Crystal structure of the α1β1 integrin I-domain: insights into integrin I-domain function

Matthias Nolte[a], R. Blake Pepinsky[a], Sergei Yu. Venyaminov[b], Victor Koteliansky[a], Philip J. Gotwals[a], Michael Karpusas[a,*]

[a] *Biogen, Inc., 14 Cambridge Center, Cambridge, MA 02142, USA*
[b] *Department of Biochemistry and Molecular Biology, Mayo Foundation, 200 First St. SW, Rochester, MN 55905, USA*

Received 30 March 1999; received in revised form 30 April 1999

Abstract **The α1β1 integrin is a major cell surface receptor for collagen. Ligand binding is mediated, in part, through a ~200 amino acid inserted 'I'-domain contained in the extracellular part of the integrin α chain. Integrin I-domains contain a divalent cation binding (MIDAS) site and require cations to interact with integrin ligands. We have determined the crystal structure of recombinant I-domain from the rat α1β1 integrin at 2.2 Å resolution in the absence of divalent cations. The α1 I-domain adopts the dinucleotide binding fold that is characteristic of all I-domain structures that have been solved to date and has a structure very similar to that of the closely related α2β1 I-domain which also mediates collagen binding. A unique feature of the α1 I-domain crystal structure is that the MIDAS site is occupied by an arginine side chain from another I-domain molecule in the crystal, in place of a metal ion. This interaction supports a proposed model for ligand-induced displacement of metal ions. Circular dichroism spectra determined in the presence of Ca^{2+}, Mg^{2+} and Mn^{2+} indicate that no changes in the structure of the I-domain occur upon metal ion binding in solution. Metal ion binding induces small changes in UV absorption spectra, indicating a change in the polarity of the MIDAS site environment.**

Key words: I-domain; Integrin; Metal binding; Collagen; Adhesion; Crystal structure

1. Introduction

Integrins are a family of large plasma membrane proteins involved in processes of cell-cell and cell-extracellular matrix adhesion and signal transduction. They exist as non-covalent αβ heterodimers consisting of different combinations of α and β chains. Currently more than 16 α chains and eight β chains are known. Many α chains contain in their extracellular region a ~200 amino acid inserted domain, known as the I-domain or A-domain. This domain is homologous to the von Willebrand factor (vWF) A-domain as well as domains that exist in other proteins such as complement factor B. Ligand binding of the I-domain-containing integrins appears to be mainly mediated through their I-domains, in a metal ion-dependent process [1–4]. Recombinant I-domains retain many of the metal ion and ligand binding properties of the parent integrin and thus are frequently used as models for studying integrin function [5–8]. Native integrins undergo an activation step that increases their affinity for their ligand, which can be mimicked through the addition of Mn^{2+} or certain specific antibodies. While the mechanism leading to activation is not well understood, it is believed to occur, at least in part, through conformational changes of the integrin.

Crystal structures have been solved for the I-domains of the αM, αL and α2 integrin chains in the presence or absence of divalent cations [9–15]. In all structures the I-domain adopts the Rossmann or dinucleotide binding fold with a central six-stranded β-sheet surrounded by seven α-helices. A metal ion binding site exists at one end of the β-sheet formed by residues from a conserved sequence motif DXSXS as well as two additional conserved residues, a D and a T. This site is known as metal ion-dependent adhesion site (MIDAS) and is implicated in ligand binding. Two modes of metal ion coordination in the MIDAS site have been observed in the available crystal structures of I-domains. The first one ('open') has been observed only in one crystal form and includes coordination of the metal ion by a glutamate residue from a neighboring molecule in the crystal [9,15]. The second one ('closed') has been observed in all other crystal forms. These differences in metal ion coordination are associated with differences in the conformation of the I-domain, involving a significant shift of the C-terminal helix and other rearrangements of residues proximal to the MIDAS site. Extensive mutagenesis work and analysis of chimeric I-domains have demonstrated that several residues surrounding the MIDAS site are also part of the ligand binding site [16].

The α1β1 integrin (very late antigen-1; VLA-1) is expressed on a variety of cell types including those of hematopoietic, neuronal and mesenchymal origin. Together with α2β1 it belongs to a subset of integrins that both bind to collagen I, collagen IV and laminin but with different affinities [17–19]. α1β1 integrin binds to collagen IV with higher affinity than to collagen I while the reverse is true for α2β1. These two integrins are also known to recognize different but adjacent sites within collagen IV [20]. A study [21] identified Arg-461 of the collagen α2(IV) chain and Asp-461 within the α1(IV) chain as essential residues for α1β1 binding. A chimeric α1 chain, in which the I-domain has been replaced with the α2 I-domain, has a ligand binding profile similar to that of the α2β1 integrin [22]. The crystal structure of the α2 I-domain showed the presence of a small helix (C-helix), adjacent to the MIDAS motif, which is a unique structural feature of the α1 and α2 I-domain subfamily. Modeling of collagen triple helix bound to the α2 I-domain suggested that the C-helix may be an important structural determinant of collagen binding and specificity [11].

*Corresponding author. Fax: (1) (617) 679-2616.
E-mail: michael_karpusas@biogen.com

PII: S0014-5793(99)00666-3

To gain more in-depth perspective on this integrin subfamily, we have determined the crystal structure of the rat α1 I-domain in the apo state at 2.2 Å resolution and directly compared it to the α2 I-domain structure. In addition, we probed the effect of divalent cations in the structure of the I-domain in solution, by optical spectroscopy.

2. Materials and methods

2.1. Expression and purification of the α1 integrin I-domain

Details of the cloning, expression and purification of the α1 integrin I-domain will be discussed elsewhere (Gotwals et al., submitted for publication). Briefly, the DNA encoding the I-domain fragment of the extracellular domain of the rat α1 integrin chain (amino acid sequence Val-127–Ala-340, referred to as α1-I in subsequent text) was amplified from full length cDNAs by PCR. The resulting PCR amplified product was purified, ligated in pGEX4t expression vector (Pharmacia) and transformed into competent DH5α *Escherichia coli* cells (Gibco BRL). The α1 I-domain was expressed as a GST fusion protein with a thrombin cleavage site at the junction of the sequences. For preparation of the purified α1 I-domain, the fusion protein was treated with thrombin and the I-domain was purified by standard techniques. The purified I-domain was >99% pure by SDS-PAGE, eluted as a single peak by size exclusion chromatography on a Superose 6 column (Pharmacia and Upjohn) and displayed its predicted mass (24 871 Da) by electrospray ionization-mass spectrometry (ESI-MS).

In preliminary studies, we found that the rat α1 integrin I-domain in this form failed to crystallize under any test condition, and, as had been observed for other I-domains (R. Liddington, personal communication), that sequences at the N-terminus of the I domain construct were problematic. A simple proteolytic method was developed to convert the purified rat I-domain into a form that could be crystallized. Briefly, 240 μl of the purified α1 integrin I-domain (16 mg/ml) was diluted with 360 μl of 50 mM Tris-HCl pH 7.5 and loaded onto a 1.2 ml V8 protease column (Pierce) that had been equilibrated in 50 mM Tris-HCl pH 7.5. The I-domain solution was left in contact with the resin for 35 min at room temperature and then recovered by washing the column with 50 mM Tris-HCl pH 7.5. The I-domain was then dialyzed overnight against 10 mM Tris pH 7.5, 10 mM 2-mercaptoethanol and concentrated to 11 mg/ml in a Centricon-10 ultrafiltration unit (Amicon). ESI-MS analysis of the V8 protease-digested product revealed that the product had been converted into a des 1–18 form, starting at Cys-143 in the fusion protein construct.

2.2. Crystallization and structure determination

Crystallization condition screenings were done with the Crystal Screen I kit from Hampton Research (Riverside, CA). In order to find conditions of crystallization, an incomplete factorial screen was set up. Crystals were grown by the vapor diffusion method [23], out of 25% w/v polyethylene glycol (PEG) 8000, 0.1 M sodium cacodylate pH 6.5, 0.2 M sodium acetate reservoir solution. The crystals are plate shaped and can reach maximum dimensions of almost 0.5 mm on one side.

Crystals were equilibrated gradually in a cryoprotectant solution of 20% glycerol, 25% w/v PEG 8000, 0.1 M sodium cacodylate pH 6.5, 0.2 M sodium acetate, and were mounted on a loop and immediately frozen in a −150° C liquid nitrogen gas stream. A native X-ray data set up to 3.0 Å resolution was collected from one crystal by using an R-AXIS II image plate detector system (Molecular Structure Corporation, Woodlands, TX). A second data set to 2.2 Å resolution was collected later by using a larger crystal. The data were integrated and reduced using the HKL program package [24]. Data processing suggested an orthorhombic unit cell with approximate cell dimensions $a = 34.77$ Å, $b = 85.92$ Å, $c = 132.56$ Å and $\alpha = \beta = \gamma = 90°$. The space group was identified as $P2_12_12_1$. The 2.2 Å data set was 91.3% complete (77.6% for 2.2–2.28 Å shell) with an $I/\sigma(I)$ of 19.98 (3.14 for 2.2–2.28 Å shell) and an R-merge of 5.6%. The Matthews volume [25] is 4.22 $Å^3$ Da^{-1}, assuming a molecular weight of 23 000 Da which suggested that there are two molecules in the asymmetric unit.

All molecular replacement computing was done with the program AMoRe [26] from the CCP4 program package [27]. Molecular graphics manipulations were done with programs QUANTA (Molecular Simulations, Inc., San Diego, CA) and 'O' [28]. The coordinates of the crystal structure of the human α2 I-domain [11] were used as a probe for rotation and translation searches using the 3 Å data set. The rotation function gave a solution with the highest correlation coefficient (cc) of 9.7. This solution was used for a translation function calculation that yielded a maximum peak with a cc of 24.6 and an R-factor of 48.7. The translation search for the second molecule, while keeping the first solution fixed, yielded a maximum peak with cc of 37.3 and an R-factor of 44.8%. Rigid body refinement on these two solutions resulted in a cc of 56.3 and an R-factor of 43.3%. The next highest solution had a cc = 36.6 and an R-factor = 49.9%. The rotation matrix between the two molecules of the asymmetric unit was determined and one molecule was used for the initial stages of model building.

Subsequent refinement computing was done with the XPLOR program [29]. 10% of the 8–3 Å data were used for the calculation of R-free. To reduce model bias, partial models were used for 3Fo-2Fc map calculation and model refinement. The initial partial model, containing a polyalanine chain of the secondary structure elements from the α2 I-domain structure, was subjected to positional and grouped B-factor refinement with strict non-crystallographic symmetry constraints. The R and R-free factors dropped to 32.3% and 39.4% respectively. Several cycles consisting of model building, positional refinement and B-factor refinement followed. When the R and R-free reached 26% and 36% respectively, no further improvement of the model could be achieved with the 3 Å data set. The 2.2 Å data set was collected at this point and was used for all subsequent model building and refinement. The R and R-free factors after the initial rigid body refinement at 2.2 Å were 41.3% and 42.2%, respectively. This larger data set allowed use of simulated annealing and torsion angle dynamics refinement. As the phases improved, more atoms were added into the model. Initially, grouped B-factors were assigned for each residue (one for main chain and the one for side chain atoms). Later, non-crystallographic symmetry constraints were removed and individual atomic B-factors were refined for each residue. In addition, bulk solvent correction was applied to the data set. Residues and side chains would be incorporated in the model if they were sufficiently well defined in 3Fo-2Fc electron density maps. Only manual structure modifications that resulted in lower R-free after refinement were accepted. When R and R-free reached 29% and 34.8%, respectively, water molecules were added by using the X-solvate utility of QUANTA. Finally, maximum likelihood refinement was used [30] and resulted in the final structure with R and R-free of 23.4% and 29.9% respectively for data between 100 and 2.2 Å resolution. The atomic coordinates and structure factors have been deposited in the Protein Data Bank (entry codes 1ck4 and r1ck4sf).

2.3. Optical spectroscopy

Circular dichroism (CD) spectra were recorded at 10°C using a J-710 spectropolarimeter (JASCO, Japan). The CD spectrum of the rat α1-I domain was measured at a protein concentration of 55 μM in 20 mM HEPES, 1 mM EDTA, 1 mM DTT, pH 7.5 in the absence of divalent cations, or the presence of either 2 mM Ca^{2+}, Mg^{2+} or Mn^{2+}. CD spectra were recorded using a scan speed of 20 nm/min, a response time of 2 s and a bandwidth of 2 nm for far- and 1 nm for near-UV measurements. Typically, five spectra were accumulated and subsequently averaged. Neither noise reduction nor smoothing programs were applied to the spectra in the near-UV range.

UV absorption spectra were measured at 25°C using a DU 640 spectrophotometer (Beckman, USA). Protein concentration was determined using a molar absorptivity of 12 950 M^{-1} cm^{-1} at 280 nm as it was calculated assuming molar absorptivities of 5500 M^{-1} cm^{-1} per tryptophan and 1490 M^{-1} cm^{-1} per tyrosine residue [31]. The spectra were corrected for turbidity by plotting the dependence of the $\log_{10}$ of the absorbance of the solution versus the $\log_{10}$ of the wavelength and extrapolating the linear dependence between these quantities in the range 340–440 nm to the absorption range 240–300 nm. The extrapolated values were then subtracted from the measured values.

3. Results and discussion

3.1. Description of structure

We have expressed the I-domain of the rat α1 integrin chain in *E. coli* as a GST fusion protein. The recombinant protein bound type IV collagen and type I collagen in a metal ion-

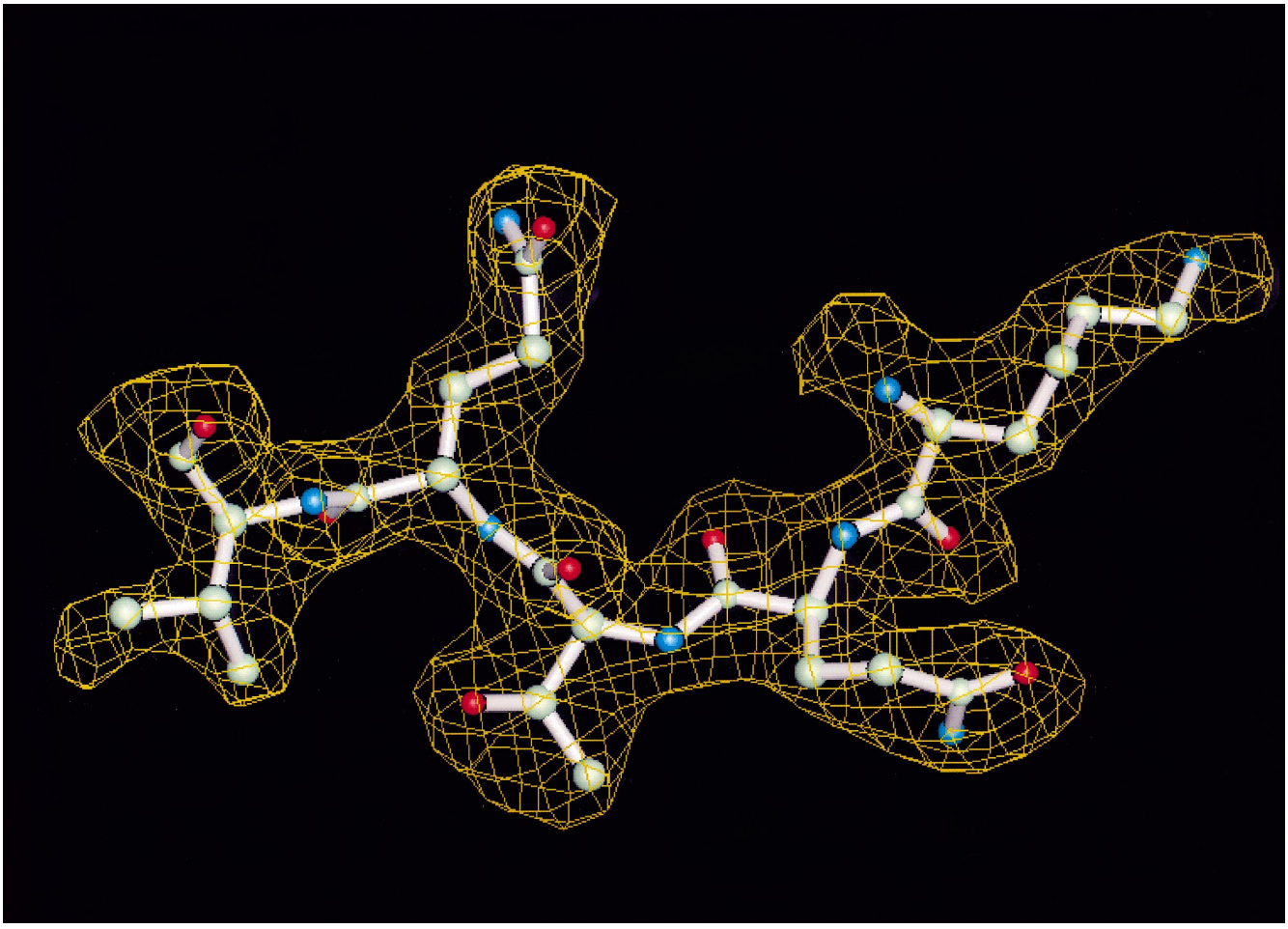

Fig. 1. Representative portion of the final 2Fo-Fc electron density map. The map is contoured at 1.0 σ and superimposed on corresponding atoms from the final refined model. The figure was made with QUANTA.

dependent manner and this interaction could be blocked with a neutralizing antibody raised against α1β1 or with EDTA (Gotwals et al., submitted for publication). Crystals of the α1 I-domain were grown and the structure was determined by molecular replacement at 2.2 Å resolution. The asymmetric unit of the crystal contains two I-domain molecules, referred to as molecules A and B. Almost all main chain amino acids, except for residues 143–144 of the N-terminus and 336–340 of the C-terminus, are well defined in the final electron density map (Fig. 1). There is no electron density for the side chain of residues Thr-145, Gln-146, Arg-234 of molecule A and Thr-145 and Arg-175 of molecule B. The current model consists of 386 amino acid residues and 215 water molecules. The stereochemistry of the model was analyzed with the program PROCHECK [32]. The Ramachandran diagram shows that all amino acid residues except Glu-192 have ϕ, ψ angles within the allowed regions. The root mean square (rms) deviations from ideal values for bonds and angles are 0.009 Å and 1.55°, respectively.

The α1 I-domain adopts the classic dinucleotide or Rossmann fold characterized by a central β-sheet consisting of five parallel and one antiparallel β-strands surrounded by seven amphipathic α-helices. The molecule presents the 'closed' conformation previously observed in most crystal forms of I-domains that have been solved to date [15]. The structure of the α1 I-domain is very similar to that of the α2 I-domain (sequence identity 49.5%). The rms positional variation of Cα atoms of superimposed α1 I and α2 I is 0.8 Å. The rms deviation between 1294 equivalent (conserved in both structures) atoms is 1.3 Å. The biggest differences between α1 and α2 I-domains are shifts in the C-terminal α7 helix that can reach up to 2 Å for the last few main chain atoms of the helix. The conformations of molecules A and B are very similar. The rms deviations in Cα positions between molecules A and B are 0.4 Å. The overall rms deviations are 1.0 Å. Most of the structural differences are rotameric shifts of surface side chains.

The structural differences between the two conformations of I-domain are dominated by the exposure to the solvent of two buried phenylalanines (Phe-275 and Phe-302 in the αM sequence) [12]. In the α1 I-domain crystal structure, a glutamic acid (Glu-321), located at the top of the α7 helix, replaces the Phe-302 of the αM I-domain. There is also a water molecule (191), not found in I-domains other than α2 I, which makes hydrogen bonds to Glu-321, the main chain nitrogen of His-288 and the MIDAS site Asp-257. However, in α1 I, this water molecule is observed only for molecule B. With respect to the second phenylalanine (Phe-275 of αM I), despite the large local structural differences between the αM I and α1 I due to a longer loop, there is a similar hydrophobic pocket in α1-I occupied by Phe-299. This Phe residue is located on the α6 helix, while Phe-275 in αM I is located on middle of the

βE-α6 loop. However, the phenyl ring of Phe-299 is in a very similar position to that of Phe-275 of αM I.

Unlike with other I-domains, the α1 I and α2 I structures contain a short 1.5 turn α-helix on the βE-α6 loop (the C-helix loop) that protrudes from the MIDAS face. The conformation of the C-helix loop (residues 286–296) of the α1 I is remarkably similar to that of the α2 I. The rms deviations between α1 I and α2 I C-helix loops are 0.78 Å for the Cα positions and 1.24 Å for all equivalent atoms. Tyr-285 of the α2 I is replaced with His-288 that adopts the same side chain conformation. Leu-286 of the α2 I is replaced by Tyr-289 that is stabilized by a hydrogen bond with the main chain nitrogen of Glu-259. These residues could play a role in different collagen specificity, since they are the only ones close to the MIDAS site that are different between the α1 and α2 I-domains.

3.2. The MIDAS site

The residues of the MIDAS site of the α1 I-domain (Asp-154, Ser-156, Ser-158, Thr-224 and Asp-257) are relatively well defined in the electron density map. The crystals of α1 I were grown in the absence of divalent cations. There is no electron density in the MIDAS site of both molecule A and B that would correspond to a divalent cation due to contaminants. A simulated annealing omit map, calculated by omitting all the residues within an 8 Å sphere from the expected cation position, did not reveal any electron density feature that could correspond to bound cation. There are only minor differences in the conformation of the MIDAS site residues to those of the α2 I structure with bound Mg^{2+}. Similarly, no major differences were observed in the MIDAS site of the αL I with and without cation [13]. Minor differences are observed, such as a rotation of the side chain of Asp-154 that results in Oδ1 of Asp-154 being 4 Å, instead of 2.6 Å, away from Ser-156 OH.

Of particular interest is a crystal contact interaction formed between the MIDAS site of molecule A and Arg-246 of molecule B (Fig. 2). The Nη1 atom of Arg-246 makes H-bonds with Ser-156 and Ser-158. The same atom also lies in a position very close to that of a water molecule that interacts with the Mg^{2+} ion in the α2 I structure. Only one of the three water molecules that would normally interact with the metal ion is present. The Nη2 atom of Arg-246 makes an H-bond with Thr-224. Thus Arg-246 seems to have some of the properties of a ligand mimetic and in the absence of divalent cations, the positive charge of the arginine may substitute the missing metal ion. Therefore this interaction might represent an intrinsic affinity of the MIDAS site for positively charged groups. This notion is consistent with the presence of an extensive negative electrostatic potential demonstrated with calculations using GRASP [33] (data unpublished). Indeed, interaction with a positively charged amino acid (Lys-181 of molecule B) is observed for molecule A too. While the presence of a positive charge in this site was not emphasized previously, on close examination of another available crystal structure of an I-domain without metal ion (αL I structure, PDB entry code 1zon) [13] we observed that the MIDAS site interacts with Lys-188 of a neighboring molecule. In particular, the Nζ1 atom of the Lys-188 forms H-bonds with the two serine residues of the MIDAS site in the similar manner as the Nη1 atom of Arg-246 in the α1 I structure.

Integrin ligands typically contain a critical acidic residue such as a glutamate or an aspartate. It has been proposed that the acidic residue coordinates the metal ion bound on the MIDAS motif during the ligand-integrin interaction [12]. Indeed, in a crystal structure of the αM I-domain, a glutamate residue from a neighboring molecule was observed to coordinate the Mg^{2+} ion in the MIDAS site [9]. It is striking that the I-domain of this particular crystal structure adopted an 'open' conformation different from all other observed I-domain structures, involving a significant shift of the C-terminal helix and other rearrangements of residues proximal to the MIDAS site. Because integrins are known to undergo activation, and this in part is likely to occur through conformational changes, it was proposed that this particular crystal structure represented the active state of the I-domain while the other structures represent the inactive state [12]. Since all other crystal structures of I-domains adopted the 'closed' conformation in the presence or absence of cation, it was hypothesized that the glutamate, acting as a ligand mimetic, and not the cation, induced the conformational change. No crystal structure of an I-domain complex with its ligand has been determined up to date, therefore this issue remains highly controversial.

While the proposed model [12] supports a state where ligand binding is through the metal ion in the MIDAS site, various studies suggest that the metal ion is either displaced upon ligand binding or in selected instances is not required for ligand binding at all. The following highlights some of these observations. In a recent study [34], the authors proposed a model in which the divalent cation is required for initial binding of the α2 I to collagen and that subsequent metal ion displacement occurs to generate a metal free I-domain ligand complex. The experiment involved use of Tb^{3+} fluorescence and demonstrated that Tb^{3+} supported ligand binding. We have observed similar effects with the α1 I-domain (unpublished data). These observations suggest the formation of a ligand-bound subsequent state of the I-domain in which the divalent cation has been displaced, a model that has also been proposed for the binding of RGD ligands to αIIbβ3 integrin [35]. The displacement of the cation may be a consequence of substitution of the acidic residue of the ligand and the metal ion that is bound to the MIDAS site, with a neighboring positively charged residue such as arginine. Interestingly, an arginine residue is frequently proximal to the acidic residue and essential for binding in many integrin ligands such as in RGD-containing proteins, collagen IV [21], collagen I [36] and VCAM-1 [37]. While the details of the mechanism leading to ligand-induced displacement of the metal ion are unknown, one possibility is that the model of interaction of the ligand acidic group with the cation in this 'activated' state [12] represents a transient intermediate in the process (Fig. 3). A prediction of this model is that the ligand-bound (I_L), metal-bound (I^+) and unoccupied form (I) are of similar energy states, which would facilitate the subsequent release of the ligand from the bound state and cycling of the process.

Several reports suggest that the I-domain can bind independently of cation, which is consistent with our model. In this situation the positively charged amino acid is acting as a ligand mimetic in the absence of cation. First, data for binding of recombinant α2 I-domain to collagen [4] suggest that divalent cations are not absolutely required for binding. It was proposed [11] that the iodination of a tyrosine residue of that study may favor a high affinity state of the I-domain, abrogating the need for a metal ion. In addition, in a recent muta-

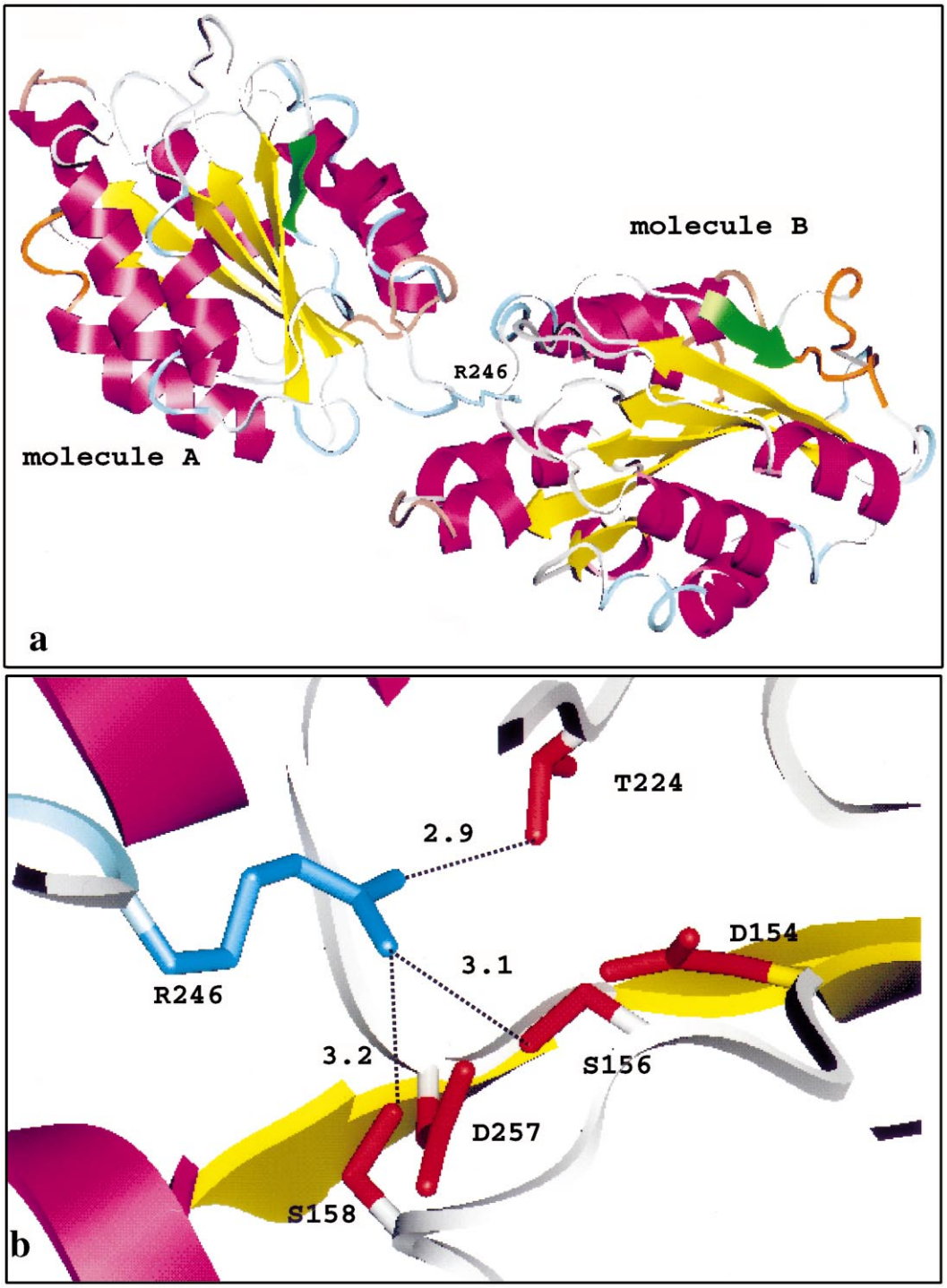

Fig. 2. The crystallographic dimer. a: Ribbon representation of the two I-domain molecules in the asymmetric unit. The Arg-246 residue involved in an interaction with the MIDAS site is also shown. b: Close-up of the Arg-246-MIDAS site interaction. Arg-246 is shown in blue and MIDAS side chains are shown in red. Hydrogen bonds are shown as dashed lines and lengths are indicated in Å. The figure was made with QUANTA.

genesis study, it was proposed that the I-domain exists in solution in equilibrium between two affinity states that can be modulated both positively or negatively by certain mutations [15].

Second, alanine substitution of MIDAS residues Asp-151 and Asp-254 of the α2 I-domain had minimal or no effect on collagen binding, while alanine substitution of Thr-221 resulted in loss of collagen binding [4]. It is interesting that the interaction of Arg-246 with the MIDAS site of α1 I involves an H-bond with the MIDAS threonine and no H-bonds with the two MIDAS aspartates.

Third, recently a cyclic peptide (CTRKKHDNAQC) was identified that competes with α2 I-domain binding to collagen [38]. Mutational analysis of the peptide showed that a sequence of three positively charged amino acids, arginine-lysine-lysine (RKK), is essential for I-domain binding and that

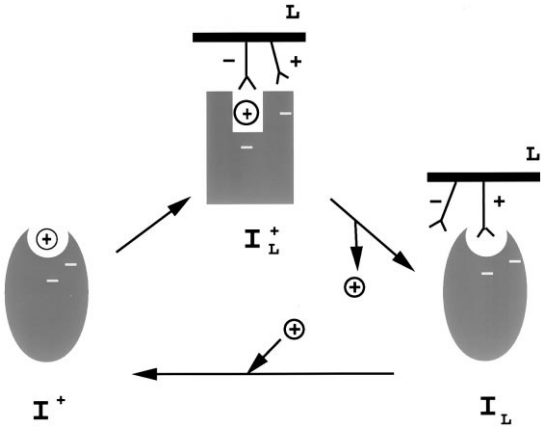

Fig. 3. Schematic of a hypothetical mechanism of ligand binding to I-domains. I = integrin I-domain, ⊕ = metal ion, L = ligand. The closed conformation of the I-domain is represented with an ellipsoid and the open with a rectangle.

mutation of the other amino acids in the peptide had little effect on binding. It is interesting that the aspartate residue of the peptide was not necessary for binding of the peptide to the I-domain.

Two other related systems provide further insights into the mechanism of I-domain function: vWF A-domains are structurally homologous to integrin I-domains and are known to bind to collagen. However, they do not require metal cation for collagen binding and their MIDAS site is only partially conserved [39,40]. The structure of the vWF A3 domain suggested that binding to collagen is achieved through interactions between negatively charged residues on A3 and positively charged residues on collagen. It is possible that the mechanism of binding of I-domain-containing integrins is evolutionarily related and shares some common aspects.

Multiple, allosterically controlled, activation states involving a cation displacement step are well-established events in the function of the, apparently closely related, G-proteins. I-domains, GTP binding proteins (such as Ras) and G_α subunits adopt a similar type of structural fold, which suggests that they diverged from a common ancestor. In addition, it has been proposed that the integrin α chains contain a propeller fold similar to G_β subunits of G-proteins [41]. Mechanistic similarities of G-proteins and I-domains in terms of Mg^{2+} binding and affinity switch mechanism have been observed and described [12]. The recent evidence of cation displacement in integrins resembles the cation displacement in G-proteins, which suggests further similarity with G-protein function. It is possible that integrins use such mechanisms to cycle between low and high affinity states resulting in rapid adhesion and de-adhesion events necessary for cell migration [42].

The glutamine-cation-MIDAS site interaction [9] and the arginine-MIDAS site interaction described here are both consistent with available experimental data and may mimic snapshots of a dynamic pathway of the integrin I-domain function. More detailed biochemical, genetic and crystallographic analysis is necessary to test these hypotheses.

3.3. Spectroscopy analysis

Because metals are involved in ligand binding and may effect I-domain structure, we investigated the effects of Mn^{2+}, Ca^{2+} or Mg^{2+} on the secondary and tertiary structure of the rat α1 I-domain in a comparative study by CD spectroscopy. Results from these studies are shown in Fig. 4a for far- and Fig. 4b for near-UV spectral range. The data from the far- and near-UV spectra are almost identical, within the error limit of the measurement, under all test conditions indicating that the secondary and tertiary structure are practically the same in the apo state and in the presence of divalent cations. The near-UV spectra are very sensitive to local protein structure in the vicinity of aromatic residues and therefore the similarity of these spectra further supports the results obtained by far-UV analysis. Thus, the CD data indicate absence of structural changes of the I-domain in solution due to divalent cation binding, supporting the crystallographic evidence [13,14].

In contrast to the lack of effect of metal binding on the CD

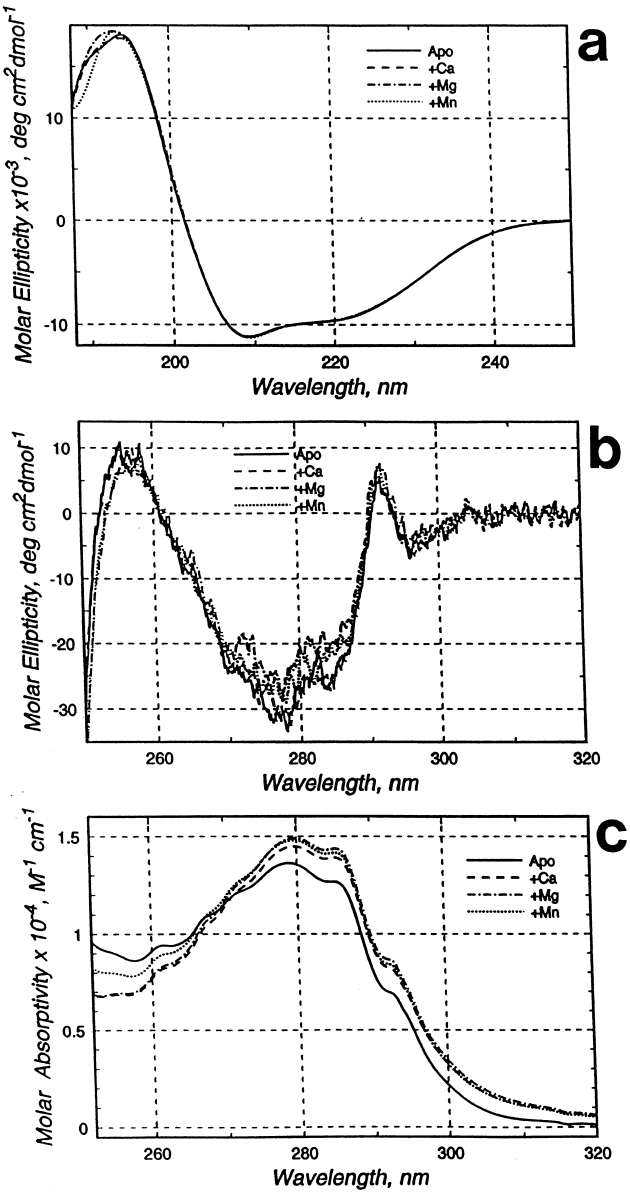

Fig. 4. Optical spectra of the α1 I-domain. a: CD spectrum in the far-UV range in the presence or absence of different divalent cations. b: CD spectrum in the near-UV range. c: UV absorption spectrum.

spectra, a small change in the absorbance spectra was detected (Fig. 4c). This suggests that in the presence of divalent cation, the environment surrounding Trp-162 is more hydrophobic than in the apo state as evident by a increase in both the intensity and I_{max}. This effect could not be accounted for by turbidity since the largest effect on the spectra was seen in the 280–290 nm spectral range while in the 240–260 nm range a decrease in intensity was seen. The crystal structure shows that Trp-162 is buried but located relatively close to the surface of the molecule and ~9 Å away from the expected position of the divalent cation. It also lies relatively close to the top of the α7 helix, the site that may undergo large conformational changes upon activation. Essentially no differences are observed in the conformation of this tryptophan residue in the structures of α1 I and α2 I (Mg^{2+}). Thus the changes in the spectra may be attributable to neutralization of charges due to cation binding.

In summary, we have determined the crystal structure of the α1 I-domain in the absence of metal ions and refined it to 2.2 Å resolution. We reviewed published evidence that I-domains can interact with ligands independently of metal ions or acidic residues and proposed a possible interpretation based on an interaction with an arginine residue observed in the α1 I crystal structure. We have also shown by CD spectroscopy that metal ion binding in solution does not affect the overall conformation of the I-domain.

Acknowledgements: We thank Dr. Robert Liddington (University of Leicester) for the coordinates of the α2 I-domain crystal structure prior to release to the Protein Data Bank; Dr. Eugene Marcantonio (Columbia) for the human α1 integrin cDNA clone; Dr. Michael Ignatious (University of California, Berkeley, CA) for the rat α1 integrin cDNA clone, and Dr. Roy Lobb (Biogen) for stimulating discussions. S.Y.V. is supported by NIH Grant GM34847 to Dr. Franklyn G. Prendergast (Mayo Foundation, Rochester, MN).

References

[1] Calderwood, D.A., Tuckwell, D.S., Eble, J., Kuhn, K. and Humphries, M.J. (1997) J. Biol. Chem. 272, 12311–12317.
[2] Champe, M., McIntyre, B.W. and Berman, P.W. (1995) J. Biol. Chem. 270, 1388–1394.
[3] Huang, C. and Springer, T.A. (1995) J. Biol. Chem. 270, 19008–19016.
[4] Kamata, T. and Takada, Y. (1994) J. Biol. Chem. 269, 26006–26010.
[5] Muchowski, P.J., Zhang, L., Chang, E.R., Soule, H.R., Plow, E.F. and Moyle, M. (1994) J. Biol. Chem. 269, 26419–26423.
[6] Rieu, P., Ueda, T., Haruta, I., Sharma, C.P. and Arnaout, M.A. (1994) J. Cell Biol. 127, 2081–2091.
[7] Ueda, T., Rieu, P., Brayer, J. and Arnaout, M.A. (1994) Proc. Natl. Acad. Sci. USA 91, 10680–10684.
[8] Tuckwell, D., Calderwood, D.A., Green, L.J. and Humphries, M.J. (1995) J. Cell Sci. 108, 1629–1637.
[9] Lee, J.-O., Rieu, P., Arnout, M.A. and Liddington, R. (1995) Cell 80, 631–638.
[10] Qu, A. and Leahy, D.J. (1995) Proc. Natl. Acad. Sci. USA 92, 10277–10281.
[11] Emsley, J., King, S.L., Bergelson, J.M. and Liddington, R.C. (1997) J. Biol. Chem. 272, 28512–28517.
[12] Lee, J.-O., Bankston, L.A., Arnout, M.A. and Liddington, R.C. (1995) Structure 3, 1333–1340.
[13] Qu, A. and Leahy, D.J. (1996) Structure 4, 931–942.
[14] Baldwin, E.T. et al. (1998) Structure 6, 923–935.
[15] Li, R., Rieu, P., Griffith, D.L., Scott, D. and Arnaout, M.A. (1998) J. Cell. Biol. 143, 1523–1534.
[16] Edwards, C.P., Champe, M., Gonzalez, T., Wessinger, M.E., Spencer, S.A., Presta, L.G., Berman, P.W. and Bodary, S.C. (1995) J. Biol. Chem. 270, 12635–12640.
[17] Hemler, M.E., Jacobson, J.G., Brenner, M.B., Mann, D. and Strominger, J.L. (1985) Eur. J. Immunol. 15, 502–508.
[18] Belkin, V.M., Belkin, A.M. and Koteliansky, V.K. (1990) J. Cell Biol. 111, 2159–2167.
[19] Dubond, J.-L., Belkin, A.M., Syfrig, J., Thiery, J.P. and Koteliansky, V.E. (1992) Development 116, 585–600.
[20] Kern, A., Eble, J., Goblik, R. and Kuhn, K. (1993) Eur. J. Biochem. 215, 151–159.
[21] Eble, J.A., Goblik, R., Mann, K. and Kuhn, K. (1993) EMBO J. 12, 4795–4802.
[22] Kern, A. and Marcantonio, E. (1998) J. Cell Physiol. 176, 634–641.
[23] Jancarik, J. and Kim, S.H. (1991) J. Appl. Crystallogr. 24, 409–411.
[24] Otwinowski, Z. (1993) in: Proceeding of the CCP4 Study Weekend: Data Collection and Processing (Sawer, L., Isaacs, N. and Bailey, S., Eds.), pp. 56–62, Daresbury Laboratory, Warrington.
[25] Matthews, B.W. (1968) J. Mol. Biol. 33, 491–497.
[26] Navaja, J. (1994) Acta Crystallogr. A 50, 157–163.
[27] Collaborative computational project No. 4, Acta Crystallogr. D 50, 760–763.
[28] Jones, T.A., Zou, J.Y., Cowan, S.J. and Kjeldgaard, M. (1991) Acta Crystallogr. A 47, 110–119.
[29] Brunger, A.T. (1992) Yale University Press, New Haven, CT.
[30] Adams, P.D., Pannu, N.S., Read, R.J. and Brunger, A.T. (1997) Proc. Natl. Acad. Sci. USA 94, 5018–5023.
[31] Pace, C.N., Vajdos, F., Fee, L., Grimsley, G. and Gray, T. (1995) Protein Sci. 4, 2411–2423.
[32] Laskowski, R.A., MacArthur, M.W., Moss, D.S. and Thornton, J.M. (1993) J. Appl. Crystallogr. 26, 283–290.
[33] Nicholls, A. (1992) Columbia University, New York.
[34] Dickeson, S.K., Bhattacharyya-Pakrasi, M., Mathis, N.L., Schlesinger, P.H. and Santoro, S.A. (1998) Biochemistry 37, 11280–11288.
[35] D'Souza, S.E. et al. (1994) Cell 79, 659–667.
[36] Knight, C.G. et al. (1998) J. Biol. Chem. 273, 33287–33294.
[37] Jones, E.Y. et al. (1995) Nature 373, 539–544.
[38] Ivaska, J., Kapyla, J., Pentikainen, O., Hoffren, A.M., Huttunen, P., Johnson, M.S. and Heino, J. (1999) J. Biol. Chem. 5, 3513–3521.
[39] Bienkowska, J., Cruz, M., Atiemo, A., Handin, R. and Liddington, R. (1997) J. Biol. Chem. 272, 25162–25167.
[40] Huizinga, E.G., van der Plas, R.M., Kroon, J., Sixma, J.J. and Gros, P. (1997) Structure 5, 1147–1156.
[41] Springer, T.A. (1997) Proc. Natl. Acad. Sci. USA 94, 65–72.
[42] Dransfield, L.C., Cabanas, C., Barrett, J. and Hogg, N. (1992) J. Cell. Biol. 116, 1527–1535.