Mechanistic Insights from a Refined Three-dimensional Model of Integrin αIIbβ3*

Marta Filizola‡*, Sergio A. Hassan‡, Andrea Artoni‡, Barry S. Coller‡, and Harel Weinstein‡*

From the ‡Department of Physiology & Biophysics, Weill Medical College of Cornell University, New York, New York 10021, the **Laboratory of Blood and Vascular Biology, The Rockefeller University, New York, New York 10021, and the †Center for Molecular Modeling, Division of Computational Bioscience (CMM/DCB/CIT), National Institutes of Health, Bethesda, Maryland 20892

The integrin αIIbβ3 plays an important role in platelet function, and abnormalities of this protein result in a serious bleeding disorder, known as Glanzmann thrombasthenia. Although crystallographic data exist for the related integrin αβ₃, to date, there are no high resolution structures of integrin αIIbβ3 available in the literature. Therefore, it is still unclear how specific elements of the αIIb subunit contribute to integrin αIIbβ3 function. Here we describe a refined model of the αIIb N-terminal portion of integrin αIIbβ3 obtained by using the αβ₃ template combined with a new method for predicting the conformations of the unique αIIb loop regions comprising residues 71–85, 114–125, and 148–164. The refined model was probed based on a structural prediction that differentiates it from standard homology models: specifically, that Lys-118 of αIIb, contacts Glu-171 of β3. To test this hypothesis experimentally, the mutant integrin chains αIIb K118C and β3 E171C were cotransfected into HEK 293 cells. We show that the cells expressed the mutants αIIbβ3 on their surface as a disulfide-linked dimer, supporting the close proximity between αIIb Lys-118 and β3 Glu-171 predicted from the refined model. This validated model provides a specific structural context for the analysis and interpretation of structure-function relations of integrin αIIbβ3. In addition, it suggests mechanistic hypotheses pertaining to both naturally occurring mutations responsible for Glanzmann thrombasthenia and to point mutations that affect ligand binding.

Integrin αIIbβ3, or glycoprotein Ib/IIaIIIa is the megakaryocyte-platelet-specific receptor that mediates the primary interactions of platelets with fibrinogen and other ligands, leading to platelet aggregation. Mutations in either the αIIb or the β3 subunit result in Glanzmann thrombasthenia, an autosomal recessive bleeding disorder affecting platelet function (see Refs. 1–3 for recent reviews). Although a large number of point mutations causing this disorder have been identified experimentally in the integrin αIIb β-propeller (sinaicentral.mssm.edu/intranet/research/glanzmann/menu), the mechanistic details of how these mutations affect receptor biogenesis, protein structure, and receptor function are not well understood.

To date, no high resolution structures of αIIbβ3 are available to provide a structural context for understanding the mechanism of action of this complex biological system. A three-dimensional model of this protein was first proposed from homology modeling using the β subunit of the G-protein transducin as a template (4). Subsequently, the crystal structures of a cognate template (αβ₃) (5, 6) with higher identity in sequence between αβ₃ and αIIb (40%) as compared with transducin (19%) enabled the construction of more accurate overall three-dimensional models of αIIbβ3 (7, 8) using standard homology-based modeling techniques. The αβ₃ crystallographic data support the folding of the N terminus of αIIb into a seven-blade β-propeller, as was originally predicted by Springer (9), and identify the propeller as the major site of interaction between αIIb and β3. However, αIIb differs from α in loop regions that are strategically located with regard to interaction with β3 and ligand binding, and standard homology-based modeling techniques do not usually provide accurate prediction of the structures of non-conserved loops. In particular, the αIIb N-terminal region contains three interacting non-conserved loop regions (residues 71–85, 114–125, and 148–164) that deserve special attention for their position in the integrin complex. Loops 114–125 and 148–164 are located at the interface between the αIIb and β3 subunits, and loop 71–85 is in proximity to, and thus potentially influences, loop 114–125.

In general, segments (or loop regions) connecting elements of defined secondary structure in proteins are likely to be the more flexible portions of the structure. This flexibility can impart significant functional roles to these loops, allowing their structural adaptation during interaction with other regions of the protein, with ligands, or with other molecules. For this same reason, however, these segments are also more difficult to predict by homology and, in many cases, are difficult to resolve by high resolution crystallographic techniques due to their large thermal fluctuations or crystal disorder (10). Thus, the three αIIb loops we have chosen for structural refinement are likely to contribute significantly to the conformational preferences may be difficult to establish even with the benefit of crystallographic data.

We present here a refined model of an initial αIIb structure obtained from a homology modeling protocol based on the αβ₃ crystal structure (5, 6). Specifically, an ab initio molecular mechanics structure prediction method recently developed (10) and evaluated (11, 12) in our laboratory was employed to explore the conformational space of the three flexible loop regions...
of \( \alpha_{IIb} \) described above. The method uses the CHARMM PAR22 force field in conjunction with the screened coulomb potential-implicit solvent model (SCP-ISM)\(^1\) for electrostatics. The refined model of the N-terminal portion of \( \alpha_{IIb} \) obtained by using this computational method was then validated experimentally. In particular, the intermolecular contact between Lys-118 of \( \alpha_{IIb} \) and Glu-171 of \( \beta_{IIb} \) predicted by our model but not a standard homology model, was supported by cross-linking of the corresponding cysteine mutants. In the absence of a high-resolution experimental structure of the \( \alpha_{IIb} \)-\( \beta_{IIb} \) complex, we propose the refined model presented here as a structural context to generate experimentally testable hypotheses and guide the design of future experiments to explore the functional roles of key structural elements.

**MATERIALS AND METHODS**

**Experimental Procedures**

Generation of cDNA Construct—pEF-\( \alpha_{IIb} \), pCDNA3.1-\( \beta_{IIb} \), pEF-\( \alpha_{IIb} \), pCDNA3.1-\( \beta_{IIb} \) R563C, and pCDNA3.1-\( \beta_{IIb} \) R563C were a generous gift from Drs. Junichi Takagi and Timothy Springer, Harvard Medical School. Vectors containing the cysteine mutants of \( \alpha_{IIb} \) Lys-118 and \( \beta_{IIb} \) Glu-171, pEF-\( \alpha_{IIb} \) K118C and pCDNA3.1-\( \beta_{IIb} \) R171C, were generated using the QuikChange XL site-directed mutagenesis kit (Stratagene, La Jolla, CA), as per the manufacturer's instructions.

**Stable Cell Lines Generation**—HEK 293 cells were transfected with either normal or mutant cDNA using PerFectin (Gene Therapy System, San Diego, CA) according to the manufacturer’s instructions. 48 h after transfection, cells were selected in media containing 500 \( \mu \)g/ml G418 for 2 weeks, and a population of clones expressing high levels of human and chimeric receptors, were labeled with the mAb 10E5 (\( \alpha_{IIb} \)-specific) (13) and sorted using a FACS caliber cell sorter (BD Biosciences).

**Flow Cytometry**—Cells (4-6 \( \times \) 10\(^5\) cells/ml) were incubated with Alexa Fluor 488-labeled 10E5 antibody (labeling as per the manufacturer’s protocol; Molecular Probes, Eugene, OR) (5 \( \mu \)g/ml) for 30 min at room temperature. Cells were washed, resuspended in phosphate-buffered saline containing 2% fetal bovine serum, and analyzed using a FACS caliber flow cytometer (BD Biosciences). Background controls were cells transfected with pEF-\( \alpha_{IIb} \) and incubated with Alexa Fluor 488-labeled 10E5.

**Biochemical Analysis of the Receptors**—To vectorially label surface molecules, 10\(^5\) transfected cells/ml were incubated on ice for 30 min with sulfo-N-hydroxysuccinimido biotin (1 mg/ml in phosphate-buffered saline) (Pierce). The reaction was stopped by adding glycine (5 mm final concentration), and then the cells were lysed with buffer (150 mm NaCl, 10 mm Tris-HCl, pH 7.5) containing 1% Triton X-100, 100 \( \mu \)m protease inhibitors (protease inhibitors mixture set III; Calbiochem). Lysates were centrifuged at 4 °C at 12,000 \( \times \) g for 30 min, and supernatants were precleared with protein A- or protein G-Sepharose (Amersham Biosciences) at 4 °C for 30 min. Immunoprecipitations were performed using mAb 10E5 (8 \( \mu \)g/ml). Protein G-Sepharose (5%) was added and incubated for 1 h at 4 °C. The beads were washed twice with lysis buffer containing 500 mm NaCl, and then the bound protein was eluted with SDS-PAGE sample buffer at 100 °C. Samples (reduced and non-reduced) were separated by SDS-PAGE (7.5% gel) and electro-transferred to polyvinylidene difluoride membranes. Membranes were blocked in 5% nonfat dry milk for 1 h at room temperature and washed. To identify biotin-labeled cell surface proteins immunoprecipitated by mAb 10E5, the membranes were hybridized with horseradish peroxidase-conjugated avidin (Amersham Biosciences) for 1 h at room temperature. Membranes were washed and developed using chemiluminescence, as per the manufacturer’s instructions (Amersham Biosciences).

**Computational Approach**

The refined three-dimensional model of the N-terminal portion of the \( \alpha_{IIIb} \) subunit of integrin \( \alpha_{IIb} \beta_{IIb} \) presented here was built in two steps. First, a standard homology modeling-based approach using the crystal structure of the cognate template \( \alpha_{IIb} \beta_{IIb} \) was employed to obtain an initial three-dimensional model of the \( \alpha_{IIIb} \)-\( \beta_{IIIb} \) \( \beta \)-propeller domain. Second, an \textit{ab initio} molecular mechanics structure prediction method was used to refine the three non-conserved loop regions (residues 71–85, 114–125, and 148–164 of human \( \alpha_{IIIb} \)) of the N-terminal portion of the \( \alpha_{IIIb} \) subunit discussed above.

**Construction of an Initial Three-dimensional Model of the \( \beta \)-Propeller Domain of Integrin \( \alpha_{IIIb} \beta_{IIIb} \)**—The initial three-dimensional model of the \( \alpha_{IIIb} \) N-terminal portion (residues 1–451) of \( \alpha_{IIIb} \beta_{IIIb} \) was based on homology with the \( \alpha_{IIIb} \) subunit (residues 1–438 of integrin \( \alpha_{IIb} \beta_{IIb} \) (sequence identity of 41%), for which atomic coordinates are available from x-ray crystallography using the MODELLER Release 9 (sali.org/modeller/modeller.html) (14, 15) and the sequence alignment shown in Fig. 1. The resulting model of \( \alpha_{IIIb} \beta_{IIIb} \) was then refined to replace the corresponding segment of \( \alpha_{IIIb} \) within the known structure of \( \alpha_{IIIb} \beta_{IIIb} \) in complex with a cyclic peptide containing the prototypical RGD sequence found in a number of integrin ligands (7) to obtain a three-dimensional model of the N-terminal portion of integrin \( \alpha_{IIIb} \beta_{IIIb} \), including \( \alpha_{IIIb} \beta_{IIIb} \) (1–451) and \( \beta_{IIIb} \) (109–352).

**Refinement of the Structure of \( \alpha_{IIIb} \beta_{IIIb} \)**—The three non-conserved loops 71–85, 114–125, and 148–164 of the N-terminal portion of integrin \( \alpha_{IIIb} \) were refined using the molecular mechanics ab initio computational method that we have recently developed to predict the structures of flexible loop regions in proteins (10). Details of this method and its application are reported elsewhere (10).

Briefly, the method employs simulated annealing Monte Carlo (SA-MC) simulations combined with biased scaled collective variables (SCV) Monte Carlo techniques (SCV-MC), which were specially designed to model segments with 8–13 amino acid residues that include parts of the internal secondary structural motifs to which the ligands are bound. The higher yield isotropic trial moves performed in the space of the scaled collective variables correspond to anisotropic trial moves in the space of the (real) dihedral angles (for details of the methodology, see Refs. 10 and 11). To incorporate solvent effects, the method uses a continuum electrostatic model based on screened coulomb potentials (SCP-ISM) (12). The screening function used in the SCP-ISM and the theoretical and experimental basis of its form have been discussed and illustrated in detail (12, 16, 17). In this ISM approach, the electrostatic representation of the protein in the solvent is determined by following a standard thermodynamic path in which the protein particles are first solvated individually and subsequently brought together to form the native protein. General expressions for the polar component of the solvation free energies for macromolecules, \( \Delta G_{\text{SASA}} \), are derived from this model. The work done in each step is evaluated, and the total electrostatic energy \( E_{\text{SCP}} \) of the macromolecule (or “effective energy” in the nomenclature of Ref. 18) is calculated (non-electrostatic effects are treated separately). The electrostatic energy is calculated from the equation,

\[
E_{\text{SCP}} = -\frac{1}{2} \sum_{i=1}^{n} \frac{q_i q_j}{r_{ij}} + \frac{1}{2} \sum_{i=1}^{n} \frac{q_i}{r_{i}} \left( \frac{1}{E(r_i)} - 1 \right)
\]

where \( q_i \) are the partial charges of the particle \( i \), \( n \) is the total number of atoms in the protein, \( D(r) \) are the screening functions of the electrostatic interactions within the protein in the solvent, and \( R_{\alpha} \) are the effective Born radii of the atom \( i \) in the protein, estimated as described in Refs. 12, 16, 17, and 19. Since a special treatment of hydrogen bonds has been shown to be necessary when continuum models are used (12, 16), the short range donor-acceptor interactions were stabilized in the SCP-ISM (independently of the strength of the long range electrostatic interactions that are controlled by the macroscopic screening \( D(r) \). The algorithm developed in the SCP-ISM stabilizes each single donor-acceptor partner available in the protein according to experimental hydrogen bond values, and a geometrical dependence of these interactions carried based on the type of hybridization of both acceptor and donor groups (16). Hydrophobic interactions (including entropic effects of the solvent and energy of cavity formation) are treated as a term proportional to the solvent-accessible surface area (SASA) of the molecule, i.e. \( E_{\text{SASA}} = \alpha + \beta \text{SASA} \), where \( \alpha \) and \( \beta \) are parameters usually obtained by fitting to solvation energies of hydrocarbons. \( E_{\text{SCP}} \) of the three loop regions were initially placed in their extended conformations. A logarithmic schedule

\( \frac{1}{2} \text{R}_{\alpha}^2 \)
was used for the annealing in which the temperature was decreased in 12 steps. Trial moves of 29,000, 22,500, and 30,000 per temperature for loops 71–85, 114–125, and 148–164, respectively, were performed by selecting at random two dihedral angles (backbone or side chain) of the segment. 10 independent runs were performed for each segment, and structures were collected at the final temperature of 300 K. Specifically, 1148 different conformations were identified for loop 148–164, 768 were identified for loop 114–125, and 1106 were identified for loop 71–85.

100 representative conformations were selected from each of these three sets based on root mean square deviation values calculated among all conformations within a specific set (for details of the procedure, see Ref. 10). These representative conformations were considered as potential candidates to fold into the native structure of the segment in the following steps of the protocol.

In the subsequent phase of the method, a first application of SCV-MC drove the C terminus of each representative loop conformation toward its final attachment point by using an adjustable harmonic constraint protocol imposed on the dummy residue attached to the C terminus. Starting with a force constant $k_0$, an SCV-MC simulation at $T_0=300$ K was carried out for each of the representative conformations obtained previously to relax the peptide around the closest local minimum in the free energy surface. This first simulation allowed the initial conformation of the segment to adjust to the characteristics of the new environment while partially preserving its intrinsic structural conformation calculated in the first step. The eigenvalues of the Hessian were updated twice to maintain the acceptance rate in the range 0.3–0.5; a total of 7000 SCV-MC steps were performed. Next, the harmonic constant $k$ was increased in successive steps to facilitate the shift of the C terminus toward the attachment point. The migration of the conformations of the segment from its intrinsic folding to its “closed” folding was simulated with the SCV-MC for each particular value of the harmonic constant $k$.

For each $k$, the simulation protocol was the same as described above for $k_0$, with an acceptance rate maintained in the range 0.3–0.5 in all cases. In previous calculations, it was observed that a power schedule of the form $k=10^{-m}k_0$, where $m$ is an integer that represents the successive increment of the constant $k (m=1, \ldots, 10)$, was appropriate in practice to slowly close the segments, and the same protocol was followed here.

After performing 250 minimization steps to eliminate bad contacts that could remain after the SCV-MC simulation, the closed conformations of each loop were sorted by energy. The lowest energy conformations within a 5 kcal/mol range were then selected for each loop. Specifically, the procedure yielded three different conformations for loop 71–85, 5 for loop 114–125 and six different conformations for loop 148–164 within this energy range. Assembly of all of these conformations into the protein yielded 90 different conformations of the three interacting loops in the IIb-(1–451) subunit. A second application of SCV-MC, using the same protocol described above for the first SCV-MC simulation, allowed relaxation of the three interacting loops in the environment created by the native protein and the solvent, where all the loops could move and interact simultaneously. This step produced the optimized structure of the IIb-(1–451) subunit with the three refined loops, which was then used to replace the initial model of IIb in the N-terminal portion of the IIbβ3 complex. Specifically, this optimized structure corresponded to the lowest energy conformation among the 90 structures that were relaxed using SCV-MC.

The three interacting loops of the IIb-(1–451) optimized structure were subjected to a final relaxation process in the presence of the
and \( \beta_3 \) Glu-171 to cysteines, the resulting heterodimer was expressed on HEK 293 cells, as judged from mAb 10E5 binding. The level of 10E5 binding to the mutant construct was comparable with that of cells transfected with normal human \( \alpha_{IIb} \) and \( \beta_3 \) or cells transfected with \( \alpha_{IIb} \) R320C and \( \beta_3 \) R563C.

When biotin-labeled surface proteins from cells transfected with normal human \( \alpha_{IIb}\beta_3 \) were immunoprecipitated with mAb 10E5 and then analyzed by SDS-PAGE, the \( \alpha_{IIb} \) and \( \beta_3 \) subunits demonstrated their typical patterns of migration under non-reducing conditions (\( \alpha_{IIb} \sim 140 \text{ kDa}, \beta_3 \sim 90 \text{ kDa} \)) and reducing conditions (\( \alpha_{IIb} \sim 120 \text{ kDa}, \beta_3 \sim 110 \text{ kDa} \)). In contrast, and consistent with the previous report by Takagi et al. (20), the \( \alpha_{IIb} \) R320C-\( \beta_3 \) R563C heterodimer migrated under non-reducing conditions as a dimer of \( \sim 220 \text{ kDa} \); with reduction of disulfide bonds, the dimer disappeared, and the individual subunits migrated at approximately their expected positions (\( \sim 120 \) and \( 110 \text{ kDa} \), respectively). Therefore, the dimer obtained under non-reducing conditions likely reflects a disulfide bond between \( \alpha_{IIb} \) R320C and \( \beta_3 \) R563C. The results with \( \alpha_{IIb} \) K118C and \( \beta_3 \) E171C are nearly identical to those with \( \alpha_{IIb} \) R320C and \( \beta_3 \) R563C, with the presence of a dimer under non-reducing conditions and the individual chains under reducing conditions. To assess whether the cross-linking was actually occurring between the \( \alpha_{IIb} \) Cys-118 and \( \beta_3 \) Cys-171, we tested the ability of antibody 7E3 to immunoprecipitate the complex since we have recently localized the 7E3 epitope to an adjacent region in \( \beta_3 \) including the Cys-177–Cys-184 loop and Trp-129 (21). 7E3 was able to precipitate \( \alpha_{IIb}\beta_3 \) from cells expressing normal \( \alpha_{IIb} \) and \( \beta_3 \) and cells expressing the combination of \( \alpha_{IIb} \) R320C and \( \beta_3 \) R563C mutants. It was not, however, able to precipitate the \( \alpha_{IIb} \) Cys-118-\( \beta_3 \) Cys-171 mutant protein. Similarly, cells expressing the \( \alpha_{IIb} \) Cys-118-\( \beta_3 \) Cys-171 mutants did not adhere to immobilized fibroinogen, whereas cells expressing native \( \alpha_{IIb}\beta_3 \) and mutant-expressing cells treated with 2 mM dithiothreitol adhered to immobilized fibrinogen (data not shown). These data support the conclusion that the disulfide cross-linking between \( \alpha_{IIb} \) Cys-118-\( \beta_3 \) Cys-171 occurred between the two newly introduced cysteines and thus reinforce the prediction from our refined model, but not the standard homology model, that \( \alpha_{IIb} \) Lys-118 and \( \beta_3 \) Glu-171 are in close proximity in the normal, solvated molecule.

**RESULTS**

Structural Features of the Refined \( \alpha_{IIb}-(1-451) \) Model—Fig. 2 shows a superposition of the \( \alpha_{IIb}\beta_3 \) crystal structure (in magenta) in complex with the RGD peptide (colored by atom code) with our proposed model of the \( \alpha_{IIb}-(1-451) \) complex (green for \( \alpha_{IIb} \) and red for \( \beta_3 \)). The three refined interacting loops of the \( \alpha_{IIb} \) subunit are shown in brown (loop 71–85), orange (114–125), and blue (148–164).

The \( \beta_3 \)-(109–352) subunit by a third application of SCV-MC, using the same protocol reported above. The protein side chains of the resulting conformation of the \( \alpha_{IIb}-(1-451) \) complex were then energy-minimized using a few steps of steepest descent followed by conjugated gradient.

**DISCUSSION**

We present a refined model of the N-terminal portion of the \( \alpha_{IIb} \) subunit of integrin \( \alpha_{IIb}\beta_3 \) that differs significantly from the homology models built on the \( \alpha_\alpha \) template (7, 8). The refinement involves the structures of the three non-conserved loop regions (residues 71–85, 114–125, and 148–164 of human \( \alpha_{IIb} \)) that deserve special attention for their strategic positions in the integrin complex, where they are likely to play functional roles in both heterodimer formation with the \( \beta_3 \) subunit and ligand binding. The refinement was obtained using an \( ab \) \textit{initio} molecular mechanics structure prediction approach based on energy calculations and simulation protocols developed recently in our laboratory (10) and evaluated for various molecular systems (11, 12).

Characteristics of the Loop Refinement Approach—The novel method we used for loop structure prediction reflects the key requirements for reliability and predictive power based on energy calculations: (i) the force field used for the calculation of energies must be able to discriminate among the many possible local minima corresponding to the different conformations of the segments in the context of the native protein and (ii) the sampling method must be able to reveal the existence of those minima despite the formidable challenges that occur as the length of the segment increases. The computational method used here to refine the N-terminal portion of \( \alpha_{IIb} \) was originally...
designed to address both of these requirements to model segments that connect regions with defined secondary structures in proteins (10). The unique features of this method, which was shown to reproduce the proper folding at the two ends of the segments as well as the correct H-bond pattern in test case peptides (e.g. see Ref. 10), are outlined below.

Unlike other loop closure techniques, the method used here is based on the assumption that segments connecting elements of secondary structure in proteins, in particular loops, have an intrinsic propensity for a particular set of conformations based on their amino acid sequence, but this intrinsic folding has to be disrupted for the best fit of the segment within the tertiary structure of the native protein. Thus, the final folding of the segments in the context of the protein is a compromise between two opposite effects: the intrinsic tendency to adopt a specific folding pattern dictated by the amino acid sequence and the partial unfolding that is imposed by the inclusion of the loop in the native conformation of the protein. The sampling methodology employed by our method follows these two processes in a rational and efficient way (see details in Ref. 10). Thus, it first
uses SA-MC simulations to find conformations that are representative of the segment structure in solution, as encoded in the primary sequence, and subsequently forces a slow unfolding of the segment to fit the final protein conformation using an adjustable force constant scheme and MC simulations with a scaled collective variables technique. The scaled collective variables technique allows the MC simulation to improve the efficiency of the search. In fact, the strong anisotropy of the free energy surface of peptides and proteins makes a standard MC simulation with a Metropolis algorithm highly inefficient due to the high probability of rejecting trial moves, especially in the neighborhood of a local minimum. If plain MC simulation were performed for such cases, small changes in each trial conformation would have to be adopted to keep the acceptance rate high enough for statistical significance and correct convergence in the calculation of thermodynamic quantities. This requirement restricts the exploration of the conformational space, compromising the quality of the results. However, it was previously shown (22) that the sampling can be improved if the trial moves are chosen anisotropically in the conformational space, giving more preference to the movements along the soft directions (low energy barriers) and reducing the movement along the hard directions (rough surface due mainly to steric effects). The scaled collective variables technique makes this possible by allowing the MC simulation to sample the space in this anisotropic way, and therefore, it improves the efficiency of the search we perform. Finally, since an accurate force field for the study of peptide and protein conformational preferences must account for the hydrophobic and electrostatic effects of the solvent, our method uses a continuum electrostatic model based on screened coulomb potentials, which has been validated in a number of systems ranging in size from small molecules to large proteins (12, 16, 17, 19).

Mechanistic Implications of Structural Details Revealed by the Refined Model

The refined model of the N-terminal portion of α₁β₃ obtained by using the computational method discussed above was supported experimentally by demonstrating that cysteine mutants of Lys-118 of α₁β₃ and Glu-171 of β₃ that are predicted from our model, but not a standard homology model, to be in close proximity resulted in a disulfide-bonded dimer of α₁β₃. To provide further support for the location of the cross-link responsible for heterodimer formation, we took advantage of our recent localization of the antibody 7E3 epitope to the region near β₃ Glu-171 (21) and demonstrated that the cross-link led to the selective loss of 7E3 binding. Although these data support our model, we recognize that they only

| α₁β₃ residue | Internal | Exposed | Interface | mAb bindingᵃ | β₃ partner(s)ᵇ |
|-------------|----------|---------|-----------|--------------|----------------|
| Asp-74      | ×        |         |           |              | LMA            |
| Leu-64      |          | ×       |           |              | LMA            |
| Asn-114     |          |         |           |              | LMA            |
| Glu-117     |          |         | ×         | ×            | LMA            |
| Lys-124     |          |         |           |              | NLMA           |
| Thr-125     |          |         |           |              | Ile-167, Ser-168, Pro-169 |
| Tyr-155     |          |         |           |              | LMA            |
| Phe-160     |          |         |           |              | LMA            |
| Asp-163     |          |         |           |              | LMA, NLMA      |
| Lys-164     |          |         |           |              | LMA, NLMA      |

ᵃ Indicated residues, when mutated to alanine, resulted in decreased binding of one or more ligand-mimetic antibodies (LMA) or one or more non-ligand mimetic antibodies (NLMA).
ᵇ The proposed interacting β₃ partners (residues of β₃ in which backbone and/or side chain atoms are within 5 Å of residues at the subunit interface) are reported for α₁β₃ residues at the interface between subunits.

Fig. 5. Residues that affect ligand binding (magenta and orange) in the context of our refined α₁β₃ (1–451) model (in green) and the β₃ subunit (in red). In particular, residues in orange belong to the loops refined here.
indicate that the cysteine-mutated residues Lys-118 of α\textsubscript{IIb} and Glu-171 of β\textsubscript{3} come close to each other sometime during bio-
genesis to form the disulfide bond.

In addition to the α\textsubscript{IIb}−Lys-118−β\textsubscript{3} Glu-171 interaction, our refined model differs from a standard homology-based model in a number of ways. Most importantly, the refined model predicts that α\textsubscript{IIb} residues Leu-116, Lys-124, and Arg-153 are close to one or more residues of the β\textsubscript{3} SDL region (residues 159−188). Specifically, β\textsubscript{3} Pro-170 is within 5 Å of α\textsubscript{IIb} Leu-116; the β\textsubscript{3} Ile-167, Ser-168, and Pro-169 are within 5 Å of α\textsubscript{IIb} Lys-124; and β\textsubscript{3} Tyr-166, Asp-179, and Met-180 are within 5 Å of α\textsubscript{IIb} Arg-153. These interactions are likely to account in large part for the ligand specificity of α\textsubscript{IIb}β\textsubscript{3} relative to α\textsubscript{IIb}β\textsubscript{3} since the shorter corresponding loops of α\textsubscript{IIb} do not interact with the β\textsubscript{3} SDL. Interestingly, replacement of one of these residues, Leu-
116, by a valine results in the Glanzmann thrombasthenia phenotype (23). Since in our model, but not the standard ho-

The refined and experimentally vali-
dated three-dimensional model of the N-terminal portion of integrin α\textsubscript{IIb}β\textsubscript{3} proposed here provides a structural context for the interpretation of elements of the α\textsubscript{IIb} subunit of integrin α\textsubscript{IIb}β\textsubscript{3} that are critical for biogenesis and function. In particu-
lar, analysis of the model offers testable mechanistic hypothe-
ses for the role of residues within three refined α\textsubscript{IIb} loops that are involved in ligand binding and subunit association; it also allows for analysis of naturally occurring mutations that pro-
duce Glanzmann thrombasthenia. Large-scale molecular dy-
namics simulations of the proposed model are currently ongo-
in our laboratory to investigate the dynamic effects of the proposed intermolecular interactions on the integrin structure.

This model has been deposited in the Protein Data Bank with the accession code 1RN0.

REFERENCES

1. Coller, B. S., and French, D. L. (2000) in Williams Hematology (Beutler, E., Lichtman, M. A., Coller, B. S., Kipps, T. J., Seligsohn, U., eds) 6th Ed., 1551–1582, McGraw-Hill, Inc., New York
2. Nair, S., Ghosh, K., Kulkarni, B., Shetty, S., and Mohanty, D. (2002) Platelets (Edinburgh) 13, 387–393
3. Belucci, S., and Caen, J. (2002) Blood Rev. 16, 193–202
4. Kamata, T., Tieu, K. R., Irie, A., Springer, T. A., and Takada, Y. (2001) J. Biol. Chem. 276, 42474–42483
5. Xia, J.-P., Stehle, T., Dardenne, B., Zhang, R., Dunker, R., Scott, D. L., Joachimiak, A., Goodman, S. L., and Arnaout, M. A. (2001) Science 294, 339–345
6. Xia, J.-P., Stehle, T., Zhang, R., Joachimiak, A., French, M., Goodman, S. L., and Arnaout, M. A. (2002) Science 296, 151–155
7. Arnaout, M. A., Goodman, S. L., and Xiong, J.-P. (2002) Curr. Opin. Cell Biol. 14, 641–651
8. Mitchell, W. B., Li, J. H., Singh, F., Michelos, A. D., Bussel, J., Coller, B. S., and French, D. L. (2003) Blood 101, 2268–2276
9. Springer, T. A. (1997) Proc. Natl. Acad. Sci. U. S. A. 94, 65–72
10. Hassan, S. A., Mehl, E. L., and Weinstein, H. (2002) Lecture Notes in Computational Science, pp. 197–231, Springer-Verlag New York Inc., New York
11. Mehl, E. L., Periole, X., Hassan, S. A., and Weinstein, H. (2002) J. Comput. Aided Mol. Des. 16, 841–853
12. Hassan, S. A., Guarnieri, F., and Mehl, E. L. (2000) J. Phys. Chem. B 104, 6478–6489
13. Coller, B. S., Peerschke, E. I., Scudder, L. E., and Sullivan, C. A. (1983) J. Clin. Invest. 72, 325–338
14. Fiser, A., Do, R. K., and Sali, A. (2000) Protein Sci. 9, 1753–1773
15. Sali, A., and Blundell, T. L. (1995) J. Mol. Biol. 324, 719–755
16. Hassan, S. A., Guarnieri, F., and Mehl, E. L. (2000) J. Phys. Chem. B 104, 6490
17. Hassan, S. A., and Mehl, E. L. (2002) Proteins 47, 45–61
18. Lazaridis, T., and Karpus, M. (2000) Curr. Opin. Struct. Biol. 10, 139–145
19. Hassan, S. A., Mehl, E. L., Zhang, D., and Weinstein, H. (2000) Proteins 51, 109–125
20. Takagi, J., Petre, B. M., Walz, T., and Springer, T. (2002) Cell 110, 599–611
21. Artoni, A., Li, J., Mitchell, W. B., Ruan, J., Coller, B. S., and French, D. L. (2002) Blood 100, 64b
