt different canonical structures (34). The presence of one more loop (C'C" loop) at the top of V set domains than I set domains, and the association of VH with VL, may help consolidate the structure at the top of V domains. In IgSF integrin ligands, the function of the extra disulfide may be to help consolidate the top of the relatively "skinny" integrin-binding domain. In two independent ICAM-1 molecules, the FG and BC loops of domain 1 appear flexible only one residue away from the "extra" disulfide. Without the disulfide, this flexibility might extend further into the domain.

Local conformation of the key integrin-binding region

An acidic residue that is key for binding to integrins has been identified by mutagenesis in all IgSF ligands studied to date. This key residue is Glu34 in ICAM-1 (35), Glu37 in ICAM-3 (36, 37), Asp40 in VCAM-1 (38, 39), and Asp42 in MAdCAM-1 (40). The aspartic acid residue in VCAM-1 and MAdCAM-1 is located in a protruded CD loop, whereas the glutamic acid residue in ICAM-1 and ICAM-2, and by homology in ICAM-3, is at the end of the C strand. We will discuss the functional meaning of this difference in the next section. Here we focus on how an exquisite hydrogen bond network maintains the local conformation around the key integrin-binding region. In VCAM-1, the segment from Thr37 to Asp40 forms a β-turn in the CD loop of domain 1. Because the β-turn resembles the turn between a C and C' strand, we may consider this first part of the CD loop as a vestige of the C' strand that is present in many of the I1 set domains. The side chain atom Oγ of Thr37 makes bifurcated hydrogen bonds to both the main chain and side chain atoms of Ser41 (9) (*Fig. 5A*). In this arrangement, the Asp40 sits at the end of the turn, and its main chain conformation is restrained by a pair of hydrogen bonds around it. The three-hydrogen-bond network ensures that the peptide segment Glu38-Ile39-Asp40-Ser41-Pro42 is firmly linked to the C strand. As will be discussed later, this peptide segment may be one key structural element for integrin binding.

In the case of ICAM-1 and ICAM-2, there is no vestige of the C' strand whatsoever. The key integrin-binding Glu34 of ICAM-1 is at the end of the C strand and is followed by a CD loop that is only four residues long, and exhibits no protrusion from the main body of the domain as in VCAM-1 and MAdCAM-1. Nevertheless, there are also two hydrogen bonds around the Glu34 within this region. These are from the side chain atoms of Thr35 and Lys39 to the main chain carbonyl oxygens of Leu37 and Ile33, respectively. Furthermore, since Glu34 is part of the CFG β-sheet, it forms two β-sheet hydrogen bonds to Met64 on the F strand. Hence, the binding surface is reinforced (*Fig. 5B*). ICAM-2 has an extremely similar conformation in this region. Since the Glu-Thr-X-X-X-Lys motif (where X can be any residue) is conserved among all ICAMs known to date (7), the Thr and Lys side chain to main chain hydrogen bonds that stabilize the conformation of the integrin-binding surface appear to be evolutionarily conserved.

An even more intriguing architecture has been revealed around the key integrin-binding residue Asp42 in MAdCAM-1 (10). The loop bearing Asp42 in MAdCAM-1 protrudes in a very different position than that in VCAM-1, probably due to a two-residue deletion. The peptide segment Gly40-Leu41-Asp42-Thr43-Ser44 constitutes the CD loop. Arg70 in the F strand is totally buried beneath this loop, and plays a critical role in maintaining the loop conformation. This arginine is surrounded by seven hydrophobic residues, and there are no negatively charged side chains in its vicinity to neutralize its strong positive charge. Instead, the guanido group of Arg70 donates three hydrogen bonds to main chain carbonyl oxygens. In particular, two hydrogen bonds to the successive carbonyl groups of Asp42 and Thr43 appear to constrain the backbone to a γ-turn-like conformation encompassing Asp42 (*Fig. 5C*). The seven-membered ring of a γ-turn is an energetically less favorable conformation than the ten-membered ring of a β-turn. However, the enhanced hydrogen bonds (the bond lengths are all less than 2.6 Å) from the arginine that is buried in an otherwise entirely hydrophobic environment are expected to consolidate the whole local structure. Buried, unneutralized arginines are rare, but have previously been reported to play this type of structural role by forming multiple hydrogen bonds to backbone oxygens (41).

The integrin-binding surface

Integrins are a large structurally and functionally complex family of αβ heterodimeric cell adhesion molecules (12, 42). They recognize a broad spectrum of extracellular and surface proteins and have a variety of biological functions. The divalent cations Mg^{2+} or Ca^{2+} are required for ligand binding. The ligand-binding regions are located in the N-terminal globular "heads" of the α and β chains (1, 43), which are connected to the membrane by stalks. The N-terminal "head" piece of the α chain contains seven repeats of about 60 residues. These are predicted to fold into a seven-bladed β-propeller domain (44), similar to that found in G protein β subunits. The β-propeller domain is made up of seven four-stranded β-sheets arranged around a pseudo-seven-fold symmetry axis, with each β-sheet corresponding to one blade of the propeller. The correct prediction of the proximity of residues involved in an antibody epitope that are 100 residues apart in the sequence has provided support for the β-propeller fold (45). In some integrins, an inserted or I domain of about 200 residues is present between β-propeller sheets 2 and 3. The I domain has an important role in ligand binding by the integrins in which it is present (46, 47).

I domain ligands

The ICAMs bind to integrins that contain I domains, whereas VCAM-1 and MAdCAM-1 bind to integrins that lack I domains. This correlates with presentation of different binding surfaces

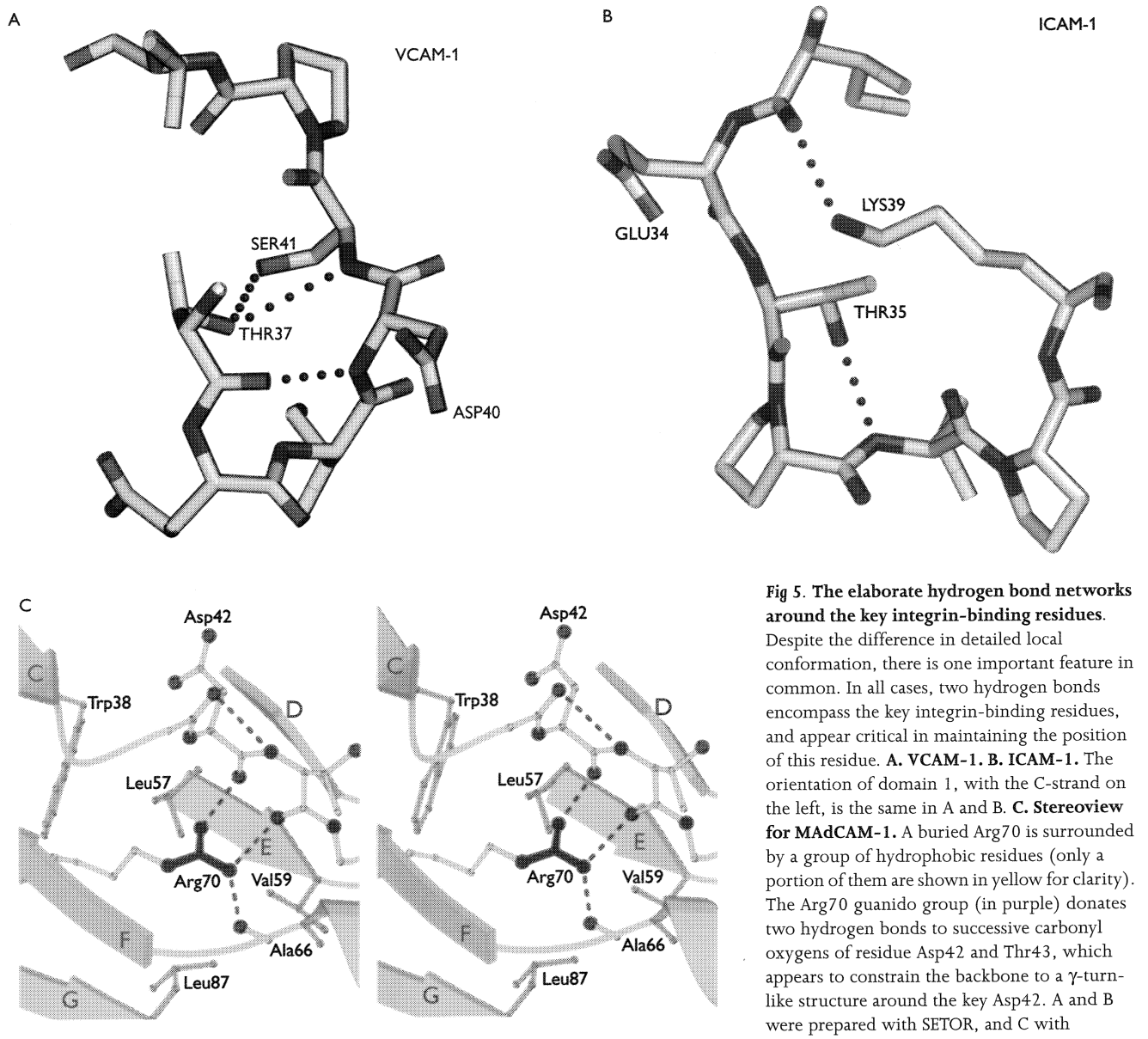

***Fig 5.* The elaborate hydrogen bond networks around the key integrin-binding residues.** Despite the difference in detailed local conformation, there is one important feature in common. In all cases, two hydrogen bonds encompass the key integrin-binding residues, and appear critical in maintaining the position of this residue. **A. VCAM-1. B. ICAM-1.** The orientation of domain 1, with the C-strand on the left, is the same in A and B. **C. Stereoview for MAdCAM-1.** A buried Arg70 is surrounded by a group of hydrophobic residues (only a portion of them are shown in yellow for clarity). The Arg70 guanido group (in purple) donates two hydrogen bonds to successive carbonyl oxygens of residue Asp42 and Thr43, which appears to constrain the backbone to a γ-turn-like structure around the key Asp42. A and B were prepared with SETOR, and C with MOLSCRIPT (101).

on the ICAMs, compared to VCAM-1 and MAdCAM-1. Mutagenesis studies suggest that the binding surface for LFA-1 on ICAM-1 and ICAM-3 (35–37) is located primarily in the C and F strands, with contributions of one residue each from the G strand and D strand. Around the key glutamic acid residue, the other important residues define a relatively flat surface covering the CD edge and CFG β-sheet of domain 1 (*Fig. 6A*) (7). This flat surface complements well with the corresponding binding surface of the I domain (48–50). The C'E loop of domain 2 is nearby, but makes no contribution to binding to LFA-1 (35), in contrast to contribution of its counterpart in VCAM-1 to binding to α4 integrins (51).

Crystal structures of I domains have recently been reported (48, 49). I domains have a "Rossmann" or dinucleotide binding fold. The structures most closely related to the I domain are the GTP-binding proteins such as the small G protein ras and the α subunit of trimeric G proteins (52, 53). I domains contain a metal ion-dependent adhesion site (MIDAS) that appears to be the ligand-binding site (48). The ligating residues in the MIDAS include a contiguous Asp-X-Ser-X-Ser sequence and a threonine and an aspartate from another part of the chain. These residues form a good octahedral co-ordination geometry for the central metal ion, Mg^{2+}, on the domain surface. Five of the six co-ordinating sites are occupied by the side chains from

protein and bound water molecules, leaving only one ligand position vacant. In one of the crystal structures, a glutamate from the neighboring molecule fills in to complete the co-ordination. This may mimic co-ordination to glutamate in a bound ligand (48). Indeed, mutagenesis experiments have shown that specificity for ligand is contributed by residues that surround the Mg^{2+} (50). These residues are present on a single, relatively flat face of the I domain, and this surface appears appropriate to bind the approximately flat surface surrounding the Glu in ICAMs. The Glu is proposed to directly ligate to the Mg^{2+}. Two different conformers of the Mac-1 I domain have been crystallized, one which binds Mg^{2+} and a Glu, and another which binds Mn^{2+} and no Glu (54). Notably, residues in G proteins that co-ordinate Mg^{2+} along with the γ-phosphate of GTP, and thus discriminate between whether GDP or GTP are bound, are structurally homologous to those in I domains that ligate Mg^{2+}. The difference in conformation between the two forms of the I domain resembles that seen in GDP and GTP-bound forms of G proteins, and thus conformational changes in I domains that regulate ligand binding may be structurally analogous to conformational changes in G proteins (54). However, this idea is not completely accepted. Similar changes have not been seen in the LFA-1 domain. It has been crystallized in the presence of Mg^{2+}, Mn^{2+}, and no metal; however, in no case was a pseudo-ligand bound, and in all cases the structure resembled that of the Mac-1 I domain with Mn^{2+} (55). Since the Glu-Mg^{2+} interaction in the crystal structure is a lattice interaction, it is possible that this conformation is a crystallographic artifact. A number of important issues remain to be resolved by further structure-function studies or an I domain-ICAM co-crystal.

Non-I domain ligands

The local conformation around the key acidic residue in ICAM-1 and ICAM-2 is remarkably distinct from that in VCAM-1 and MAdCAM-1 (*Fig. 2*). Structurally, sequences of the I domain ligands and non-I domain ligands do not align in this region (*Fig. 3*) (7). The key integrin-binding residue Asp40 in VCAM-1 is located on a protruded loop, not a flat surface (8, 9). A cyclic, but not a linear peptide, Cys-Gln-Ile-Asp-Ser-Pro-Cys, inhibits binding by α4β1 integrin-bearing cells binding to VCAM-1 (9). The cyclic peptide may somehow mimic the conformation of the integrin-binding motif around the Asp40. Other synthetic cyclic peptides can competitively inhibit α4β1 integrin binding to VCAM-1, with potency extending to the subnanomolar range (56, 57).

MAdCAM-1 and VCAM-1 both bind members of the α4 integrin family, with preference for α4β7 and α4β1, respectively. However, when domains 1 of VCAM-1 and MAdCAM-1

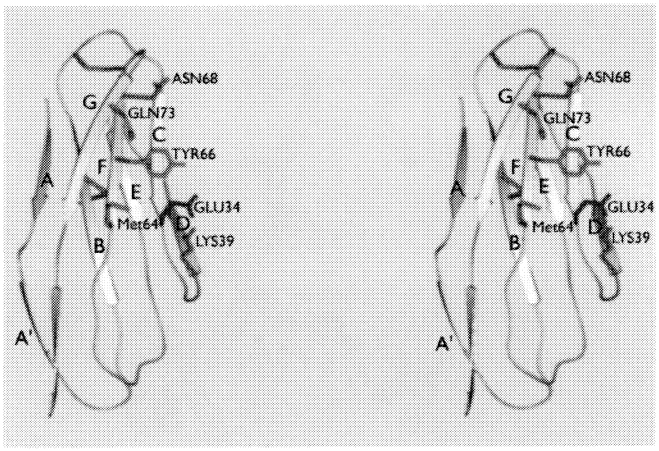

Fig. 6. Stereoview of the integrin-binding site of ICAM-1. Only domain 1 is shown, which contains all of the known integrin-binding residues (35, 104). The key Glu34 on the C strand is red, and other single amino acids that when mutated lower binding to LFA-1 by >50% or >2 s.d. are shown in green. The residues are roughly distributed on a relatively flat surface of domain 1. The figure was prepared with SETOR (103).

are superimposed, their integrin-binding loops differ markedly in position (*Fig. 7*). Asp42 in MAdCAM-1 and Asp40 in VCAM-1 lie 8 Å apart. The Asp42 in MAdCAM-1 is also less protruded compared to Asp40 in VCAM-1. However, if the five-residue peptide fragments centered at the key aspartate residue are superimposed, they match very well (*Fig. 7*). This five-residue fragment forms a "W"-shaped motif, with the acidic side chain of Asp markedly exposed for integrin binding. It is possible that this "W"-shaped local conformation is an important structural requirement for presenting the key α4 integrin-binding residue. Since this "W"-shaped motif occupies a markedly different position relative to the rest of domain 1, it may partly account for the preferences of α4β1 and α4β7 for VCAM-1 and MAdCAM-1, respectively.

The ICAMs differ from VCAM-1 and MAdCAM-1 in use of Glu rather than Asp, and presentation of the acidic residue on a relatively flat rather than a protruded surface. This correlates with recognition by integrins that contain and lack I domains, respectively. β-propeller enzymes all have their substrate-binding sites at the "top" of the domain, a face of the domain surrounding one end of the pseudo-symmetry axis, and which contains 14 loops in the case of seven β-sheets. Side-by-side loops run in opposite directions: a feature that is thought to be important in enzyme-active sites. The top of the integrin β-propeller domain functions in ligand binding by α4 integrins and other integrins that lack I domains (44, 58). Both the integrin α subunit β-propeller domain and a conserved domain in the integrin β subunit that may bind Mg^{2+} similarly to the I domain (14, 43) may contribute to ligand binding by those integrins

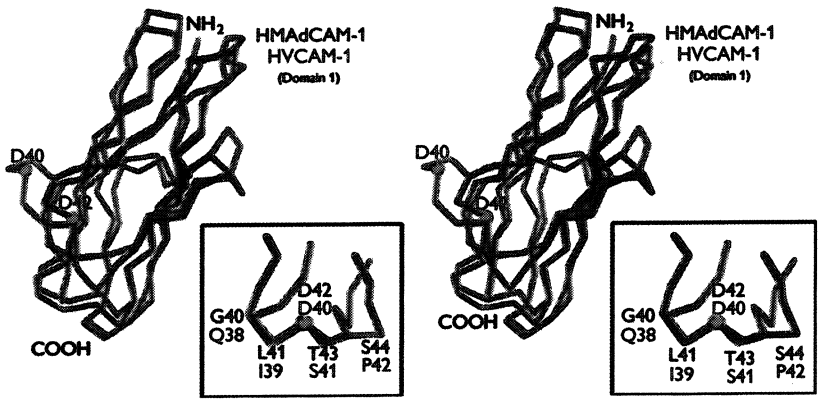

Fig. 7. Stereoview of the superposition of domain 1 between MAdCAM-1 and VCAM-1. The structures were aligned with 3DMALIGN in MODELLER (102) as in *Fig.* 3. The key Asp residues are not aligned, and are 8 Å away from one another. Inset: a superposition of local structure around the key Asp from the two molecules. A "W"-shaped conformation is present in both structures, which may be important for α4 integrin binding. The figure was prepared using MOLSCRIPT (101).

that lack I domains. These structures may differ from the I domain, in containing a binding site with a pocket that is complementary to the protruded CD loops of VCAM-1 and MAd-CAM-1, where the key aspartate residue sits.

The involvement of domain 2 of MAdCAM-1 and VCAM-1 in integrin binding

Mutagenesis has shown that residues in domain 2 of VCAM-1 as well as those in domain 1 may play a role in integrin binding (38, 39, 59). More recently, it has been shown that mutations of Asp143, Ser148, and Glu150 in VCAM-1 affect binding to α4 integrins (51). These residues are clustered in the C'E loop of domain 2, near the CD loop of domain 1. It is interesting that similar but non-identical residues in domain 1 of VCAM-1 are involved in the interaction with α4β1 and α4β7 (60). Strikingly, the residues that are differentially important, Asn44 and Glu66, are at the bottom of domain 1 and point towards domain 2 (60). The newly determined crystal structure of MAdCAM-1 demonstrates marked differences of its domain 2 with domain 2 of VCAM-1 (10). Domain 2 of MAdCAM-1 belongs to the I1 set, just like domain 1; i.e. domain 2 has only a very short C' strand and a long D strand. The long C' strand in domain 2 of VCAM-1 in effect switches to the other sheet in MAdCAM-1, and becomes the D strand adjacent to strand E. The DE loop is long and extended, and assumes a twisted, slightly "unzipped" β-ribbon-like structure (*Fig.* 2). This is an extraordinarily acidic structural motif, and a number of mutations of acidic residues in this structure affect binding to integrin α4β7 (M. Briskin, personal communication). The distance between Asp42 in domain 1 and the closest exposed acidic residue in the DE loop, Glu157, is about 20 Å. It is tempting to speculate that this motif in domain 2 and the key Asp42 and its surrounding region in domain 1 may interact with different regions of integrins, perhaps with different integrin subunits. In the next section, we will discuss another possible biological role this unique β-ribbon may play.

How integrin-binding surfaces are presented

The IgSF molecules that bind integrins reveal a novel recognition modality for Ig domains. Previous examples of recognition of IgSF proteins include the loops at the top of the domain in antibodies (61) and αβ T-cell receptors (62–64), the C'CFG face in CD2, LFA-3, and CD48 (27, 65), loops on the outside of the elbow between domains in the natural killer cell receptor (66), and the C'C" top corner and C" edge of CD4 (30, 31). The IgSF integrin ligands bind through a bottom loop and the edge of the domain. The critical acidic residue is closer to the bottom than to the top of the domain. Each CAM molecule apparently has its own way of orienting this binding surface so that it is properly exposed for recognition by a cognate integrin on an opposing cell.

Tripod-like glycans in domain 2 of ICAM-2

ICAM-2 is the only IgSF integrin ligand with just two Ig domains. A seven-residue stalk links domain 2 to the cell membrane. Two features of the ICAM-2 structure appear important to promote upward exposure of the key integrin-binding residue Glu37. First is the bend of about 35° between the two domains. In contrast to VCAM-1 and ICAM-1, domain-domain movement in ICAM-2 may be very limited. Although the conserved Pro117, Leu172, Leu170, Tyr85, and Ala13 hydrophobic patch exists just as in the other CAMs, on the other side of the domain interface there is a distinctive interdomain hydrogen bond network instead of a hydrophobic interdomain contact. A buried Arg175 plays a key role by forming multiple salt bridges with both domains. Ser83, Lys11, Glu116, Ser174, and a few water molecules are involved in an intricate network of

16 hydrogen bonds (7). These extensive polar interactions would appear to prevent flexion of the interdomain linkage, although there is no direct evidence for this because only one example of an ICAM-2 molecular structure is known. The second unique feature of ICAM-2 is the distribution of its glycans on the molecular surface (*Fig. 8A*). All six possible N-linked glycosylation sites have attached glycans. Three of these are uniformly distributed around the periphery of domain 2 near its bottom. This tripod-like arrangement of glycans is predicted to help orient domain 2 vertically on the cell membrane. The number of carbohydrate residues attached per N-linked site appears unusually large for ICAM-2, with an estimate of approximately 5,000 M_r or 25 residues/glycan from SDS-PAGE (67). Only two residues/glycan are visualized in the crystal structure (7), and thus one must imagine much longer and more elaborate glycans than shown in *Fig. 8*. These residues may have extensive hydrogen bond interactions with membrane phospholipid head groups, and their extensive bulk would also hinder bending of domain 2 with respect to the cell surface. Collectively, the glycans and the relatively rigid interdomain bend are predicted to maintain an orientation of domain 1 that exposes the integrin-binding site upward, ideally suited for binding to an integrin on an opposing cell (*Fig. 8*).

Orientation by glycans may be of general significance for cell-surface proteins. CD2, for example, has a glycan located at the bottom part of domain 2, which points towards the cell membrane (27). Human CD4 has two potential N-glycosylation sites, of which one is near the bottom of the membrane proximal domain, domain 4 (68). These glycans may help these cell-surface molecules assume an orientation in which they extend away from the membrane, and are accessible for interactions with other cells.

Dimeric ICAM-1 exposes its binding surface upwards

Apparently, the elegant architecture of ICAM-2 is suitable for orientation of a two-domain molecule. CAMs with a larger number of Ig domains, with the binding surface extending further from the cell surface, may require different orientation mechanisms. ICAM-1, for instance, has five Ig-like domains. In the crystal structure (5), there are two independent copies of the N-terminal two-domain fragment. One important observation is that, although these two molecules differ by 17° in the angle between domains 1 and 2, they both form dimers with their crystallographic 2-fold-symmetry mates. In both cases, domain 1 dominates dimer formation, with the symmetry-related ABED β-sheets packing face-to-face in an almost identical way. The dimer observed in the crystal may be biologically significant, because ICAM-1 is dimeric on the cell surface (69, 70). Furthermore, soluble dimeric ICAM-1 binds with higher avidity to LFA-1 than monomeric ICAM-1 (69, 71). The residue at the center of the dimer interface is Val51 in strand E, which is conserved as a Val or Ile in all known species of ICAM-1 molecules, but is hydrophilic in ICAM-2, which does not dimerize (70). The angle between the axes of the ABED β-sheets in the dimer interface is about 120°, whereas the angle is about 60° in interfaces between domains in antibodies (24, 25). For antiparallel β-sheets, packing at 120° places the two sheets at an equivalent angle to packing at 60°, i.e. the two modes are complementary. Interestingly, in the absence of constraints from other domains, Vδ (72) and Vβ domains (73) of the T-cell receptor can also pack in the complementary mode at 120° to form a dimer. This complementary mode of dimerization for ICAM-1 permits each of the monomers to have a 60° tilt relative to the dimer axis, which would be perpendicular to the cell surface if both monomers have the same orientation on the cell surface (*Fig. 8*). In this packing mode, the key integrin-binding Glu34 residues of each monomer are 42 Å apart and pointing away from each other on the opposite side of the dimer interface, and hence are well exposed towards the top of the molecules in a way reminiscent to the orientation of the bent ICAM-2 molecules. Electron micrographs of soluble five-domain ICAM-1 molecules demonstrate that ICAM-1 is bent (35), and the bend occurs between domain 3 and domain 4 (74). In the crystal dimer, the two C-termini in domain 2 point away from each other. The bend, however, might bring the C-terminal domains 4 and 5 back together so as to facilitate the association of the transmembrane domains (*Fig. 8B*). There is experimental evidence that these transmembrane domains contribute to ICAM-1 dimerization (69, 70).

Electrostatics may optimally orient the binding surface of MAdCAM-1

MAdCAM-1 appears to have yet another way to orient its binding surface. Human MAdCAM-1 has, in addition to the two Ig-like domains, a mucin-like domain of 115 residues between the second Ig-like domain and the cell membrane (75) (*Fig. 8B*). This sequence contains 36 predicted O-linked glycosylation sites (76). Electron microscopic measurements of similarly heavily O-glycosylated, mucin-like proteins have shown them to assume an extended, unstructured, rod-like shape, with an extension on average of 1.8 to 2.0 Å/residue (77, 78). The mucin-like region of MAdCAM-1 should thus extend approximately 22 nm, and allow the integrin-binding IgSF domains to be positioned well above the cell surface (*Fig. 8B*). Most or all of the O-linked glycans should bear sialic acid. Fur-

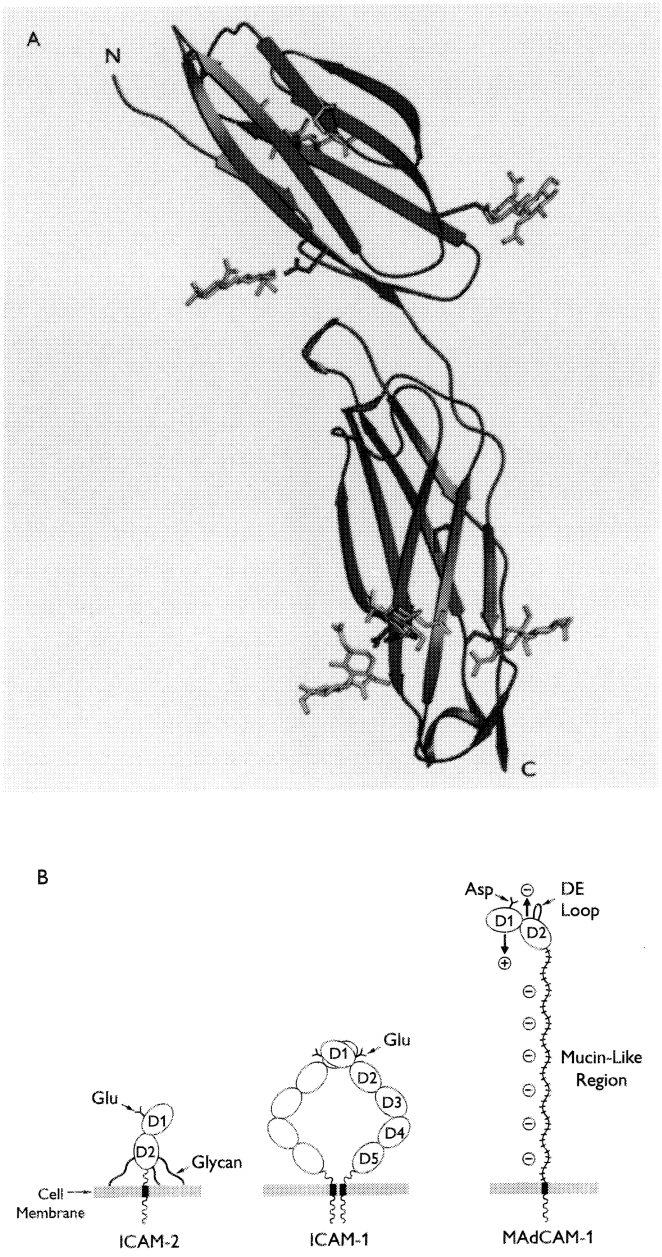

Fig. 8. Orientation of IgSF integrin ligands on cell surfaces.

A. ICAM-2. Six glycans are shown in yellow. Three glycans are uniformly distributed around the periphery of the bottom of domain 2, forming a tripod-like architecture. This may position domain 2 roughly vertical to the cell membrane. A relatively rigid bend between domains 1 and 2 maintains a tilt of domain 1 that is appropriate to expose Glu37 upwards for recognition by an integrin on an opposing cell. The figure was prepared with SETOR (103).

B. Cartoons drawn roughly to scale of ICAM-1, ICAM-2, and MAdCAM-1. The model of ICAM-1 is based on the dimerization seen in the X-ray crystal study (5), the bend between domains 3 and 4 visualized in the electron microscope (74), and the dimerization at the transmembrane region deduced from studies on intact cells (69, 70). Dimerization presents the critical Glu34 upward, and offers bivalent binding. The mucin-like region of MAdCAM-1 is predicted to extend the two Ig-like domains well above the cell membrane. The unique charge distribution, in particular the highly negatively charged DE β-ribbon of domain 2, is predicted to help optimally orient the integrin-binding region on the upper face toward an interacting cell.

thermore, the first segment of the mucin-like domain, 65 residues immediately following the second IgSF domain, contains 15 acidic residues and no basic residues. Since cell surfaces have a net negative charge, it is likely that repulsive electrostatic forces make the mucin-like fragment point away from and roughly perpendicular to the membrane to present integrin-binding IgSF domains. As described above, the most striking feature of the MAdCAM-1 structure is a highly negatively charged DE β-ribbon in domain 2. Seven of the eleven residues in this extension, from position 149 to 159, are Glu or Asp. We have proposed that this DE ribbon functions as a negatively charged "antenna", to orient the Ig domains. The coulombic field established by the cell surface and mucin-like region will repel the "antenna" on domain 2, and orient it facing away from the cell surface. This is predicted to optimally orient domain 2 and domain 1 such that the key integrin-binding residue Asp42 in the CD loop and the surrounding region of domain 1 and the DE β-ribbon in domain 2 both will be exposed for interaction with integrins on other cells (*Fig. 8B*) (10).

Orientation as a mechanism for preventing interactions in *cis*

Orientation of cell-surface molecules may be a mechanism not only for enhancing interactions in *trans* with other cells but also for preventing interactions in *cis* with neighboring molecules on the same cell. ICAM-1, ICAM-2, and ICAM-3 can all be expressed on leukocytes, and almost all leukocytes also express the integrin LFA-1. If interactions between ICAMs and LFA-1 on the same cell were possible, they would be highly favored entropically over interactions in *trans*. Such interactions do not appear to occur. For example, B-cell lines established from healthy individuals or LFA-1-deficient patients adhere to LFA-1 equally well. Similarly, treatment of cells that co-express LFA-1 and ICAM-1 with mAb to LFA-1 or ICAM-1 does not affect binding to substrates bearing LFA-1 or ICAM-1, respectively (79). If *cis* interactions occurred, blocking interactions in *cis* would be expected to enhance interactions in *trans*. In the absence of a complete structure for LFA-1, we do not know how its binding site is oriented. However, we predict that it too will be oriented directly away from the cell on which it is expressed. This then would prevent interaction in *cis*, which would require a sideways orientation of both binding sites, or an upward orientation of one and downward orientation of the other. The distance above the cell surface at which the binding sites are presented may also be important since, if binding sites are confined within different elevations above the surface, this could be another mechanism for preventing *cis* interactions.

The electrostatic mechanism we have proposed for orientation of MAdCAM-1 atop its mucin-like domain may be less precise than the architectural mechanisms for orienting ICAM-1 and ICAM-2. Thus, the electrostatic mechanism might be less effective in preventing interactions in *cis*. It is interesting that the expression of MAdCAM-1 in tissues is much more restricted than that of the ICAMs (4, 80). There are to date no cell types that co-express MAdCAM-1 and the integrin $\alpha4\beta7$, and therefore it may not be important to inhibit interactions in *cis*.

Discriminating between *cis* and *trans* interactions is particularly important for homophilic cell adhesion molecules, i.e. those that engage in like-like interactions. Cadherins are homophilic cell adhesion molecules for which the structure has revealed how this is accomplished (81, 82). These molecules are highly rigid. A Ca^{2+} ion is co-ordinated between each pair of adjacent cadherin domains, which have a somewhat Ig-like fold. The Ca^{2+} assures a rigid interdomain linkage. Cadherins can associate laterally on the same cell in *cis*, as well as in *trans* with molecules on other cells. However, the structures show that the two types of binding sites are distinct. Molecules must assume a parallel orientation to use the *cis* interaction site, and must assume an antiparallel orientation to use the *trans* binding site. Since molecules expressed on the same cell will have a largely parallel orientation, the *trans* type of interaction will be disfavored for molecules on the same cell.

IgSF cell adhesion molecules subverted as virus receptors

Ironically, some of the IgSF molecules that are important for immune defenses are subverted by viruses as receptors. The most notorious example is CD4, the receptor for HIV (83, 84). Another well-studied example is ICAM-1, a receptor for approximately 90 out of the 100 different serotypes of HRV (85, 86). HRV causes about 50% of common colds. HIV is an enveloped virus with its gp120/gp41 counter-receptor inserted in its lipid envelope. HRV has an icosahedral protein capsid, with a canyon or depression in the capsid around the 5-fold symmetry axis that contains the binding site for ICAM-1 (87). The attachment of a virus to its receptor is the first step in infection. The receptor-virus interaction has thus been an extensive research subject for pathogenesis of virus infection and an attractive target for antiviral drug design. The most advanced is work on influenza hemagglutinin binding to sialic acid (88). Here we compare two protein viral receptors: ICAM-1 and CD4. We show how the binding regions for physiological ligands and viruses differ, and speculate on how viruses may take advantage of some unique structural feature of these cell-surface IgSF proteins.

An extraordinarily protruded C'C" ridge of CD4 domain 1 as the gp120-binding motif

The HIV envelope glycoprotein consists of the extracellular gp120 subunit associated with a transmembrane gp41 subunit. The gp120 directs viron binding to the initial host cell surface receptor, CD4, with high affinity ($Kd = 10^{-9}$ M) and subsequent binding to chemokine co-receptors (19). The gp41 mediates viral-cell membrane fusion (89). The crystal structure of the N-terminal two-domain fragment of CD4 (30, 31) shows that, compared to a typical V set Ig domain, CD4 has an extraordinarily protruded C'C" loop and C" strand. This prominent feature has been termed the "C'C" ridge" (90). The unusually protruded C'C" ridge is supported at the top of the domain by the bulky side chain of Trp62, which sits in the middle of an α-helix inserted into the DE loop (Fig. 9D). Structure-guided mutants of CD4 have been examined for effect on gp120-binding equilibria and kinetics. These experiments delineated a binding region of about 900 Å^2 around the C'C" ridge. Phe43 is a hot spot (500-fold affinity reduction for F43A

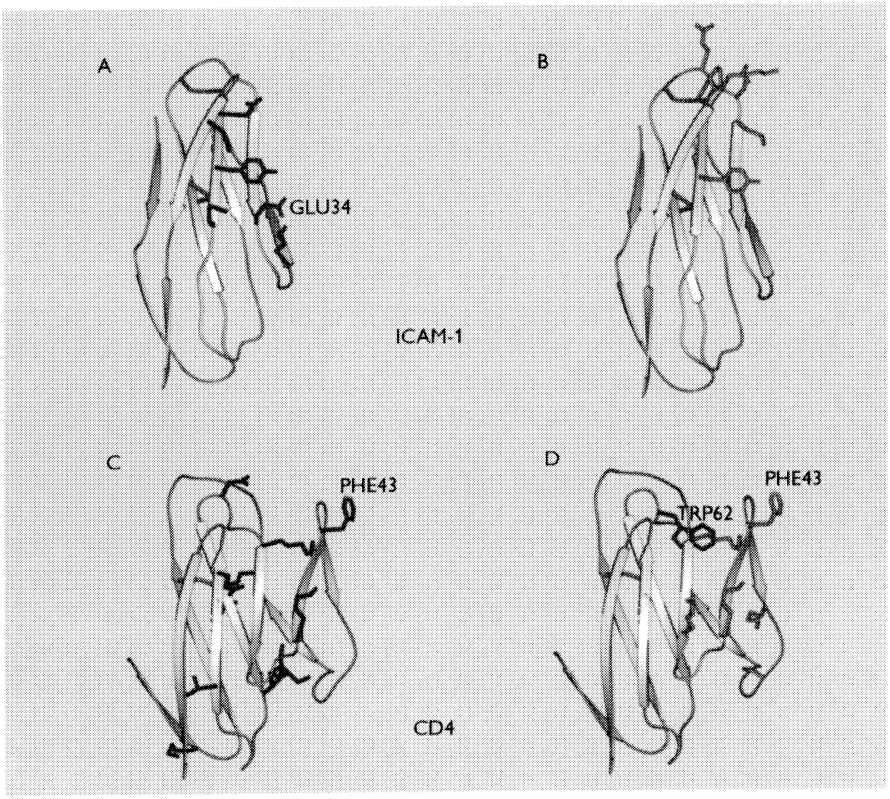

Fig. 9. Comparison of ligand-binding regions of ICAM-1 and CD4.
A. The integrin-binding sites of ICAM-1 as described in *Fig.* 6.
B. The HRV-binding sites of ICAM-1. Single amino acid substitutions are shown that decrease binding by HRV 3, 14, 15, 36, or 41 (35, 97, 98) by >50% or >2 s.d. These residues are on the top part of the molecule, in particular in the mobile BC and FG loops at the very tip of domain 1 just beyond the "extra" disulfide bond.
C. The MHC molecule-binding sites on CD4. The important residues are widely spread on the "right" side around the C'C" edge.
D. The HIV gp120-binding sites on CD4. The binding sites mostly overlap with the MHC-binding sites (91, 93). The "hot" spot Phe43 is labeled. Note that the Trp62 in black is buried. It sits on a short helical region, and its bulky side chain pushes the C'C" edge outward. The figures were prepared with SETOR (103).

or F43I mutants) on the upper, outer corner of the C'C" ridge (91). This has led to a prediction that the gp120 molecule should have a complementary groove for the ridge, and a specificity pocket within the groove to receive the phenylalanine side chain (92). Because gp120 is heavily shielded by glycans, an unusually protruded structure such as the C'C" ridge may be needed to reach into its binding groove. On CD4, the binding regions for gp120 and MHC class II molecules appear to overlap, but the contact region for MHC is suggested by mutagenesis to be much broader (*Fig.* 9 C & D) (93).

ICAM-1 binding to a surface depression of HRV

HRV is an icosahedral assembly of 60 copies of protein protomers and an RNA. Each protomer consists of three external subunits (VP1, VP2, and VP3) and one internal peptide (VP4). The crystal structure of HRV-14 has revealed that the VP subunits pack to create a surface depression 12 Å deep and 12–15 Å wide, encircling each pentagonal vertex (94). The residues that line this depression or "canyon" are more conserved than other surface residues among different serotypes; therefore, it has been proposed to be the receptor-binding site (94, 95) and experimental evidence has emerged in support of this concept (87, 96). There are over 100 HRV serotypes, and ICAM-1 serves as the receptor for approximately 90% of these (85, 86). Cryoelectron microscopic images of the complex between HRV and an N-terminal two-domain fragment of ICAM-1 clearly show virus particles decorated with 60 small projections (87). The difference map between the density from electron microscopy image reconstruction and a low-resolution X-ray structure suggests that an Ig-like two-domain fragment could fit into this projection, and localized the binding site to a specific region of the HRV canyon. Ironically, it was the structure of CD4, which has a markedly larger N-terminal Ig domain than ICAM-1, that was originally used as the model for docking to the canyon. The canyon is also appropriate for docking of ICAM-1, and approximately the upper half of domain 1 could fit into the canyon (6). A dimer of ICAM-1 could dock to the proposed binding site in the canyon through one of its monomers, because the upper half of domain 1 is unobscured in the dimer (5).

The residues important for binding to HRV have been mapped with five different HRV serotypes by mutagenesis experiments (35, 97, 98). Six of these residues map to BC and FG loops on the tip of domain 1, and one maps part-way down the side of domain 1 in the F strand (*Fig.* 9*B*). The binding regions for LFA-1 and virus are quite different, although at least one residue overlaps both sites (*Fig.* 9 *A* & B). The "extra" disulfide bond tethers the BC loop to the FG loop. Interestingly, the

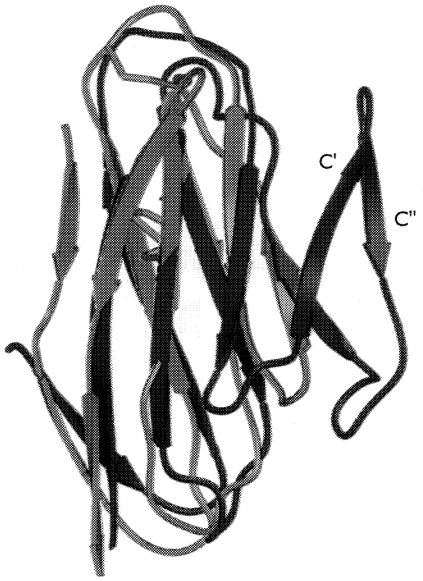

Fig. 10. Superposition of domains 1 of ICAM-1 and CD4. The striking distinction in shape between ICAM-1 (green) and CD4 (red) is apparent. The critical gp120-binding region is concentrated on the extraordinarily protruded C'C" ridge of the fat domain 1 of CD4, whereas the HRV-binding region is concentrated at the tip of the skinny domain 1 of ICAM-1. The molecules are shown in the same orientation as in *Fig.* 9. The figure was prepared with SETOR (103).

loop residues important in HRV binding are located just beyond the "extra" disulfide bond (see *Fig. 9B*), and the BC and FG loops are remarkably flexible in this region. Some well-defined residues in these loops have different conformations between two copies of molecule in the crystal structure, and overall the loops have a high B factor (5). It is interesting that tests on different HRV serotypes suggest significant differences between serotypes in which residues on ICAM-1 are important for binding (98). It is possible that the conformation of the BC and FG loops differs when bound to different HRV serotypes, resulting in differences in the contact residues on ICAM-1.

One of the biggest questions with HRV is how so many different serotypes, which differ markedly in capsid amino acid sequence, can bind the same receptor. If this receptor-binding site were conserved, it might be expected that antibodies could be elicited that would crossreact between serotypes and neutralize infectivity. One possibility to explain the lack of such antibodies was raised by the observation of the canyon in the HRV X-ray structure. It was proposed that the receptor-binding site was in a depression that was too narrow to admit an antibody (94). However, cryoelectron microscopy of an antibody Fab bound to HRV has shown that the antibody footprint extends into the canyon (99). This points out that the important feature is not that the receptor bind in a surface depression, but that it be smaller than an antibody. The HRV-binding tip of ICAM-1 has a markedly smaller footprint than an Fab fragment. As long as the receptor is smaller than the Fab, mutations can arise that are outside and adjacent to the receptor-binding site, but within the footprint of neutralizing antibodies. In this way, the HRV can mutate and escape blockade of the receptor-binding site by antibodies. With further mutations elsewhere in the capsid proteins, large numbers of non-crossreactive serotypes could emerge that all bind the same receptor. The structure of two independent ICAM-1 molecules presents yet another scenario for how a large number of serotypes could emerge (5). Because of the flexibility of the BC and FG loops, which are at the center of the HRV-binding site on ICAM-1, it is quite conceivable that different HRV serotypes could bind different conformations of these loops, and thus differ in amino acid sequence even near the center of the receptor-binding site. This would provide yet another mechanism for evasion of immune responses.

Further analysis of cryoelectron microscopy images using the newly available ICAM-1 electron density maps (5, 6), and X-ray structures of ICAM-1 HRV complexes, would shed light on the interesting relationships between ICAM-1 and the evolution of such a large number of HRV serotypes. ICAM-1 binding triggers HRV disruption (16–18). It would be interesting to learn whether the flexible BC and FG loops adapt to conformational changes in the virus capsid during uncoating, and whether ICAM-1 dimerization plays any role in the uncoating process.

Domains 1 of CD4 and ICAM-1 can be readily superimposed based on the alignment of their intersheet disulfide bonds and surrounding residues (*Fig.* 10). It is striking to see how the whole C' strand, C'C" loop, and C" strand are missing in ICAM-1. Receptor binding by HIV and HRV appears to take advantage of different topological features of these two cell-surface adhesion molecules.

References

1. Springer TA. Traffic signals for lymphocyte recirculation and leukocyte emigration: the multi-step paradigm.
Cell 1994;**76**:301–314.
2. Chothia C, Jones EY. The molecular structure of cell adhesion molecules.
Annu Rev Biochem 1997;**66**:823–862.
3. Tian L, Yoshihara Y, Mizuno T, Mori K, Gahmberg CG. The neuronal glycoprotein telencephalin is a cellular ligand for the CD11a/CD18 leukocyte integrin.
J Immunol 1997;**158**:928–936.
4. Dustin ML, Springer TA. Intercellular adhesion molecules (ICAMs). In: Kreis T, Vale R, eds. Guidebook to the extracellular matrix and adhesion proteins. New York: Sambrook and Tooze; 1998 (In press).
5. Casasnovas JM, Stehle T, Liu J-h, Wang J-h, Springer TA. A dimeric crystal structure for the N-terminal two domains of ICAM-1.
Proc Natl Acad Sci USA 1998;**95**:4134–4139.
6. Bella J, Kolatkar PR, Marlor C, Greve JM, Rossmann MG. The structure of the two amino-terminal domains of human ICAM-1 suggests how it functions as a rhinovirus receptor and as an LFA-1 integrin ligand.
Proc Natl Acad Sci USA 1998;**95**:4140–4145.
7. Casasnovas JM, Springer TA, Liu J-h, Harrison SC, Wang J-h. The crystal structure of ICAM-2 reveals a distinctive integrin recognition surface.
Nature 1997;**387**:312–315.
8. Jones EY, et al. Crystal structure of an integrin-binding fragment of vascular cell adhesion molecule-1 at 1.8 Å resolution.
Nature 1995;**373**:539–544.
9. Wang J-h, et al. The crystal structure of an N-terminal two domain fragment of VCAM-1: a cyclic peptide based on the domain 1 C-D loop can inhibit VCAM-1/α4 integrin.
Proc Natl Acad Sci USA 1995;**92**:5714–5718.
10. Tan K, Casasnovas JM, Liu J-h, Briskin MJ, Springer TA, Wang J-h. The structure of MAdCAM-1 reveals novel features in domains 1 and 2 important for integrin recognition.
Structure 1998;**6**:793–801.
11. Sethi SK, Bianco A, Allen JT, Knight RA, Spiteri MA. Interferon-gamma (IFN-gamma) down-regulates the rhinovirus-induced expression of intercellular adhesion molecule-1 (ICAM-1) on human airway epithelial cells.
Clin Exp Immunol 1997;**110**:362–369.
12. Springer TA. Adhesion receptors of the immune system.
Nature 1990;**346**:425–433.
13. Diamond MS, Springer TA. The dynamic regulation of integrin adhesiveness.
Curr Biol 1994;**4**:506–517.
14. Schwartz MA, Schaller MD, Ginsberg MH. Integrins: emerging paradigms of signal transduction.
Annu Rev Cell Dev Biol 1995;**11**:549–599.
15. Alon R, Chen S, Puri KD, Finger EB, Springer TA. The kinetics of L-selectin tethers and the mechanics of selectin-mediated rolling.
J Cell Biol 1997;**138**:1169–1180.
16. Greve JM, et al. Mechanisms of receptor-mediated rhinovirus neutralization defined by two soluble forms of ICAM-1.
J Virol 1991;**65**:6015–6023.
17. Hoover-Litty H, Greve JM. Formation of rhinovirus-soluble ICAM-1 complexes and conformational changes in the virion.
J Virol 1993;**67**:390–397.
18. Martin S, Casasnovas JM, Staunton DE, Springer TA. Efficient neutralization and disruption of rhinovirus by chimeric ICAM-1/immunoglobulin molecules.
J Virol 1993;**67**:3561–3568.
19. Wilkinson D. Cofactors provide the entry keys.
Curr Biol 1996;**6**:1051–1053.
20. Bork P, Downing AK, Kieffer B, Campbell ID. Structure and distribution of modules in extracellular proteins.
Q Rev Biophys 1996;**29**:119–167.
21. Barclay AN, et al. The leucocyte antigen facts book. London: Academic Press; 1993.
22. Parker MJ, Dempsey CE, Hosszu LLP, Waltho JP, Clarke AR. Topology, sequence evolution and folding dynamics of an immunoglobulin domain.
Nat Struct Biol 1998;**5**:194–198.
23. Chothia C, Janin J. Relative orientation of close-packed β-pleated sheets in proteins.
Proc Natl Acad Sci USA 1981;**78**:4146–4150.
24. Chothia C, Novotny J, Bruccoleri R, Karplus M. Domain association in immunoglobulin molecules. The packing of variable domains.
J Mol Biol 1985;**186**:651–663.
25. Miller S. Protein-protein recognition and the association of immunoglobulin constant domains.
J Mol Biol 1990;**216**:965–973.
26. Harpaz Y, Chothia C. 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 1994;**238**:528–539.
27. Jones EY, Davis SJ, Williams AF, Harlos K, Stuart DI. Crystal structure at 2.8 Å resolution of a soluble form of the cell adhesion molecule CD2.
Nature 1992;**30**:232–239.
28. Wang J-h, Stehle T, Pepinsky B, Liu J-h, Karpusas M, Osborn L. Structure of a functional fragment of VCAM-1 refined at 1.9 angstrom resolution.
Acta Crystallogr 1996;**D52**:369–379.
29. Wu H, Kwong PD, Hendrickson WA. Dimeric association and segmental variability in the structure of human CD4.
Nature 1997;**387**:527–530.
30. Wang J, et al. Atomic structure of a fragment of human CD4 containing two immunoglobulin-like domains.
Nature 1990;**348**:411–418.
31. Ryu SE, et al. Crystal structure of an HIV-binding recombinant fragment of human CD4.
Nature 1990;**348**:419–426.
32. Brady RL, et al. Crystal structure of domains 3 and 4 of rat CD4: relation to the NH_2-terminal domains.
Science 1993;**260**:979–983.
33. McCalden P, Argos P. Oligopeptide biases in protein sequences and their use in predicting protein coding regions in nucleotide sequences.
Proteins 1988;**4**:99–122.
34. Chothia C, Lesk AM. Canonical structures for the hypervariable regions of immunoglobulins.
J Mol Biol 1987;**196**:901–917.
35. Staunton DE, Dustin ML, Erickson HP, Springer TA. The arrangement of the immunoglobulin-like domains of ICAM-1 and the binding sites for LFA-1 and rhinovirus.
Cell 1990;**61**:243–254.
36. Holness CL,et al. Analysis of the binding site on intercellular adhesion molecule 3 for the leukocyte integrin lymphocyte function-associated antigen 1.
J Biol Chem 1995;**270**:877–884.
37. Klickstein LB, York MB, de Fougerolles AR, Springer TA. Localization of the binding site on intercellular adhesion molecule-3 (ICAM-3) for lymphocyte function-associated antigen-1 (LFA-1).
J Biol Chem 1996;**271**:23920–23927.
38. Osborn L, et al. Arrangement of domains, and amino acid residues required for binding of vascular cell adhesion molecule-1 to its counter-receptor VLA-4 (α4β1).
J Cell Biol 1994;**124**:601–608.
39. Vonderheide RH, Tedder TF, Springer TA, Staunton DE. Residues within a conserved amino acid motif of domains 1 and 4 of VCAM-1 are required for binding to VLA-4.
J Cell Biol 1994;**125**:215–222.
40. Viney JL, et al. Mucosal addressin cell adhesion molecule-1. A structural and functional analysis demarcates the integrin binding motif.
J Immunol 1996;**157**:2488–2497.

41. Borders CL Jr, et al. A structural role for arginine in proteins: Multiple hydrogen bonds to backbone carbonyl oxygens. Protein Sci 1994;**3**:541–548.
42. Hemler ME, Elices MJ, Parker C, Takada Y. Structure of the integrin VLA-4 and its cell-cell and cell-matrix adhesion functions. Immunol Rev 1990;**114**:45–65.
43. Loftus JC, Smith JW, Ginsberg MH. Integrin mediated cell adhesion: the extracellular face. J Biol Chem 1994;**269**:25235–25238.
44. Springer TA. Folding of the N-terminal, ligand-binding region of integrin α-subunits into a β-propeller domain. Proc Natl Acad Sci USA 1997;**94**:65–72.
45. Oxvig C, Springer TA. Experimental support for a β-propeller domain in integrin α-subunits and a calcium binding site on its lower surface. Proc Natl Acad Sci USA (Submitted).
46. Diamond MS, Garcia-Aguilar J, Bickford JK, Corbi AL, Springer TA. The I domain is a major recognition site on the leukocyte integrin Mac-1 (CD11b/CD18) for four distinct adhesion ligands. J Cell Biol 1993;**120**:1031–1043.
47. Michishita M, Videm V, Arnaout MA. A novel divalent cation-binding site in the A domain of the β2 integrin CR3 (CD11b/CD18) is essential for ligand binding. Cell 1993;**72**:857–867.
48. Lee J-O, Rieu P, Arnaout MA, Liddington R. Crystal structure of the A domain from the α subunit of integrin CR3 (CD11b/CD18). Cell 1995;**80**:631–638.
49. Qu A, Leahy DJ. Crystal structure of the I-domain from the CD11a/CD18 (LFA-1, $\alpha_L\beta 2$) integrin. Proc Natl Acad Sci USA 1995;**92**: 10277–10281.
50. Huang C, Springer TA. 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 1995;**270**:19008–19016.
51. Newham P, et al. Alpha4 integrin binding interfaces on VCAM-1 and MAdCAM-1. Integrin binding footprints identify accessory binding sites that play a role in integrin specificity. J Biol Chem 1997;**272**:19429–19440.
52. Edwards YJK, Perkins SJ. Assessment of protein fold predictions from sequence information: the predicted α/β doubly wound fold of the von Willebrand Factor type A domain is similar to its crystal structure. J Mol Biol 1996;**260**:277–285.
53. Murzin AG, Brenner SE, Hubbard T, Chothia C. SCOP: a structural classification of proteins database for the investigation of sequences and structures. J Mol Biol 1995;**247**:536–540.
54. Lee J-O, Bankston LA, Arnaout MA, Liddington RC. Two conformations of the integrin A-domain (I-domain): a pathway for activation? Structure 1995;**3**:1333–1340.
55. Qu A, Leahy DJ. The role of the divalent cation in the structure of the I domain from the CD11a/CD18 integrin. Structure 1996;**4**:931–942.
56. Jackson DY, et al. Potent alpha 4 beta 1 peptide antagonists as potential anti-inflammatory agents. J Med Chem 1997;**40**:3359–3368.
57. Vanderslice P, et al. A cyclic hexapeptide is a potent antagonist of α_4 integrins. J Immunol 1997;**158**:1710–1718.
58. Irie A, Kamata T, Takada Y. Multiple loop structures critical for ligand binding of the integrin α4 subunit in the upper face of the β-propeller models. Proc Natl Acad Sci USA 1997;**94**:7198–7203.
59. Renz ME, et al. Structural requirements for adhesion of soluble recombinant murine vascular cell adhesion molecule-1 to α4β1. J Cell Biol 1994;**125**:1395–1406.
60. Chiu HH, et al. Similar but nonidentical amino acid residues on vascular cell adhesion molecule-1 are involved in the interaction with α4β1 and α4β7 under different activity states. J Immunol 1995;**155**:5257–5267.
61. Padlan EA. Anatomy of the antibody molecule. Mol Immunol 1994;**31**:169–217.
62. Garcia KC, et al. An αβ T cell receptor structure at 2.5 angstrom and its orientation in the TCR-MHC complex. Science 1996;**274**:209–219.
63. Garboczi DN, Ghosh P, Utz U, Fan QR, Biddison WE, Wiley DC. Structure of the complex between human T-cell receptor, viral peptide and HLA-A2. Nature 1996;**384**:134–141.
64. Wang JH, et al. Atomic structure of an alpha beta T cell receptor (TCR) heterodimer in complex with an anti-TCR Fab fragment derived from a mitogenic antibody. EMBO J 1998;**17**:10–26.
65. Van der Merwe PA, McNamee PN, Davies EA, Barclay AN, Davis SJ. Topology of the CD2-CD48 cell-adhesion molecule complex: implications for antigen recognition by T cells. Curr Biol 1995;**5**:74–84.
66. Fan QR, Mosyak L, Winter CC, Wagtmann N, Long EO, Wiley DC. Structure of the inhibitory receptor for human natural killer cells resembles haematopoietic receptors. Nature 1997;**389**:96–100.
67. De Fougerolles AR, Stacker SA, Schwarting R, Springer TA. Characterization of ICAM-2 and evidence for a third counter-receptor for LFA-1. J Exp Med 1991;**174**:253–267.
68. Maddon PJ, et al. Structure and expression of the human and mouse T4 genes. Proc Natl Acad Sci USA 1987;**84**:9155–9159.
69. Miller J, Knorr R, Ferrone M, Houdei R, Carron CP, Dustin ML. Intercellular adhesion molecule-1 dimerization and its consequences for adhesion mediated by lymphocyte function associated-1. J Exp Med 1995;**182**:1231–1241.
70. Reilly PL, Woska JR Jr, Jeanfavre DD, McNally E, Rothlein R, Bormann B-J. The native structure of intercellular adhesion molecule-1 (ICAM-1) is a dimer: correlation with binding to LFA-1. J Immunol 1995;**155**:529–532.
71. Weber C, Lu C-F, Casasnovas JM, Springer TA. Role of $\alpha_L\beta_2$ integrin avidity in transendothelial chemotaxis of mononuclear cells. J Immunol 1997;**159**:3968–3975.
72. Li H, Lebedeva MI, Liera AS, Fields BA, Brenner MB, Mariuzza RA. Structure of the Vδ domain of a human γδ T-cell antigen receptor. Nature 1998;**391**:502–506.
73. Bentley GA, Boulot G, Karjalainen K, Mariuzza RA. Crystal structure of the β chain of a T cell antigen receptor. Science 1995;**267**:1984–1987.
74. Kirchhausen T, Staunton DE, Springer TA. Location of the domains of ICAM-1 by immunolabeling and single-molecule electron microscopy. J Leukoc Biol 1993;**53**:342–346.
75. Shyjan AM, Bertagnolli M, Kenney CJ, Briskin MJ. Human mucosal addressin cell adhesion molecule-1 (MAdCAM-1) demonstrates structural and functional similarities to the a4b7-integrin binding domains of murine MAdCAM-1, but extreme divergence of mucin-like sequences. J Immunol 1996;**156**:2851–2857.
76. Hansen JE, Lund O, Engelbrecht J, Bohr H, Nielsen JO. Prediction of O-glycosylation of mammalian proteins: specificity patterns of UDP-GalNAc: polypeptide N-acetylgalactosaminyltransferase. Biochem J 1995;**308**:801–813.
77. Cyster JG, Shotton DM, Williams AF. The dimensions of the T lymphocyte glycoprotein leukosialin and identification of linear protein epitopes that can be modified by glycosylation. EMBO J 1991;**10**:893–902.
78. Li F, Erickson HP, James JA, Moore KL, Cummings RD, McEver RP. Visualization of P-selectin glycoprotein ligand-1 as a highly extended molecule and mapping of protein epitopes for monoclonal antibodies. J Biol Chem 1996;**271**:6342–6348.

79. Marlin SD, Springer TA. Purified intercellular adhesion molecule-1 (ICAM-1) is a ligand for lymphocyte function-associated antigen 1 (LFA-1).
Cell 1987;**51**:813–819.
80. Briskin M, et al. Human mucosal addressin cell adhesion molecule-1 is preferentially expressed in intestinal tract and associated lymphoid tissue.
Am J Pathol 1997;**151**:97–110.
81. Shapiro L, et al. Structural basis of cell-cell adhesion by cadherins.
Nature 1995;**374**:327–337.
82. Nagar B, Overduin M, Ikura M, Rini JM. Structural basis of calcium-induced E-cadherin rigidification and dimerization.
Nature 1996;**380**:360–364.
83. Dalgleish AG, Beverley PCL, Clapham PR, Crawford DH, Greaves MF, Weiss RA. The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus.
Nature 1984;**312**:763–767.
84. Klatzmann D, et al. T-lymphocyte T4 molecule behaves as the receptor for human retrovirus LAV.
Nature 1984;**312**:767–768.
85. Greve JM, et al. The major human rhinovirus receptor is ICAM-1.
Cell 1989;**56**:839–847.
86. Staunton DE, Merluzzi VJ, Rothlein R, Barton R, Marlin SD, Springer TA. A cell adhesion molecule, ICAM-1, is the major surface receptor for rhinoviruses.
Cell 1989;**56**:849–853.
87. Olson NH, et al. Structure of human rhinovirus complexed with its receptor molecule.
Proc Natl Acad Sci USA 1993;**90**:507–511.
88. Eisen MB, Sabesan S, Skehel JJ, Wiley DC. Binding of the influenza A virus to cell-surface receptors: structures of five hemagglutinin-sialyloligosaccharide complexes determined by X ray crystallography.
Virology 1997;**232**:19–31.
89. Gallaher WR. Detection of a fusion peptide sequence in the transmembrane protein of human immunodeficiency virus.
Cell 1987;**50**:327–328.
90. Harrison SC, et al. Structure and interactions of CD4.
Cold Spring Harb Symp Quant Biol 1992;**57**: 541–548.
91. Moebius U, Clayton LK, Abraham S, Harrison SC, Reinherz EL. The human immunodeficiency virus gp120 binding site on CD4: delineation by quantitative equilibrium and kinetic binding studies of mutants in conjunction with a high-resolution CD4 atomic structure.
J Exp Med 1992;**176**:507–517.
92. Harrison SC. CD4: Structure and interactions of an immunoglobulin superfamily adhesion molecule.
Acc Chem Res 1993;**26**: 449–453.
93. Moebius U, et al. Human immunodeficiency virus gp120 binding C'C" ridge of CD4 domain 1 is also involved in interaction with class II major histocompatibility complex molecules.
Proc Natl Acad Sci USA 1992;**89**: 12008–12012.
94. Rossmann MG. The Canyon Hypothesis: Hiding the host cell receptor attachment site on a viral surface from immune surveillance.
J Biol Chem 1989;**264**:14587–14590.
95. Rossmann MG, Palmenberg AC. Conservation of the putative receptor attachment site in picornaviruses.
Virology 1988;**164**:373–382.
96. Colonno RJ, Condra JH, Mizutani S, Callahan PL, Davies M-E, Murcko MA. Evidence for the direct involvement of the rhinovirus canyon in receptor binding.
Proc Natl Acad Sci USA 1988;**85**:5449–5453.
97. McClelland A, DeBear J, Yost SC, Meyer AM, Marlor CW, Greve JM. Identification of monoclonal antibody epitopes and critical residues for rhinovirus binding in domain 1 of intercellular adhesion molecule 1.
Proc Natl Acad Sci USA 1991;**88**:7993–7997.
98. Register RB, Uncapher CR, Naylor AM, Lineberger DW, Colonno RJ. Human-murine chimeras of ICAM-1 identify amino acid residues critical for rhinovirus and antibody binding.
J Virol 1991;**65**:6589–6596.
99. Smith TJ, et al. Structure of human rhinovirus complexed with Fab fragments from a neutralizing antibody.
J Virol 1993;**67**:1148–1158.
100. Kabsch W, Sander C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features.
Biopolymers 1983;**22**:2577–2637.
101. Kraulis P. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structure.
J Appl Crystallogr 1991;**24**:946–950.
102. Sali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints.
J Mol Biol 1993;**234**:779–815.
103. Evans SV. SETOR: hardware lighted three-dimensional solid model representations of macromolecules.
J Mol Graph 1993;**11**:134–138.
104. Fisher KL, Lu J, Riddle L, Kim KJ, Presta LG, Bodary SC. Identification of the binding site in intercellular adhesion molecule 1 for its receptor, leukocyte function-associated antigen 1.
Mol Biol Cell 1997;**8**:501–515.