



Published in final edited form as:

*Science*. 2001 October 12; 294(5541): 339–345. doi:10.1126/science.1064535.

## Crystal Structure of the Extracellular Segment of Integrin $\alpha V\beta 3$

Jian-Ping Xiong<sup>1</sup>, Thilo Stehle<sup>1,2</sup>, Beate Diefenbach<sup>3</sup>, Rongguang Zhang<sup>4</sup>, Reinhardt Dunker<sup>3</sup>, David L. Scott<sup>1</sup>, Andrzej Joachimiak<sup>4</sup>, Simon L. Goodman<sup>3</sup>, and M. Amin Arnaout<sup>1,\*</sup>

<sup>1</sup>Renal Unit, Leukocyte Biology & Inflammation Program, Structural Biology Program, Massachusetts General Hospital and Harvard Medical School, 149 13th Street, Charlestown, MA 02129, USA.

<sup>2</sup>Laboratory of Developmental Immunology, Massachusetts General Hospital and Harvard Medical School, 55 Fruit Street, Boston, MA 02114, USA.

<sup>3</sup>Departments of Biotechnology and Biomedical Research Immunology/Oncology, Merck KGaA, Darmstadt 64271, Germany.

<sup>4</sup>Biosciences Division, Argonne National Laboratory, IL 60439, USA

### Abstract

Integrins are  $\alpha\beta$  heterodimeric receptors that mediate divalent cation-dependent cell-cell and cell-matrix adhesion through tightly regulated interactions with ligands. We have solved the crystal structure of the extracellular portion of integrin  $\alpha V\beta 3$  at 3.1 Å resolution. Its 12 domains assemble into an ovoid “head” and two “tails.” In the crystal,  $\alpha V\beta 3$  is severely bent at a defined region in its tails, reflecting an unusual flexibility that may be linked to integrin regulation. The main inter-subunit interface lies within the head, between a seven-bladed  $\beta$ -propeller from  $\alpha V$  and an A domain from  $\beta 3$ , and bears a striking resemblance to the  $G\alpha/G\beta$  interface in G proteins. A metal ion-dependent adhesion site (MIDAS) in the  $\beta A$  domain is positioned to participate in a ligand-binding interface formed of loops from the propeller and  $\beta A$  domains. MIDAS lies adjacent to a calcium-binding site with a potential regulatory function.

---

Integrins are large heterodimeric cell surface receptors found in many animal species ranging from sponges to mammals [reviewed in (1)]. These receptors are involved in fundamental cellular processes such as attachment, migration, proliferation, differentiation, and survival. Integrins also contribute to the initiation and/or progression of many common diseases including neoplasia, tumor metastasis, immune dysfunction, ischemia-reperfusion injury, viral infections, osteoporosis, and coagulopathies [reviewed in (2,3)]. An integrin is ~280 Å long and consists of one  $\alpha$  (150 to 180 kD) and one  $\beta$  (~90 kD) subunit, both of which are type I membrane proteins. Eighteen  $\alpha$  and eight  $\beta$  mammalian subunits are known, which assemble noncovalently into 24 different heterodimers. Contacts between the  $\alpha$  and  $\beta$  subunits primarily involve their  $NH_2$ -terminal halves [reviewed in (1)], which together form a globular head; the remaining portions form two rod-shaped tails (4–7) that span the plasma membrane.

Like other receptors, integrins transmit signals to the cell interior (so-called “outside-in” signaling), which regulate organization of the cytoskeleton, activate kinase signaling cascades, and modulate the cell cycle and gene expression [reviewed in (8)]. Unlike other receptors, however, ligand binding with integrins is not generally constitutive but is regulated to reflect the activation state of the cell. This “inside-out” regulation of integrin affinity protects the host

---

\*To whom correspondence should be addressed. arnaout@receptor.mgh.harvard.edu.

from pathological integrin-mediated adhesion (2). Inside-out and outside-in signaling are associated with distinct conformational changes in the integrin extracellular segment. These changes vary with cell type and the state and nature of the ligand, and are modulated by divalent cations that are also required for integrinligand interaction (9–11). The structural bases of activation and regulation of integrins are unknown. Previous attempts to address this question were limited to the structure determination of the ligand-binding, ~180 –amino acid A-type domain ( $\alpha$ A) (12) present in one-half of the integrin  $\alpha$  subunits (13). The structural basis of ligand binding to  $\alpha$ A-lacking integrins and its regulation are unknown. One such integrin is  $\alpha$ V $\beta$ 3 (CD51/CD61), a receptor important in tumor angiogenesis and metastasis, inflammation, and bone resorption [reviewed in (14)].  $\alpha$ V $\beta$ 3 is one of the most promiscuous integrins, as it binds to multiple ligands including vitronectin, angiostatin, and osteopontin, and also serves as a receptor for several viruses such as foot-and-mouth disease virus, adenovirus, and human immunodeficiency virus [reviewed in (15)]. Here we report the crystal structure of the extracellular region of human  $\alpha$ V $\beta$ 3. The structure has both known and previously unknown domains with creative designs, and reveals unexpected quaternary arrangements that provide insights into integrin function and regulation. These studies also have major bearing on development of drugs to control angiogenesis, inflammation, viruses, and osteoporosis by targeting integrins.

## Structure Determination and Refinement

Expression and purification of a soluble  $\alpha$ V $\beta$ 3 fragment, which terminates after Pro<sup>957</sup> (in  $\alpha$ V) and Asp<sup>692</sup> (in  $\beta$ 3) and therefore lacks the transmembrane and short cytoplasmic tails of both subunits, was carried out essentially as described (16). Soluble  $\alpha$ V $\beta$ 3 displayed the same epitopes, ligand-binding specificities, and divalent cation requirements as the active form of native  $\alpha$ V $\beta$ 3 [(16) and data not shown]. Protein crystals were grown by vapor diffusion using 100 mM MES, pH 6.0, 100 mM NaCl, 5 mM CaCl<sub>2</sub>, and 10% polyethylene glycol 4000 as precipitant. The structure was solved with the use of single isomorphous replacement-anomalous scattering (SIRAS) and multiwavelength anomalous diffraction (MAD) (Table 1). The final model contains all the extracellular residues with the exception of  $\alpha$ V residues 839 – 867 and 957, and  $\beta$ 3 residues 1–54, 435–531, and 691–692. The final model contains 58 cysteines, all of which are paired. It also contains 6 calcium ions and 11 glycans [with either one or two *N*-acetyl glucosamine (Glc-NAc) modeled at each site]. The final crystallographic *R* factors between 20 and 3.1 Å are 25.5% (working set) and 33.5% (free set). Refinement statistics are listed in Table 1. Figure 1A shows a representative electron density map. The secondary structure assignment and an alignment with integrins  $\alpha$ 5,  $\alpha$ IIB, and  $\beta$ 1 are given in Web fig. 1 (17).

## Overall Structure of the Heterodimer

The NH<sub>2</sub>-terminal segments of the  $\alpha$  and  $\beta$  subunits assemble into an ovoid “head” from which two nearly parallel “tails” emerge (Fig. 1, B and C). The head consists of a seven-bladed  $\beta$ -propeller from  $\alpha$ V and a  $\beta$ A domain looping out from a unique immunoglobulin (Ig)-like “hybrid” domain in  $\beta$ 3. The integrin head has dimensions of ~90 Å by 60 Å by 45 Å. The  $\alpha$ V tail is composed of three  $\beta$ -sandwich domains: an Ig-like “thigh” domain and two very similar domains that form the “calf” module. The  $\beta$ 3 tail consists of a PSI [for plexins, semaphorins, and integrins (18)] domain, four epidermal growth factor (EGF) domains, and a  $\beta$ -tail domain ( $\beta$ TD). The  $\alpha$  and  $\beta$  integrin tails fold back at a ~135° angle, forming a V-shaped structure with a “genu” between the thigh domain and the calf module of  $\alpha$ V. This profound bending of both subunits is surprising but unambiguous. When extended, each integrin tail forms a thin cylinder, ~160 Å long and ~20 Å in diameter. These values are in agreement with those derived from rotary shadowing images (4–7).

## Domain Architecture and Interdomain Interfaces $\alpha V$ Subunit Domains

### $\beta$ -propeller domain

The  $\beta$ -propeller is formed from the NH<sub>2</sub>-terminal seven-fold ~60 residue sequence repeats of  $\alpha V$  and consists of seven radially arranged “blades,” each formed from a four-stranded antiparallel sheet (Fig. 2A). The inner strand (strand A) of each blade lines the channel at the center of the propeller, with strands B and C of the same repeat radiating outward, and strand D of the next repeat forming the outer edge of the blade [Fig. 2A and Web fig. 2 (17)]. The  $\beta$ -propeller is circularized by juxtaposition of the C7 and D1 strands. The central channel is lined predominantly with amide and carbonyl oxygen groups from the seven A strands, with only a few side chains projecting into the cavity.

Superposition of the seven integrin blades reveals a unique consensus sequence repeat, the “cage” motif [Fig. 2, B through D, and Web fig. 2 (17)]. This motif ( $X_{17-33}\{\phi\phi G\phi X_{13-20}PX_{2-15}GX_{5-8}\}_7$ , where X is any residue and  $\phi$  is an aromatic residue) is defined by a mostly aromatic four-residue “cup” ( $\phi\phi G\phi$ ) that precedes strand A, a proline immediately following strand B (Pro-B), and a glycine at the beginning of strand C (Gly-C). The first and fourth residues of the cup form the upper and lower rings of a cage-like structure that lies at the center of the contact region between  $\alpha V$  and  $\beta 3$ , holding Arg<sup>261</sup> of  $\beta 3$  in place (Fig. 1A and Fig. 2, C and D). The second and third residues form the floor of the cup. Pro-B introduces an acute turn and is tightly sandwiched between sequential cups, providing a rigid scaffold for the upper and lower ring residues. Pro-B is replaced with Phe<sup>427</sup> in blade 7, pushing the side chains of the lower ring residues Phe<sup>21</sup> and Tyr<sup>406</sup> toward Arg<sup>261</sup> (Fig. 2C). The small side chains of the lower ring residues Ala<sup>96</sup> and Ser<sup>342</sup> are required to avoid clashes in this region. Gly-C is required because any side chain at this position would interfere with Pro-B.

Four solvent-exposed Ca<sup>2+</sup>-binding sites are found in the A-B  $\beta$  hairpin loops of blades 4-7 at the propeller's bottom (opposite the  $\alpha\beta$  interface) [see metal coordination in Web fig. 3 (17)]. The calcium-binding residues span a nine-residue segment with the consensus sequence Asp-h-Asp/Asn-x-Asp/Asn-Gly-h-x-Asp, where “h” is hydrophobic and “x” is any residue; the invariant Gly is required because of its unusual main chain torsion angles. Ca<sup>2+</sup> is usually coordinated by oxygen atoms from side chains of residues 1, 3, 5, and 9 and the carbonyl oxygen of residue 7. Sequential  $\beta$  hairpin loops contact each other in a chain-like arrangement. The calcium-containing loop of blade seven makes extensive contacts with the thigh domain; the presence of calcium is likely to make this interface more rigid.

### $\beta$ sandwich domains

The remainder of  $\alpha V$  consists of the thigh and calf-1 and -2 domains [Fig. 1C and Web fig. 4 (17)]. The  $\beta$  sheet structure of the thigh domain is similar in topology to the C2-set Ig fold (19), although the thigh domain is significantly larger. The contact area between the propeller and thigh domains buries a ~700 Å<sup>2</sup> surface area. The elongated shape and intermediate size of this interface suggest that some inter-domain movement, likely as a rotation around the long axis of the contact area, is possible. Such movement could be regulated by calcium ions in the propeller.

Each calf domain contains two antiparallel  $\beta$  sheets, one with four strands (ABGD) and the other with five (EFCHI). The two domains are related by an approximate twofold screw axis symmetry. The known proteolytic cleavage site after Arg<sup>860</sup>, which generates the heavy and light chains of  $\alpha V$ , is located in the EE' loop of the calf-2 domain. The residues flanking the cleavage site (Gln<sup>839</sup>-Gly<sup>867</sup>) are not seen in the electron density map and are presumably disordered.

A small  $\sim 200 \text{ \AA}^2$  interface exists in the crystal at the genu between the thigh and calf-1 domains. The base of the thigh domain and the top of the calf-1 domain both contain acidic patches, and these would be expected to face each other in an extended integrin. A well-coordinated calcium ion lies at the genu; in an extended structure it may help to neutralize the negative charge at this interface. The  $\sim 500 \text{ \AA}^2$  interface between calf-1 and -2 domains is fairly extensive and largely hydrophobic, suggesting that these two domains form a rigid structural entity, the calf module.

## $\beta 3$ Subunit Domains

### $\beta A$ and hybrid domains

The  $\beta 3$  portion of the integrin head is composed of the  $\beta A$  and hybrid domains (Fig. 1C). The  $\beta A$  domain (Fig. 3A) is inserted into the B-C loop of the hybrid domain [Fig. 1C and Web fig. 5 (17)] and assumes the nucleotide-binding (or Ross-mann) fold found in  $G\beta$  and integrin  $\alpha A$  domains (13). The  $\beta A$  domain consists of a central six-stranded  $\beta$  sheet surrounded by eight helices. A MIDAS motif occupies a crevice at the top of the central  $\beta$  strand, as in  $\alpha A$  domains (13).  $\beta 3$  MIDAS is formed by the side chains of Asp<sup>119</sup>, Ser<sup>121</sup>, Ser<sup>123</sup>, Glu<sup>220</sup> (equivalent to Thr<sup>209</sup> in  $\alpha A$  of CD11b), and Asp<sup>251</sup> (equivalent to Asp<sup>242</sup> in  $\alpha A$  of CD11b) (Fig. 3B). Thus, the overall geometry of the metal-ligand coordination is similar to that of  $\alpha A$ . However, we do not clearly see a metal ion in  $\beta 3$  MIDAS: although there is a density feature present, it likely represents a water molecule or a metal ion with partial occupancy. Adjacent to MIDAS lies a metal ion-binding site (ADMIDAS) (Fig. 3C). We assigned a calcium ion at ADMIDAS because calcium is present in the crystallization buffer and because the ion is octahedrally coordinated. In addition, the ion can be replaced by the calcium analog Lu<sup>3+</sup> in the derivatized crystal. Calcium is coordinated by the carbonyl oxygens of Ser<sup>123</sup> (from A'- $\alpha 1$  loop) and Met<sup>335</sup> (from  $\alpha 7$ ) and by the side chains of Asp<sup>126</sup> and Asp<sup>127</sup> (from  $\alpha 1$ ) (Fig. 3C), thus linking  $\alpha 1$  to the top of  $\alpha 7$ . Asp<sup>126</sup> and Asp<sup>127</sup> are invariant in all  $\beta A$  domains except that of  $\beta 8$ , where they are both replaced with asparagines.

$\alpha A$  exists in two conformations, "open" and "closed" (20,21), corresponding to "high" and "low" affinity states, respectively (22,23). It has been suggested that both forms are present in an equilibrium, with a switch from the "closed" to the "open" state being a necessary feature of activation in  $\alpha A$ -integrins (22). The "closed" and "open" conformations are distinguished primarily by the length and relative position of helix  $\alpha 7$ . In the "closed" form,  $\alpha 7$  has an extra NH<sub>2</sub> terminal turn and is tethered to the domain body through hydrophobic contacts that include insertion of a conserved isoleucine (I<sup>316</sup> in  $\alpha A$  of CD11b) from  $\alpha 7$  into the socket-for-isoleucine (SILEN) (23). Removal of this contact shifts  $\alpha 7$  by  $\sim 10 \text{ \AA}$ , switching  $\alpha A$  to the "open" high-affinity form.

Except for the two insertions in the B-C and D- $\alpha 5$  loops, the  $\beta A$  domain is largely superimposable on that of  $\alpha A$ . However,  $\alpha 7$  of the  $\beta A$  domain is more similar to  $\alpha 7$  in the "open" state of  $\alpha A$ : it lacks an extra turn at its top (Fig. 3D) and is positioned similarly relative to the body of the domain (Fig. 3E). Both features suggest that the  $\beta A$  domain is "open" in our  $\alpha V\beta 3$  structure, though it is not liganded. The NH<sub>2</sub>- and COOH-termini connecting the  $\beta A$  domain to the hybrid domain appear to be rigidly held in place through numerous contacts with the hybrid domain, suggesting that the  $\beta A$  domain, in contrast to  $\alpha A$ , always exists in the "open" state and may, therefore, be regulated differently.

The hybrid domain (Fig. 1, B and C) is similar to the I-set Ig domains (19). Its core most closely resembles that of the I-set domain 1 of CD102 [see Web fig. 5 (17)]. The hybrid and  $\beta A$  domains make extensive contacts with each other. The  $\beta A$  domain-hybrid interface is circular (diameter,  $\sim 15 \text{ \AA}$ ), fairly large ( $650 \text{ \AA}^2$ ), and has a mixed hydrophilic and hydrophobic nature. Its

architecture and size suggest that interdomain movement across the  $\beta$ A domain–hybrid interface is minimal.

### PSI domain

On the basis of sequence homology, the NH<sub>2</sub>-terminal 54 amino acids of  $\beta$ 3 are predicted to fold into a PSI domain (18), located in the region connecting the hybrid domain to the expected EGF-1 domain (Fig. 1, B and C). The electron density for PSI is weak, and its C $\alpha$  chain cannot be traced with certainty at the current resolution. Thus, the PSI domain is not included in the present model, although the nature of the electron density suggests that it is partially ordered.

### EGF domains

Four tandem cysteine-rich repeats follow the hybrid domain in the primary sequence. Good electron density allowed us to build a model for cysteine-rich repeats 3 and 4 [Fig. 1, B and C, and Web fig. 6 (17)]. Each assumes the structure of a class I EGF fold (24) and contains the three disulfide bonds typical for EGF domains. An additional pair of cysteines links the two domains. EGF-4 contains an extra disulfide bridge that helps define the COOH-terminal boundary of the short D strand. EGF-3 and -4 form a rod-like, extended module in which the two domains are related by an approximate two-fold screw axis symmetry (Fig. 1C). A consensus site found in calcium binding (cb) EGF domains such as fibrillin (24) is lacking in the integrin EGF domains, and there is no evidence in our electron density map for calcium (or heavy atom) binding. A small ( $\sim 200 \text{ \AA}^2$ ), but presumably rigid, interface exists between EGF-3 and -4: it includes the disulfide bridge as well as main chain hydrogen bonds, hydrophobic interactions, and a glycan.

The electron density is not sufficiently featured to trace the predicted EGF-1 and -2 domains. Most likely, these domains are partially disordered because of the severe bending of the integrin in this region. Examination of their sequence [Web fig. 6 (17)] shows, however, that EGF-1 and -2 are highly homologous to EGF-3 and -4; each is expected to contain the typical three cysteine pairs and to lack the metal ion site found in cbEGF domains. Thus, EGF-1 and -2 are also likely to form a non-calcium-binding rod-shaped module.

**$\beta$ TD.**  $\beta$ TD consists of a four-stranded  $\beta$  sheet that contains antiparallel and parallel strands and faces an NH<sub>2</sub>-terminal  $\alpha$  helix (Fig. 1C). Interactions between the  $\alpha$  helix and the  $\beta$  sheet are mostly hydrophobic and involve a disulfide bond [see Web fig. 7 (17)]. The rear of the  $\beta$  sheet is covered with a long A-B loop. A DALI search (25) showed only a weak homology to cystatin C, an inhibitor of the papain-family cysteine pro-teases (26). However, the  $\beta$ TD structure differs from that of cystatin C in the direction of one of the  $\beta$  strands in the sheet. Hence  $\beta$ TD has to be classified as a new fold. Only two weak hydrophobic contacts are found between  $\beta$ TD and EGF-4, suggesting flexibility at this interface.

### The $\alpha$ V $\beta$ 3 Interface

The  $\beta$ A domain–propeller interface bears a striking resemblance to the  $G\alpha$ - $G\beta$  interface of G proteins (Fig. 4). The seven-bladed propellers of  $\alpha$ V and  $G\beta$  have similar dimensions, and they also have almost identical orientations of the  $\beta$  strands with respect to the central channel (Fig. 4). Furthermore, both  $\beta$ A and  $G\alpha$  domains rest off-center with respect to the propellers, and in each case contact six of the seven blades. In both cases, the axis of the propeller's channel is aimed directly at a short helix: the  $3_{10}$ - and  $\alpha 2$ -helices of  $\beta$ A and  $G\alpha$  domains, respectively (Fig. 4). In the case of  $G\alpha$ , the  $\alpha 2$ -helix marks switch region II, which undergoes a major conformational change upon activation of G proteins.

Residue Arg<sup>261</sup> lies at the core of the  $\beta$ A domain–propeller interface (Fig. 1A, Fig. 2, C and D, and Fig. 5). Arg<sup>261</sup> (Lys in all other integrin  $\beta$  subunits except  $\beta 4$ ) protrudes off-center into

the propeller's channel, and is caged into place by two concentric rings of predominantly aromatic propeller residues (Fig. 1A and Fig. 2, C and D). Residues in the upper and lower rings lie at positions 1 and 4, respectively, of the seven cups in  $\alpha V$  (Fig. 2, B and C). Side chains of Phe<sup>21</sup>, Phe<sup>159</sup>, Tyr<sup>224</sup>, Phe<sup>278</sup>, and Tyr<sup>406</sup> form the lower ring and contact Arg<sup>261</sup> directly (Fig. 2C). Some of these contacts occur through cation- $\pi$  bonding (27). Residues Tyr<sup>18</sup>, Trp<sup>93</sup>, Tyr<sup>221</sup>, Tyr<sup>275</sup>, and Ser<sup>403</sup> in the upper ring (Fig. 2D) contact side chains in the lower ring and also provide a hydrophobic interface for residues flanking Arg<sup>261</sup> in the  $3_{10}$ -helix. The  $\beta A$  domain-propeller interface has a number of additional contacts (Fig. 5) and buries a large surface area of  $\sim 1600 \text{ \AA}^2$ .

Naturally occurring mutations that disrupt the  $\alpha\beta$  heterodimer in integrins have been described [reviewed in (2,28)]. These occur most frequently in the  $\beta A$  domain of  $\beta 2$  and  $\beta 3$  integrin subunits, resulting in life-threatening bacterial infections and thrombasthenia (a severe bleeding disorder), respectively. Our model helps explain how such mutations cause disruption of the heterodimer. For example, Gly<sup>251</sup> to Arg or Pro<sup>156</sup> to Leu mutations are known to prevent formation of  $\beta 2$  integrin heterodimers. Gly<sup>260</sup> in  $\beta 3$  (corresponding to Gly<sup>251</sup> in  $\beta 2$ ) lies in the  $3_{10}$ -helix next to Arg<sup>261</sup>, which makes the central contact with the propeller (Fig. 2, C and D). A Gly to Arg substitution here would result in clashes with other  $\beta A$  residues and would likely alter the position of the  $3_{10}$  helix relative to the propeller. Pro<sup>163</sup> in  $\beta 3$  (equivalent to Pro<sup>156</sup> of  $\beta 2$ ) lies in a loop adjacent to the  $3_{10}$ -helix and contacts the propeller (Fig. 5); the longer leucine side chain would disrupt the  $\alpha\beta$  interface. In  $\beta 3$ , a Ser<sup>162</sup> to Leu mutation results in an unstable  $\alpha IIb\beta 3$  heterodimer and causes thrombasthenia. Ser<sup>162</sup> lies above blade 2 of the propeller; replacing it with leucine would result in unfavorable contacts at the  $\alpha\beta$  interface.

Although the propeller- $\beta A$  domain interface is the main contact between the  $\alpha V$  and  $\beta 3$  subunits, in the crystal, additional contacts are seen between the propeller-EGF-3 and -4 domains, thigh-EGF-3, calf-2-EGF-4, and calf-2- $\beta TD$ . All these contacts are small, not very hydrophobic, and for the most part discontinuous. Therefore, they are not expected to occur in the membrane-bound receptor.

## Implications for Integrin Function and Regulation

The structure described in this communication is that of  $\alpha V\beta 3$  in its active (ligand-competent) state. Activation in this case is artificially induced by truncating the transmembrane and cytoplasmic segments that normally restrain integrins into a default low-affinity state (29,30). Naturally occurring  $\beta 3$  mutants (from some patients with thrombasthenia) and mutagenesis studies of the related integrin  $\alpha IIb\beta 3$  [reviewed in (1,31)] identified three ligand-binding regions. Two of these are located in the  $\alpha IIb$  propeller and the third involves the  $\beta A$  domain MIDAS. The  $\beta A$  domain also contains a fourth "ligand specificity" region. Replacing this region in  $\beta 1$  with that of  $\beta 3$  switched ligand-binding specificity of integrin  $\alpha V\beta 1$  into that of  $\alpha V\beta 3$  [reviewed in (1,31)]. The two  $\beta 3$  regions and the  $\alpha V$  residues that correspond to the two regions identified in  $\alpha IIb$  cluster at the top of the integrin head (Fig. 6). We propose that ligand binding occurs in this area. Confirmatory evidence comes from rotary-shadowed preparations where fibrinogen is found associated with the top of  $\alpha IIb\beta 3$  (6). The putative  $\alpha V\beta 3$  ligand-binding site is formed primarily by residues in loops D3-A3 (between blades 2 and 3) and B3-C3 (within blade 3) of the propeller and by residues in loop B-C and the MIDAS motif of the  $\beta A$  domain. The  $\alpha A$  domain loops out from the COOH-terminal end of the D3-A3 loop in  $\alpha A$ -integrins.

Major conformational changes throughout the  $\beta$  subunit and reorientation of the extracellular domains of the  $\alpha$  and  $\beta$  subunits appear to coincide with integrin activation, whether induced physiologically or artificially (9,32). Our crystal structure reveals a severely bent  $\alpha V\beta 3$  conformation. This arrangement differs dramatically from the extended conformations seen in

cryoelectron microscopy images of integrins reconstituted in lipid bilayers (7) or rotary shadowing of isolated integrins (4–6). However, a minority of rotary shadowing images suggests that the tails may have collapsed on top of (or beneath) the head (5,6). The dimensions of this oblong form (120 Å by 80 Å) agree nicely with our structure, suggesting that similar bending may occur in solution. The fact that extended and severely bent conformations of the same molecule are seen suggests that a highly flexible site, the genu, exists in the integrin. The extreme degree of bending found in the crystal structure is less likely to occur in a membrane-anchored integrin because such a conformation would likely constrain access of polymerized extracellular matrix to the ligand-binding site. Nevertheless, the high degree of flexibility at the genu provides a first glimpse of the spectrum of the quaternary changes possible during bi-directional signaling in integrins. The presence of a metal ion at the genu suggests one mechanism by which these changes could be regulated.

The structural similarities between the propeller- $\beta$ A domains and the  $G\alpha$ - $G\beta$  domains are intriguing (Fig. 4). In G proteins, GTP hydrolysis induces a major conformational shift in  $G\alpha$ , a component of which is the destabilization of the  $\alpha$ 2-helix in switch region II. This results in physical separation of the two subunits in the active state and unmasks ligand-binding sites on the propeller (33). Does the structural homology between integrins and G proteins translate into analogous models of allosteric control? A large interface between the propeller and  $\beta$ A domains is seen in our structure, in contrast to G proteins where in the active state, the equivalent  $G\beta$  and  $G\alpha$  are completely dissociated. Other observations suggest, however, that the propeller- $\beta$ A domain contact area may also be a dynamic interface that can be modulated. First, integrin  $\alpha$ I**II** $\beta$ 3 can be reversibly dissociated into its individual subunits by brief treatment with EDTA (34). Second, a conformational shift of the propeller relative to the  $\beta$ A domain has been reported (32). Third, a split of the  $\alpha$ I**II** $\beta$ 3 head into two distinct knobs was observed in the presence of RGD peptides (11). In our structure, cation- $\pi$  bonding between Arg<sup>261</sup> and surrounding aromatic residues contributes to the  $\alpha\beta$  interface. Amide-aromatic interactions can be attractive or repulsive (27), unlike ion pairs. Thus, it is conceivable that small changes in side chain orientation could modulate the stability of the  $\alpha\beta$  interface, perhaps allowing for dissociation under certain conditions.

The location of the metal ion-binding sites in  $\alpha$ V $\beta$ 3 helps clarify some of the complex and unresolved regulatory effects exerted by metal ions on integrin-ligand binding. Cells expressing a mutant  $\alpha$ 4 $\beta$ 1 (lacking one or more metal-binding sites in the propeller) detach easily from substrate under shear flow (35). The calcium ions in the propeller are not close to the proposed ligand-binding site shown in Fig. 6, but they may help to make more rigid the propeller-thigh interface and may thus regulate integrin-ligand interactions in an allosteric manner. The calcium ion at the thigh-calf-1 interface may play a similar role. Studies with  $\beta$ 3 and  $\alpha$ 5 $\beta$ 1 integrins have also shown that a high-affinity  $Ca^{2+}$  site is required for ligand-binding, and that a low-affinity  $Ca^{2+}$  site allosterically inhibits ligand binding (36,37). The presence of two adjacent metal-binding sites (MIDAS and ADMIDAS) in the  $\beta$ A domain suggests underlying mechanisms for these effects.

The structure of  $\alpha$ V $\beta$ 3 reveals new domains, previously unpredicted domains, and creative use of the Ig scaffold. Integrin-ligand interactions are regulated not only by changes in affinity but also by altered avidity (receptor clustering) [reviewed in (38)]. Integrins also bind in cis to several membrane receptors, an interaction that modulates their signaling functions [reviewed in (39)]. The mosaic of domains revealed in the integrin structure can now serve as a foundation for future investigations into the structural basis of these interactions.

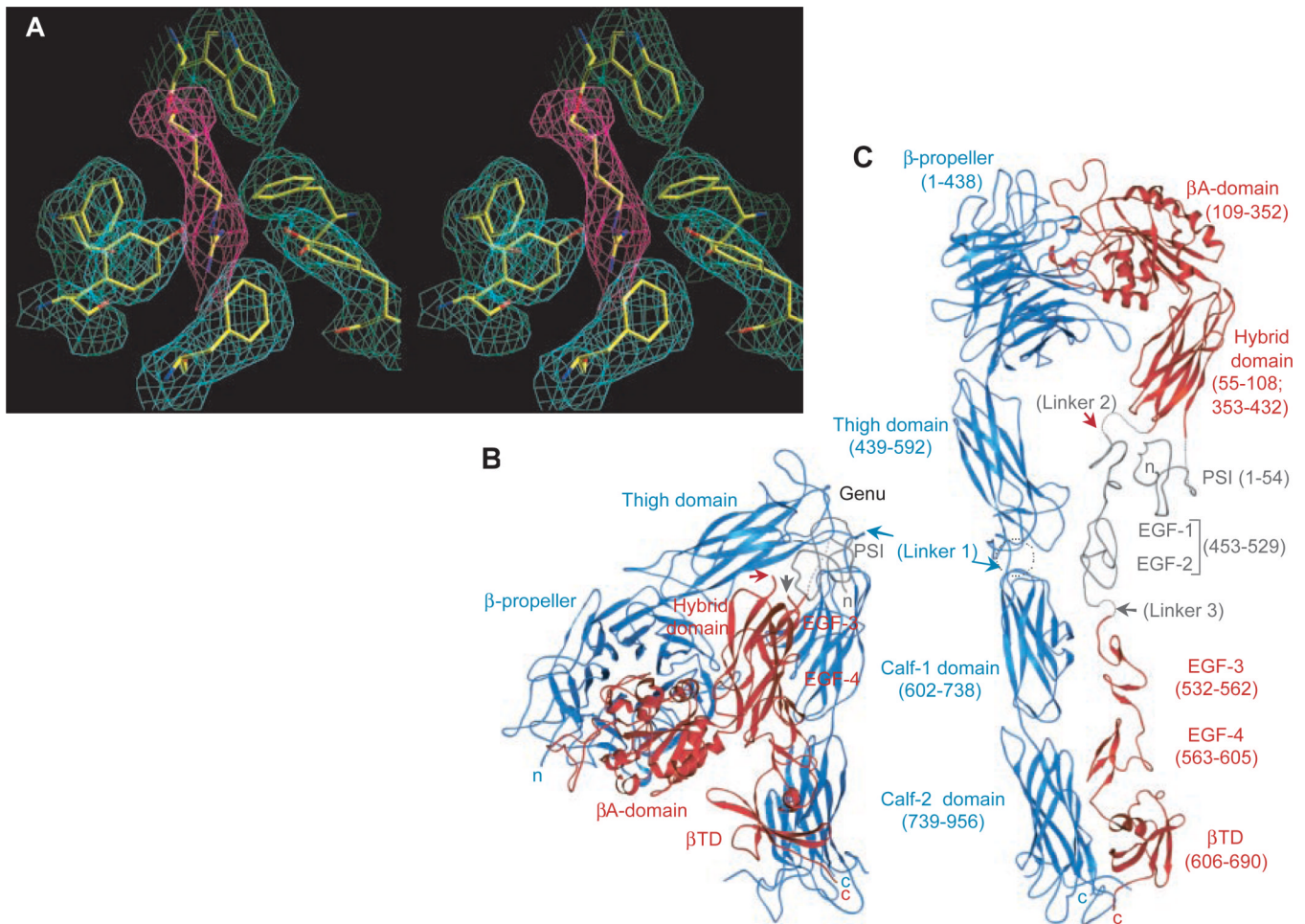
## References and Notes

1. Humphries MJ. *Biochem. Soc. Trans* 2000;28:311. [PubMed: 10961914]

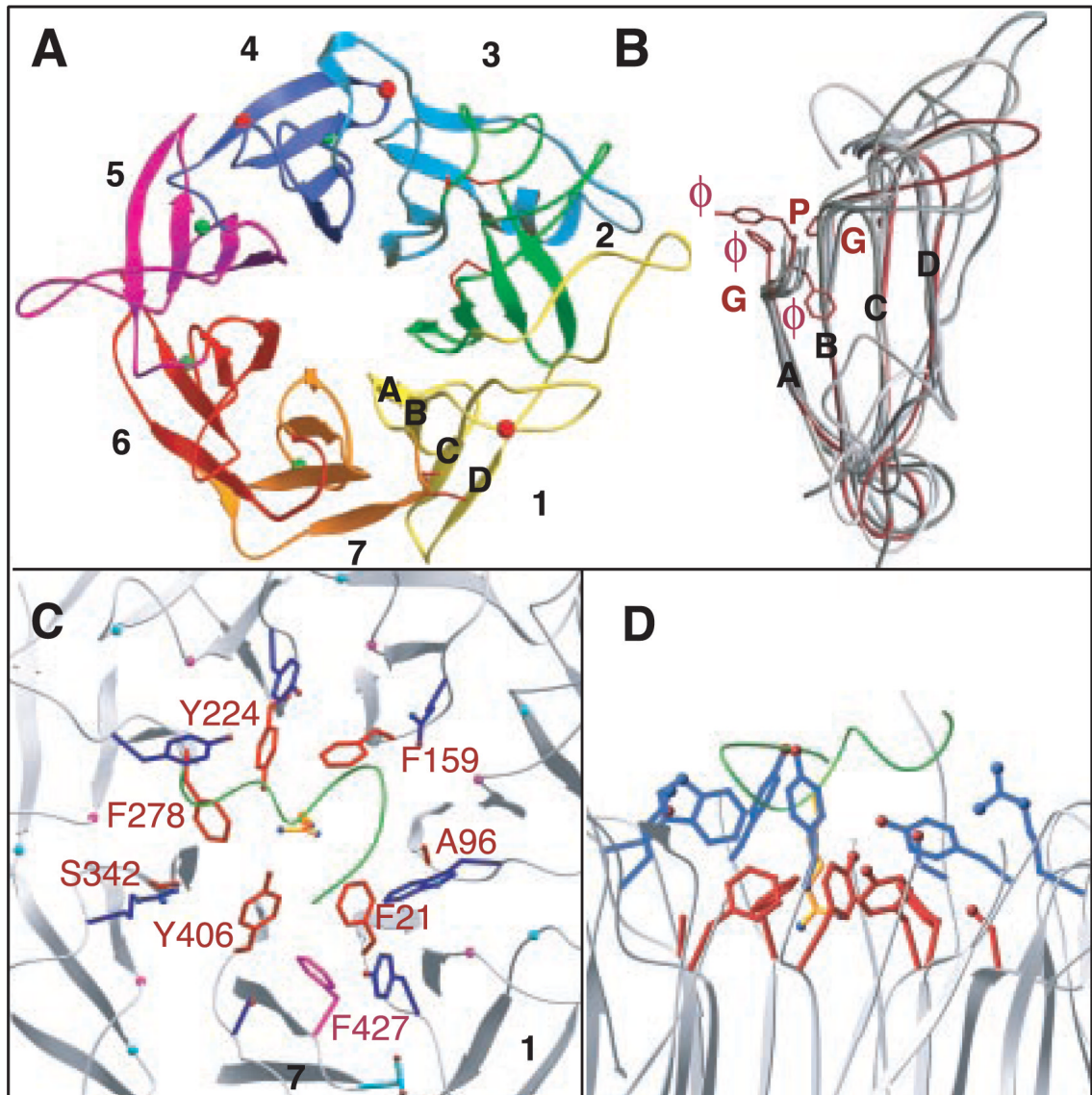
2. Arnaout MA. *Immunol. Rev* 1990;114:145. [PubMed: 1973407]
3. Hynes RO. *Cell* 1992;69:11. [PubMed: 1555235]
4. Parise LV, Phillips DR. *J. Biol. Chem* 1985;260:1750. [PubMed: 3155739]
5. Nermut MV, Green NM, Eason P, Yamada SS, Yamada KM. *EMBO J* 1988;7:4093. [PubMed: 2977331]
6. Weisel JW, Nagaswami C, Vilaire G, Bennett JS. *J Biol Chem* 1992;267:16637. [PubMed: 1644841]
7. Erb E-M, Tangemann K, Bohrmann B, Muller B, Engel J. *Biochemistry* 1997;36:7395. [PubMed: 9200686]
8. Yamada KM, Geiger B. *Curr. Opin. Cell Biol* 1997;9:76. [PubMed: 9013677]
9. Sims PJ, Ginsberg MH, Plow EF, Shattil SJ. *J. Biol. Chem* 1991;266:7345. [PubMed: 1902217]
10. Du X, et al. *J. Biol. Chem* 1993;268:23087. [PubMed: 7693683]
11. Hantgan RR, Paumi C, Rocco M, Weisel JW. *Biochemistry* 1999;38:14461. [PubMed: 10545168]
12. Michishita M, Videm V, Arnaout MA. *Cell* 1993;72:857. [PubMed: 8458080]
13. Lee J-O, Rieu P, Arnaout MA, Liddington R. *Cell* 1995;80:631. [PubMed: 7867070]
14. Eliceiri BP, Cheresch DA. *J. Clin. Invest* 1999;103:1227. [PubMed: 10225964]
15. Plow EF, Haas TA, Zhang L, Loftus J, Smith JW. *J. Biol. Chem* 2000;275:21785. [PubMed: 10801897]
16. Mehta RJ, et al. *Biochem J* 1998;330:861. [PubMed: 9480902]
17. Online at [www.sciencemag.org/cgi/content/full/1064535/DC1](http://www.sciencemag.org/cgi/content/full/1064535/DC1) For additional information, see Web figs. 1 through 7, available at *Science*
18. Bork P, Doerks T, Springer TA, Snel B. *Trends Biochem. Sci* 1999;24:261. [PubMed: 10390613]
19. Chothia C, Jones EY. *Annu. Rev. Biochem* 1997;66:823. [PubMed: 9242926]
20. Lee J-O, Anne-Bankston L, Arnaout MA, Liddington RC. *Structure* 1995;3:1333. [PubMed: 8747460]
21. Emsley J, Knight CG, Farndale RW, Barnes MJ, Liddington RC. *Cell* 2000;100:47. [PubMed: 10778855]
22. Li R, Rieu P, Griffith DL, Scott D, Arnaout MA. *J. Cell Biol* 1998;143:1523. [PubMed: 9852148]
23. Xiong JP, Li R, Essafi M, Stehle T, Arnaout MA. *J. Biol. Chem* 2000;275:38762. [PubMed: 11034990]
24. Downing AK, et al. *Cell* 1996;85:597. [PubMed: 8653794]
25. Holm L, Sander C. *J. Mol. Biol* 1993;223:123. [PubMed: 8377180]
26. Janowski R, et al. *Nature Struct. Biol* 2001;8:316. [PubMed: 11276250]
27. Gallivan JP, Dougherty DA. *Proc. Natl. Acad. Sci. U.S.A* 1999;96:9459. [PubMed: 10449714]
28. Hogg N, Bates PA. *Matrix Biol* 2000;19:211. [PubMed: 10936446]
29. Dana N, Fathallah DF, Arnaout MA. *Proc. Natl. Acad. Sci. U.S.A* 1991;88:3106. [PubMed: 1673028]
30. Hughes PE, et al. *J. Biol. Chem* 1996;271:6571. [PubMed: 8636068]
31. Calvete JJ. *Proc. Soc. Exp. Biol. Med* 1999;222:29. [PubMed: 10510244]
32. Mould AP, Garratt AN, Puzon-McLaughlin W, Takada Y, Humphries MJ. *Biochem J* 1998;331:821. [PubMed: 9560310]
33. Wall MA, et al. *Cell* 1995;83:1047. [PubMed: 8521505]
34. Fitzgerald LA, Phillips DR. *J. Biol. Chem* 1985;260:11366. [PubMed: 3161896]
35. Pujades C, et al. *Mol. Biol. Cell* 1997;8:2647. [PubMed: 9398682]
36. Mould AP, Akiyama SK, Humphries MJ. *J. Biol. Chem* 1995;270:26270. [PubMed: 7592835]
37. Hu DD, Barbas CF, Smith JW. *J. Biol. Chem* 1996;271:21745. [PubMed: 8702970]
38. van Kooyk Y, van Vliet SJ, Figdor CG. *J. Biol. Chem* 1999;274:26869. [PubMed: 10480895]
39. Hemler ME. *Curr. Opin. Cell Biol* 1998;10:578. [PubMed: 9818167]
40. Carson M. *J. Mol. Graph* 1987;5:103.
41. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.



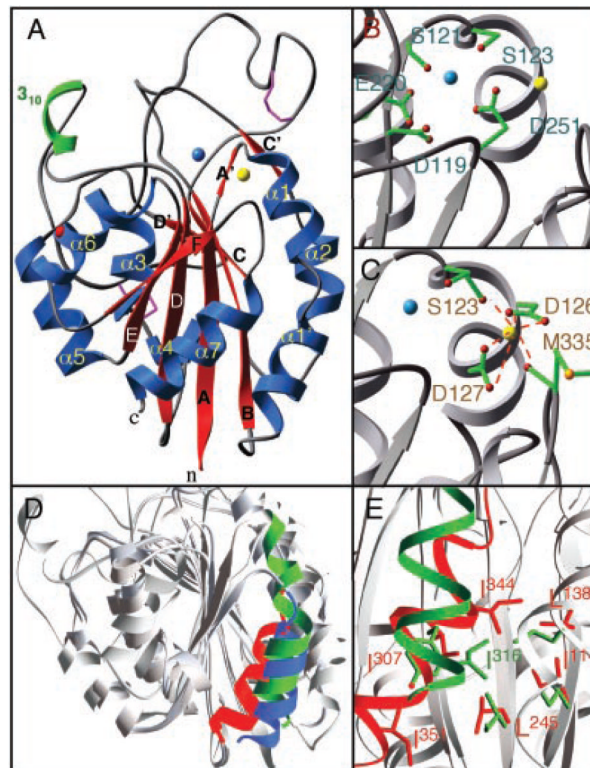
42. Nicholls A, Sharp KA, Honig B. *Proteins* 1991;11:281. [PubMed: 1758883]
43. Otwinowski Z, Minor W. *Methods Enzymol* 1997;276:307.
44. de La Fortelle E, Bricogne G. *Methods Enzymol* 1997;276:47.
45. CCP4. *Acta Crystallogr. D* 1994;50:760. [PubMed: 15299374]
46. Jones TA, Zou JY, Cowan SW, Kjeldgaard M. *Acta Crystallogr. A* 1991;47:110. [PubMed: 2025413]
47. Brünger AT, et al. *Acta Crystallogr. D* 1998;54:905. [PubMed: 9757107]
48. We thank A. Viel, M. Frech, and D. Cheresch for valuable assistance. Supported by NIH grants DK48549, DK50305, HL54227, and AI45716 and a contract from the DOE under contract W-31-109-Eng-38. The coordinates have been deposited in the Protein Data Bank (PDB1JV2).



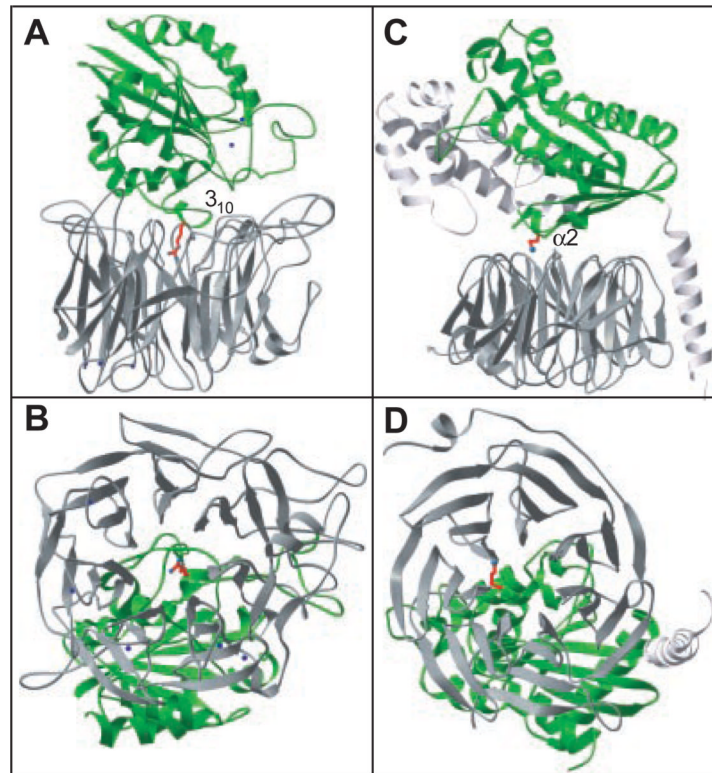
**Fig. 1.** Structure of the extracellular segment of  $\alpha V\beta 3$ . **(A)** Stereoview of a simulated-annealing omit map at  $1.5 \sigma$ , in the vicinity of Arg<sup>261</sup> (from  $\beta 3$ ) (magenta). Surrounding densities (cyan) (from  $\alpha V$ ) are from the same map. **(B)** Ribbon drawing (40) of crystallized  $\alpha V\beta 3$  [shown in blue ( $\alpha V$ ) and red ( $\beta 3$ )]. **(C)** Model of the straightened extracellular segment of  $\alpha V\beta 3$ . The two tails would extend into the plasma membrane in the native integrin. Translated and rotated EGF-3 and -4 show the approximate location of EGF-1 and -2 (gray). The PSI tracing (gray) is approximate. Connections of the untraced domains are in dotted lines.  $\alpha V$  was straightened by extending the structure by  $135^\circ$  at the thigh–calf-1 interface (circled) and then rotating the calf module  $\sim 120^\circ$  around its “long” axis to avoid clashes at the thigh–calf-1 interface. The same transformations were then applied to  $\beta 3$  (residue 445 onward). Arrows point to the position of the three longest inter-domain linkers 1, 2, and 3 in the structure. Amino acid domain boundaries are indicated in parenthesis. In this and other figures, “n” and “c” indicate NH<sub>2</sub>- and COOH-terminus, respectively.



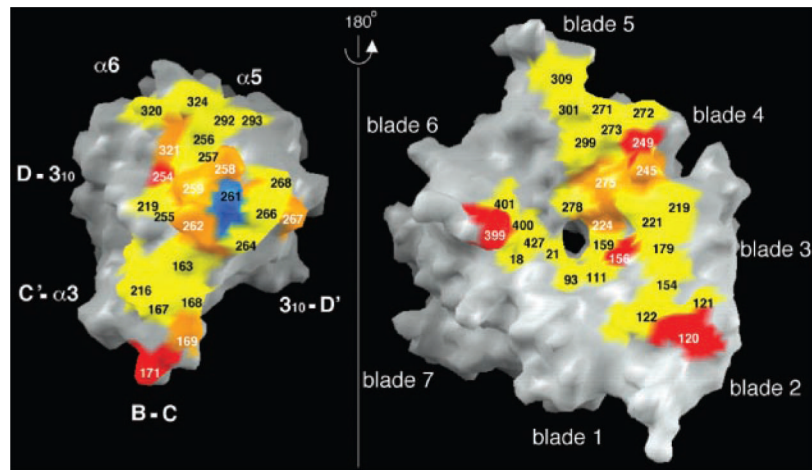
**Fig. 2.** Structure of the integrin  $\beta$ -propeller. **(A)** Bottom view of the seven-bladed (numbered)  $\alpha$ V propeller. Disulfides (sticks) and glycans (spheres) are in red and  $\text{Ca}^{2+}$  ions are in green. **(B)** Super-imposition of the seven blades, with blade five shown in red. The cage residues ( $\phi$ , aromatic; G, Gly; and P, Pro) are in red. The  $\phi\phi G\phi$  sequence assumes the shape of a cup. Strands A to D are labeled. Top **(C)** and side **(D)** views of the central propeller- $\beta$ A domain interface. The propeller's lower (red) and upper (blue) aromatic rings surrounding  $\text{Arg}^{261}$  of  $\beta 3$  are shown. The  $3_{10}$ -helix of  $\beta 3$  is in green. The lower ring residues are labeled in red. The prolines and glycines of the cage motif are shown as magenta and cyan spheres, respectively.  $\text{F}^{427}$  and  $\text{D}^{430}$  replace the proline and glycine residues in blade seven, respectively (41).



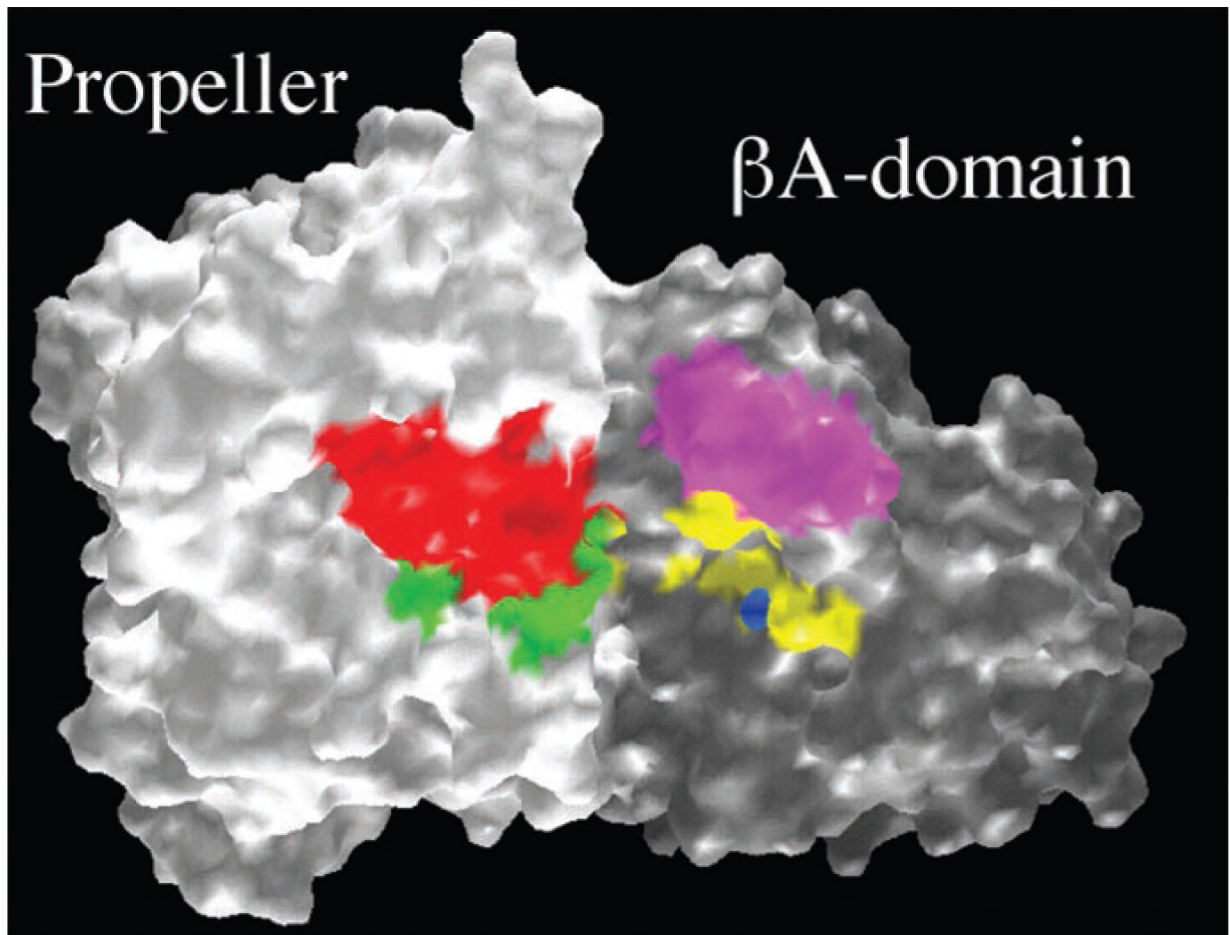
**Fig. 3.** Structural features of the  $\beta$ A domain. **(A)** Ribbon drawing of the  $\beta$ A domain. Disulfides and the glycan are in purple and red, respectively. **(B)** Residues [single letter (41)] forming MIDAS that could participate in metal ion (blue) coordination. **(C)** Calcium (yellow) coordination at ADMIDAS. **(D)** Superimposition, based on the central  $\beta$  sheet, of the  $\beta$ A domain (minus the insertion in the B-C loop) and the “open” and “closed” forms of  $\alpha$ A of CD11b (all gray). Shown are the  $\alpha$ 7-helices of the  $\beta$ A domain (red) and the “open” (blue) and “closed” (green) forms of  $\alpha$ A of CD11b. **(E)** Coordination of SILEN in the  $\beta$ A domain (red) and in “closed”  $\alpha$ A of CD11b (green). The  $\beta$ A domain SILEN residues [single letter, (41)] I<sup>114</sup>, L<sup>138</sup>, L<sup>245</sup>, and I<sup>307</sup> (red) [equivalent to I<sup>135</sup>, L<sup>164</sup>, I<sup>236</sup>, and Y<sup>267</sup>, respectively, in  $\alpha$ A of CD11b (green) (23)] coordinate I<sup>344</sup> (equivalent in sequence to L<sup>312</sup> in  $\alpha$ A of CD11b). I<sup>351</sup> (corresponding to I<sup>316</sup> in  $\alpha$ A of CD11b) is not coordinated in  $\beta$ 3 SILEN but is partially exposed at the bottom of the structure.



**Fig. 4.** Architecture of integrin and G protein interfaces. The nucleotide-binding folds are in green, and the propellers are in gray. The integrin is shown in (A) and (B) and the G protein in (C) and (D). The helices at the core of each interface are indicated. Integrin's Arg<sup>261</sup> and G protein's Lys<sup>210</sup> project toward the respective propeller's central cavity. G $\alpha$  (Ala<sup>30</sup>–Leu<sup>348</sup>) and G $\beta$  (Ser<sup>2</sup>–Asn<sup>340</sup>) are shown. G protein coordinates were from (33).



**Fig. 5.** Surface representation, done with GRASP (42), of the main  $\alpha\text{V}\beta\text{3}$  interface in the head region. Contacting residues (distance cutoff 3.5 Å for hydrogen bonds and salt bridges, and 4.0 Å for van der Waals contacts) are shown. Arg<sup>261</sup> is in blue. Hydrophobic, ionic, and mixed contacts are in yellow, red, and orange, respectively. The  $\beta\text{3}$  strands and helices contributing to the interface are indicated.



**Fig. 6.**

The putative ligand-binding site at the integrin  $\alpha\beta$  interface. GRASP representation of the  $\alpha$ V propeller (white) and the  $\alpha$ 3 A-domain (gray), viewed from the top of the head (the tails lying in the back). Arg<sup>143</sup>-Phe<sup>154</sup> (red) and Gly<sup>172</sup>-Gly<sup>181</sup> (green) of  $\alpha$ V correspond to the identified ligand-binding residues Arg<sup>147</sup>-Tyr<sup>166</sup> and Gly<sup>184</sup>-Gly<sup>193</sup>, respectively, in  $\alpha$ IIb. Ligand-binding residues in  $\beta$ 3 (Asp<sup>179</sup>-Thr<sup>183</sup>, “ligand specificity” region (magenta); Asp<sup>119</sup>, Ser<sup>121</sup>, Ser<sup>123</sup>, Glu<sup>220</sup> (all MIDAS residues), and Arg<sup>214</sup> (yellow) are shown. Asp<sup>217</sup> in  $\beta$ 3 (also in yellow) is involved in binding of a ligand-mimetic monoclonal antibody. The expected MIDAS metal ion is in blue.

Crystallographic data and refinement statistics. Methods are as follows:  $\alpha$ V $\beta$ 3 crystals belong to space group P3<sub>2</sub>2<sub>1</sub>, with unit cell dimensions of  $a = b = 130.0$  Å, and  $c = 307.3$  Å. Each asymmetric unit contains a single heterodimer and ~61% solvent. Crystals were cryoprotected in glycerol, and native and derivative data were collected at 100 K at ID-19 beamline at the Argonne National Laboratory. Data were processed with HKL (43). Heavy atom refinement and phase calculations were performed using SHARP (44), and phases were then improved using SOLOMON (45). A high degree of nonisomorphism between the Pt<sup>2+</sup> and Lu<sup>3+</sup> derivative data sets prevented phase combination, and so the polypeptide chain was traced using two separate electron density maps, one based on MAD phases (Lu<sup>3+</sup>) and the other calculated with SIRAS phases (Pt<sup>2+</sup>). Most of the main chain and many side chains could be recognized in these maps, with the exception of the PSI and the predicted EGF-1 and -2 domains of  $\beta$ 3. Model building was carried out in O (46), and refinement was performed with XPLOR (47) using bulk solvent correction and simulated annealing protocols. After several rounds of refinement and model building into the resulting  $\sigma_A$ -weighted electron density maps, atomic B-factor refinement (grouping residues) was carried out. The average B factor for the whole structure is 60 Å<sup>2</sup>. The average B factors for the individual domains are as follows: propeller, 47 Å<sup>2</sup>; thigh, 78 Å<sup>2</sup>; calf-1, 75 Å<sup>2</sup>; calf-2, 60 Å<sup>2</sup>;  $\beta$ A domain, 56 Å<sup>2</sup>; hybrid, 68 Å<sup>2</sup>; EGF-3, 82 Å<sup>2</sup>; EGF-4, 66 Å<sup>2</sup> and  $\beta$ TD, 71 Å<sup>2</sup>. Values in parentheses indicate the highest resolution shell. n/a, not available; No., number.

Table 1

Space group	Data collection			
	P3 <sub>2</sub> 2 <sub>1</sub>	130.0 Å, 130.0 Å, 307.3 Å, 90°, 90°, 120°		
Unit cell dimensions	Native	MAD(Lu <sup>3+</sup> )		SIRAS(Pt <sup>2+</sup> )
	Remote	Edge	Peak	Peak
Wavelength (Å)	1.0332	1.3412	1.3407	0.9840
Resolution (Å)	50-3.1	50-3.6	50-3.2	50-3.5
Unique reflections	53802	34310	46024	32413
% completeness	96.9 (95.0)	91.4 (70.4)	99.5 (99.1)	91.6 (90.7)
Redundancy	6.1	9.0	10.0	8.9
<i>I</i> / $\sigma$	8.7 (3.7)	5.8 (1.0)	5.5 (2.7)	6.6 (2.2)
$R_{\text{sym}}(\%)^*$	10.5 (55.1)	12.1 (63.6)	14.5 (58.2)	14.7 (56.0)
				9.6 (48.4)
No. of heavy atoms			3	8
Phasing power <sup>†</sup>				
Centrics	n/a <sup>‡</sup>	0.69	0.99	0.75
Acentrics				
Isomorphous	n/a	0.79	1.20	0.72
Anomalous	0.94	2.55	1.72	1.37
Mean figure of merit (overall)	0.43			0.23



Space group	Data collection		
	Native	Remote	Edge
P3 <sub>2</sub> 21			
Unit cell dimensions	130.0 Å, 130.0 Å, 307.3 Å, 90°, 90°, 120°		
		MAD(Lu <sup>3+</sup> )	SIRAS(Pt <sup>2+</sup> )
$R_{\text{cullis}}^{\ddagger}$			
Centrics	n/a	0.84	0.77
Acentrics			
Isomorphous	n/a	0.84	0.78
Anomalous	0.91	0.70	0.80
<i>Refinement statistics</i>			
Resolution range (Å)	20.0-3.1		
$R_{\text{factor}}^{\S}$ (%) (work set)	25.5		
$R_{\text{factor}}^{\S}$ (%) (free set)	33.5		
Average B value (Å <sup>2</sup> )	60		
No. of Glc-NAc	18		
No. of Ca <sup>2+</sup>	6		
		Peak	Peak
			0.96
			1.0
			0.83
			0.008
			1.6

\*  $R_{\text{sym}} = \sum |I - \langle I \rangle| / \sum I$ , where  $I$  is the observed intensity and  $\langle I \rangle$  is the average intensity from multiple observations of symmetry-related reflections.

<sup>†</sup> Phasing power =  $F_{\text{H}}(\text{calc})/E$ , where  $E$  is phase-integrated lack-of-closure. The three values for each wavelength are for centric/acentric isomorphous/acentric anomalous contribution.

<sup>‡</sup>  $R_{\text{cullis}} = \sum_{\text{hkl}} |F_{\text{PH}} + F_{\text{P}}| - F_{\text{H}}(\text{calc}) / \sum_{\text{hkl}} |F_{\text{PH}} + F_{\text{P}}|$

<sup>§</sup>  $R_{\text{factor}} = \sum_{\text{hkl}} |F_{\text{obs}}(\text{hkl}) - F_{\text{calc}}(\text{hkl})| / \sum_{\text{hkl}} F_{\text{obs}}(\text{hkl})$ .