MOLECULAR DYNAMICS SIMULATIONS OF BINDING, UNFOLDING, AND GLOBAL CONFORMATIONAL CHANGES OF SIGNALING AND ADHESION MOLECULES

A Dissertation
Presented to
The Academic Faculty

by

Wei Chen

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
Woodruff School of Mechanical Engineering

Georgia Institute of Technology
May 2009

MOLECULAR DYNAMICS SIMULATIONS OF BINDING,

UNFOLDING, AND GLOBAL CONFORMATIONAL CHANGES OF

SIGNALING AND ADHESION MOLECULES

Approved by:

Dr. Cheng Zhu, Advisor
Department of Biomedical Engineering
Georgia Institute of Technology

Dr. Stephen Harvey
School of Biology
Georgia Institute of Technology

Dr. Ting Zhu
School of Mechanical Engineering
Georgia Institute of Technology

Dr. Nicholas V. Hud
School of Chemistry
Georgia Institute of Technology

Dr. Evan Zamir
School of Mechanical Engineering
Georgia Institute of Technology

Date Approved: March 23, 2009

To my parents and my wife

ACKNOWLEDGEMENTS

Firstly, I would like to thank my advisor, Dr. Cheng Zhu, for his advice, guidance, and support during my PhD study. He has been leading me into this exciting field that I was not familiar with at all six years ago and teaching me how to perform scientific research.

Secondly, I acknowledge Drs. Ting Zhu, Evan Zamir, Stephen Harvey, and Nicholas V. Hud for serving on my thesis committee and giving advice. Also thank Dr. Marc E. Levenston, who was serving on my thesis committee but left Georgia Tech. I specially thank Dr. Harvey for giving a lot of advice and providing computer resources for part of computations in this work.

Thirdly, I would like thank Dr. Jizhong Lou for teaching me MD simulations and giving lots of advice. I also thank Dr. Fabien Cailliez for teaching me the MM-PB/GBSA method. Thanks to my collaborators, Dr. Klaus Schulten and his student Jen Hsin, for their help and suggestions on MD simulations of integrins.

Furthermore, I thank Dr. Ning Jiang for providing me CD16 affinity data. Also, thank Dr. Periasamy Selvaraj for answering my questions about CD16. I acknowledge Dr. Ingeborg Schmidt-Krey for teaching me electron microscopy (EM) and her postdoctor Dr. Gengxiang Zhao for helping me with EM experiments. Thank Dr. Allen Orville and his students, especially Patrick Shaw, for teaching me X-ray crystallography.

Thank Tao Wu and Jiangguo Lin, Fang Kong, and Wei Jack Chen for providing experimental data for VWF, $\alpha_5\beta_1$ integrin, and $\alpha_L\beta_2$ integrin, respectively. Also, I thank

all the Zhu lab members. With them, I had a happy time during the past years and got lots of help and support.

I also acknowledge High Performance Computing Group in Georgia Tech and NSF TeraGrid for providing computer resources.

Finally, I greatly thank my parents, my parents in law, and my wife. Without their love and support, I would not be able to stand here and finish my PhD study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

2D	Two Dimension(al)
3D	Three Dimension(al)
ADAMTS-13	A Disintegrin And Metalloprotease with a ThromboSpondin type 1 motifs 13
AFM	Atomic Force Microscopy
BFP	Biomembrane Force Probe
COM	Center Of Mass
DOPE	Discrete Optimized Protein Energy
ECM	ExtraCellular Matrix
EM	Electron Microscopy
FcR	Fc Receptor
FcγR	Fcγ Receptor
GB	Generalized Born model
GPI	GlycosylPhosphatidylInositol
GPIb	Glycoprotein Ib
Ig	Immunoglobulin
IgG	Immunoglobulin G
ITAM	Immunoreceptor Tyrosine-based Activation Motif
ITIM	Immunoreceptor Tyrosine-based Inhibitory Motif
MD	Molecular Dynamics
MIDAS	Metal Ion-Dependent Adhesion Site
MM-PB/GBSA	Molecular Mechanics-Poisson Boltzmann/Generalized Born Surface Area
NMR	Nuclear Magnetic Resonance

PB	Poission-Boltzmann equation
PDB	Protein Data Bank
PME	Particle Mesh Ewald
RMSD	Root Mean Square Deviation
RMSF	Root Mean Square Fluctuation
SASA	Solvent Accessible Surface Area
SMD	Steered Molecular Dynamics
TTP	Thrombotic Thrombocytopenia Purpura
ULVWF	UltraLarge von Willebrand Factor
VWD	von Willebrand Disease
VWF	von Willebrand Factor

SUMMARY

Molecular dynamics simulations were used to investigate the structural basis for the functions of three proteins: Fcγ receptor III (CD16), von Willebrand factor (VWF), and integrin.

CD16, a heavily glycosylated protein expressed on human immune cells, plays a crucial role in immune defense by linking antibody-antigen complexes with cellular effector functions. Experimental data demonstrated that glycosylation of CD16 decreases its affinity for IgG. To understand the underlying structural basis, molecular dynamics (MD) simulations were performed for a CD16-IgG Fc complex with or without an N-glycan attached *in silico* on residue Asn159 of CD16. The simulated complex structures show different conformations for the two forms of CD16, which induces changes in the IgG Fc binding pocket. Besides, CD16 itself was observed to transition between two conformations. In order to evaluate quantitatively the changes of CD16-IgG binding upon CD16 glycosylation, Molecular Mechanics-Poisson Boltzmann/Generalized Born Surface Area (MM-PB/GBSA) approach was applied to calculate the binding free energy of the two CD16-IgG Fc complexes. Although they gave opposite results for the total binding free energy due to large errors of solute entropy calculation, PB and GB methods predicted the same tendency of the changes of the effective energy, in which solute entropy part is excluded. According to the decomposition of the effective energy on each residue, the critical residues for CD16-IgG binding were identified and the changes of residue-residue interactions due to CD16 glycosylation were clarified.

VWF, a multimeric multidomain glycoprotein, initiates platelet adhesion at sites of vascular injury. The three tandem A domains (A1, A2, and A3) of VWF play critical roles for its functions. The A1 and A3 domains contain respective binding sites for platelet glycoprotein Ib (GPIb) and collagen. The A2 domain hosts a proteolytic site for the VWF-cleavage enzyme A Disintegrin And Metalloprotease with a ThromboSpondin type 1 motifs 13 (ADAMTS-13). Previous studies suggested that shear flow or denaturants assists the ADAMTS-13 cleavage of VWF by unfolding the A2 domain and thus exposing the cryptic proteolytic site. Here steered MD (SMD) and free MD were used to simulate the unfolding of A1 and A2 by tensile force and under high temperature, respectively. The forced unfolding of A2 started from the C-terminus because of its specific topology. The β-strands of A2 were pulled out sequentially, generating sawtooth-like peaks in the force-extension curves. The disulfide bond between A1 N- and C-termini prevented it from being unfolded under low force. After eliminating the disulfide bond, A1 was unfolded similarly as A2 in terms of the β-strand pullouts, but differed in the unfolding of helices. The major resistance of A1 and A2 to unfolding came from the hydrogen bond networks of the central β-sheets. Two different unfolding pathways of the β-strands were observed, where the sliding pathway encountered much higher energy barrier than the unzipping pathway. The thermal unfolding of A1 and A2 was very different from the unfolding by tensile force. At 500K, A2 was unfolded from the C-terminal portion, but A1 was unfolded from the N-terminal portion, reflecting their sequence variations and thus stability difference.

Integrins mediate cell-cell, cell-extracellular matrix, and cell-pathogen adhesions essential in a wide variety of physiological and pathological processes. Experiments have

suggested that integrins undergo a large-scale global conformational change from a bent to an extended conformation upon activation. However, experiments show only static structures of integrins, and a complete understanding of the dynamic process of integrin conformational changes is still lacking. Here, SMD was applied to simulate in atomic details how integrins undergo structural transition from the bent to the extended conformation in response to tensile force. Homology models were first constructed for the unresolved I-EGF1/2 domain in the bent $\alpha_V\beta_3$ crystal structure to obtain a model of the complete ectodomain. Then, in an SMD simulation with a full, solvated integrin, a force was applied on the βA domain while the βTD domain was constrained to unbend integrin gradually without domain unfolding. A major force peak was observed, corresponding to the breakage of interactions at the Hybrid/EGF4 and Hybrid/βTD interfaces. Furthermore, free MD simulations were carried out to test stabilities of different conformations along the unbending pathway. Two partially extended conformations, existing immediately after the major force peak observed in SMD, were shown to be unstable and tended to bend back. One of them almost returned to the bent conformation. In contrast, two fully extended conformations was observed to be stable in free MD, their stabilities possibly due to the participation of residue Asp457 of the Thigh domain in the Ca^{2+} ion coordination at α-genu.

In order to test whether leg separation affects the unbending of integrins as proposed by the switchblade model, SMD simulations were run to separate the two lower legs of the simulated $\alpha_V\beta_3$ integrin. A much larger force peak than that in the unbending simulations indicated that leg separation was much harder because of more extensive contacts between the two legs. After leg separation, more interactions were formed

between the headpiece and tailpiece of β-subunit. As a result, two additional force peaks were observed during the unbending with separated legs, showing more resistance. The free MD simulations for different conformations with separated legs showed contrary results to those with closed legs: a partially extended conformation did not return to its bent conformation and a fully extended conformation was not stable and bent back.

CHAPTER 1

OBJECTIVES AND SPECIFIC AIMS

The overall objective of this thesis is to reveal structural basis by molecular dynamics simulations for the functions of three proteins: Fcγ receptor III (CD16), von Willebrand factor (VWF), and integrin.

1. Study Structural Basis for the Glycosylation Effects on CD16-IgG Binding

CD16-IgG binding acts as a linkage between the humoral and the cellular immune responses. Glycosylation of CD16 modulates CD16-IgG binding affinity and thus immune responses. It is unclear how the glycosylation of CD16 affects the CD16-IgG binding. To understand the structural basis for the glycosylation effects, first, molecular dynamics (MD) simulations were used to study the structures and dynamics of CD16-IgG complexes with or without CD16 glycosylation. Second, Molecular Mechanics-Poisson Boltzmann/Generalized Born Surface Area (MM-PB/GBSA) approach was utilized to calculate the binding free energy of CD16-IgG complexes and provide energetic information for CD16-IgG binding.

2. Investigate the Unfolding Pathways of VWF A Domain under Tensile Force or High Temperature

VWF A2 domain contains a cryptic proteolytic site for a plasma protease, <u>A</u> <u>D</u>isintegrin <u>A</u>nd <u>M</u>etalloprotease with a <u>T</u>hrombo<u>S</u>pondin type 1 motifs 13 (ADAMTS-13). Only after A2 unfolding, ADAMTS-13 can access the proteolytic site and cut ultralarge VWF (ULVWF) into short pieces to generate the full spectrum of circulating

plasma VWF species. To understand how ADAMTS-13 binds and cleaves the A2 domain, it is necessary to investigate how A2 is unfolded. First, steered molecular dynamics (SMD) was performed to simulate the unfolding of A2 by a tensile force, which is similar to the condition of shear flow in physiologic environment. As a comparison, the homologous A1 domain of VWF was also unfolded by SMD simulations. Second, MD simulations at 500 K were run to study the unfolding of A1 and A2 by thermal force, which mimic the unfolding by denaturants (e.g., urea) that was shown experimentally to increase ADAMTS-13 cleavage of VWF.

3. Simulate the Unbending Process of $\alpha_V\beta_3$ Integrin Induced by Force

Integrins mediate cell-cell, cell-extracellular matrix, and cell-pathogen adhesion and transmit signals bi-directionally across cell membranes. During signaling, integrins undergo a large-scale global conformational change from a bent to an extended conformation. To understand how integrin functions are regulated by its conformations, it requires studying the transitions between different conformations. Starting from a bent crystal structure of $\alpha_V\beta_3$ integrin, a complete ectodomain structure was obtained by building homology models of two unresolved domains. Then, SMD simulations were used to induce the unbending of $\alpha_V\beta_3$ integrin from the bent to the extended conformation. Stabilities of different conformations along the unbending pathway were tested. Furthermore, with separation of the legs of $\alpha_V\beta_3$ integrin that mimics inside-out signaling, the relation between extension and leg separation was studied.

CHAPTER 2

BACKGROUND

2.1 CD16-IgG Binding

Roles of Fc Receptors in Immune Responses
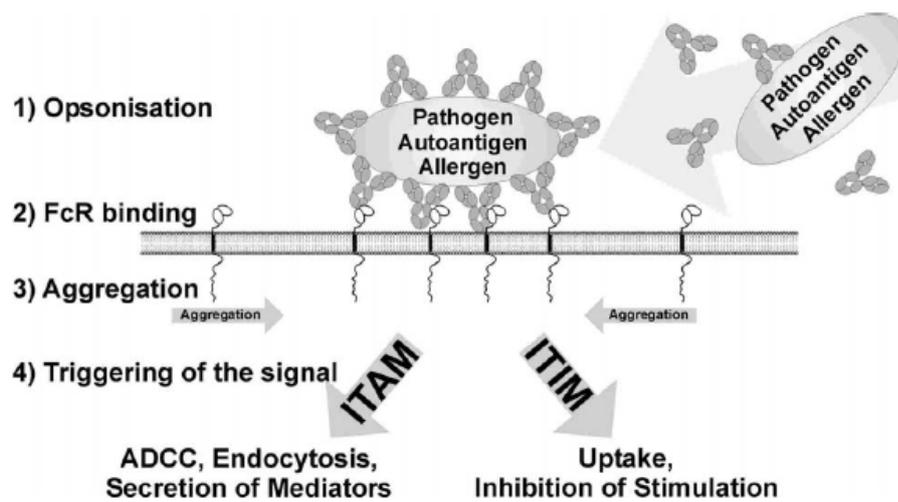
Figure 2.1 Functions of Fc receptors. Fc receptors expressed on immune cells connect antibody-mediated humoral with cellular immune responses. Figure adapted from Sondermann et al. [1]

Fc receptors (FcRs) are membrane proteins expressed on immune competent cells. They play critical roles in an antibody-mediated immune response: Firstly, immunoglobulins (Igs or antibodies) circulating in blood recognize and bind antigens to form immune complexes; Secondly, these immune complexes are captured by FcRs on cell membranes; Thirdly, FcRs aggregate; Finnally, through intrinsic Immunoreceptor Tyrosine-based Activation Motif (ITAM) or Immunoreceptor Tyrosine-based Inhibitory Motif (ITIM) or associated ITAM containing polypeptides, FcRs trigger intracellular signaling cascades, resulting in stimulation or suppression of an immune response (Fig. 2.1) [1].

Fcγ receptors (FcγRs) are one type of FcRs that specifically bind the Fc region of immunoglobulin G (IgG) [2]. FcγRs consist of three members: FcγRI, FcγRII, and FcγRIII. FcγRIII, also named CD16, is of interest in this thesis.

CD16 Isotypes and Structures
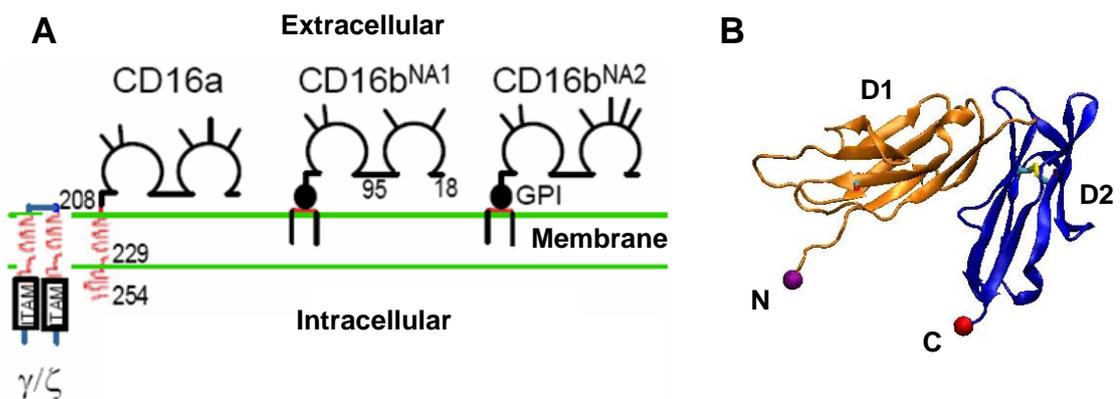
Figure 2.2 CD16 isotypes and structures. A. CD16 isotypes. Two Ig-like domains of each CD16 molecule are depicted with N-glycosylation sites shown as sticks. Figure adapted from Chesla et al. [3] **B.** The CD16 crystal structure (pdb 1e4j). N- and C-termini are indicated. Disulfide bonds connecting two β sheets are shown as sticks.

CD16 includes two subtypes: CD16a and CD16b, which differ in their anchors on cell membranes (Fig. 2.2A) [3]. CD16a contains a transmembrane anchor, but CD16b has a glycosylphosphatidylinositol (GPI) anchor. Each CD16 molecule has two Ig-like domains (D1 and D2) with β sandwich folding (Fig. 2.2B), in which only six amino acids are different between CD16a and CD16b. Two major alleles of CD16b exist: NA1 and NA2, which differ by only four amino acids. CD16 is highly glycosylated with a molecular weight of 50-80 KDa. CD16a, CD16b^{NA1}, and CD16b^{NA2} have 5, 4, and 6 N-glycosylation sites, respectively. The glycosylation patterns are also cell type specific. NK cell CD16a but not monocyte CD16a was found to have high mannose-type oligosaccharides [4].

Crystal Structures of CD16-IgG Fc Complex
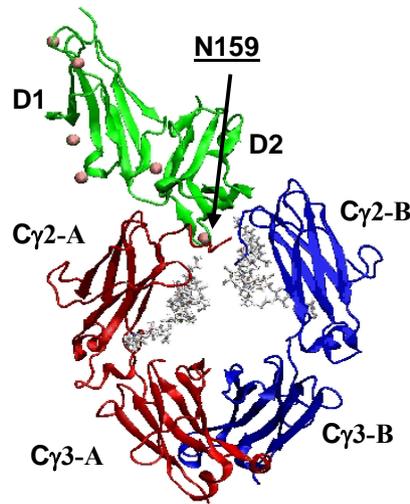
Figure 2.3 The crystal structure of a CD16b^{NA2}-IgG Fc complex (pdb 1e4k). CD16b^{NA2} in green color, chain A of IgG Fc in red, and chain B in blue. Gray sticks represent saccharides on Fc. Pink spheres indicate N-glycosylation sites.

The crystal structure of a CD16b^{NA2}-IgG Fc complex shows that CD16 and IgG forms a 1:1 stoichiometric complex [5, 6]. The D2 domain and the D1/D2 hinge of CD16 contact with the Cγ2 domains of IgG Fc. The Fc portion of IgG has an asymmetric opening upon binding, where Cγ2-B displaces more away from the symmetric axis of Fc than Cγ2-A. Two saccharide chains on Fc interact with each other and also have a few contacts with CD16. However, no saccharides on CD16 because the CD16 used for crystallization was produced in *Escherichia coli*. One of the six N-glycosylation sites of CD16b^{NA2}, ASN159 (numbering according to Sondermann et al. [5]), locates right in the binding pocket, which most likely affect the CD16-IgG binding if glycosylated.

Glycosylation Effects on CD16-IgG Binding

Different glycosylation patterns among CD16 isotypes or cell types result in varied affinities of CD16 to IgG. NK cell CD16a, which contains high mannose-type
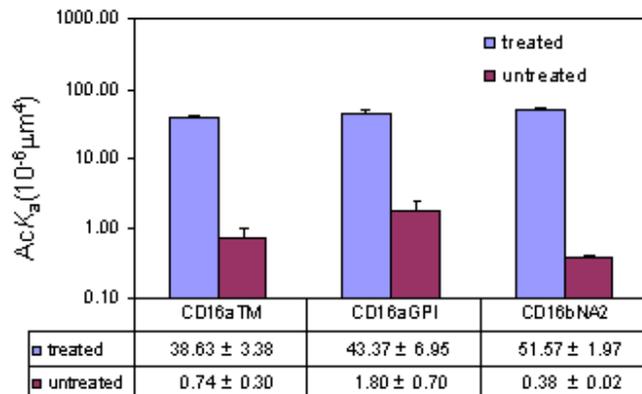
Figure 2.4 2D affinities of CD16 to IgG measured by micropipette assay. CHO CD16 was treated or untreated with N-glycosylation inhibitor, Tunicamycin. Figure provided by Ning Jiang.

oligosaccharide, has higher affinity to IgG than monocytes CD16a [4]. Ning Jiang's

experiments show that N-glycosylation of CD16 decreases its affinity to IgG by >10

folds (Fig. 2.4, unpublished). Particularly, the N-glycosylation at residue ASN159

contributed mostly to the decrease of CD16 affinity to IgG [7].

However, the structural basis of glycosylation effects on CD16-IgG binding is

still not quite clear. The glycosylation at ASN159 could introduce a steric hinder to the

CD16-IgG binding. But how it changes the binding in more details is not understandable

with the structure of the unglycosylated CD16-IgG Fc complex.

2.2 VWF A Domains

Functions of VWF in Hemostasis

von Willebrand factor (VWF) is a multidomain multimeric plasma glycoprotein

that is synthesized and secreted by vascular endothelial cells and megakaryocytes [8, 9].

A major function of VWF multimers is to mediate the adhesion of platelets to

subendothelial extracellular matrices at sites of vascular injury. When a blood vessel is

injuried, subendothelial matrix containing collagen is exposed. VWF, which circulates in
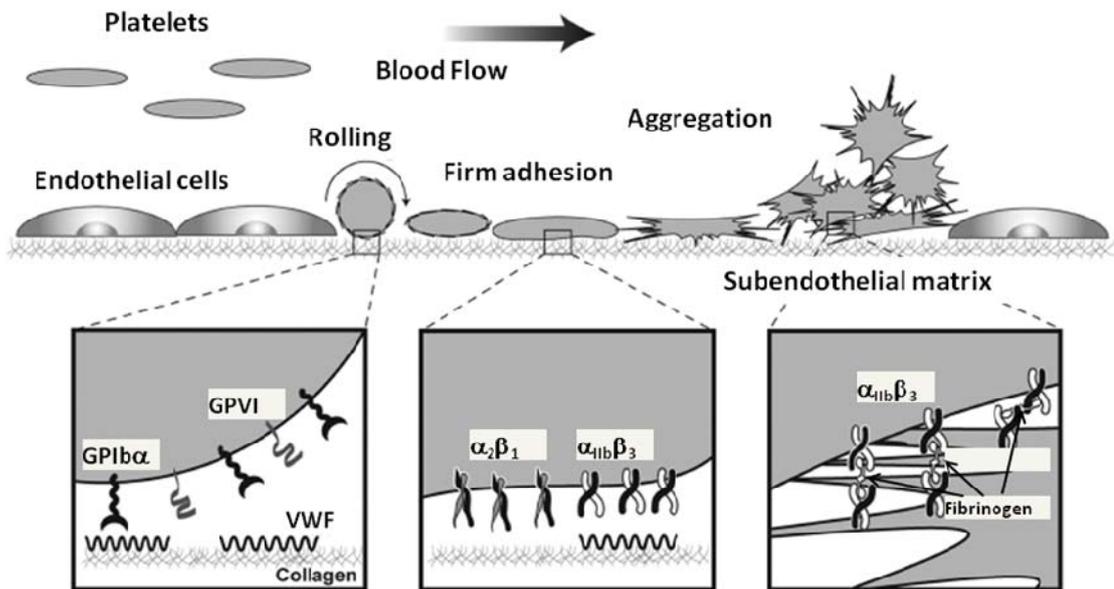
Figure 2.5 Sequence of events in platelet-subendothelial interactions at a site of vascular injury. Platelets first tether to and roll on the subendothelium through GPIb-VWF interaction, followed by integrin-mediated firm adhesion and aggregation. Modified figure from Chen et al. [10]

blood, binds to exposed collagen and is immobilized. Platelets then bind to immobilized VWF through GPIb receptors. GPIb-VWF interactions trigger intracellular signaling in platelets, which activate integrins $\alpha2\beta1$ and $\alpha IIb\beta3$. Active $\alpha2\beta1$ and $\alpha IIb\beta3$ interact with collagen and VWF, respectively, resulting in firm adhesion. Furthermore, platelets aggregate by $\alpha IIb\beta3$-fibrinogen interactions. Eventually, thrombus is formed to stop bleeding. (Fig. 2.5) [10]

Structures of VWF and Its A Domains

A VWF monomer consists of twelve domains: D1, D2, D', D3, A1, A2, A3, D4, B, C1, C2, and CK from N-terminus to C-terminus (Fig. 2.6) [9]. Two monomers form a dimer through disulfide bonds between two CK domains. Then dimers are polymerized, by disulfide bonds connecting D3 domains, into large multimers with molecular weights of up to 20,000 kDa. The D' and D3 domains bind to clotting factor VIII [11]. The C1 domain contains a RGD sequence, which binds to $\alpha IIb\beta3$ integrin [12]. The A1 domain
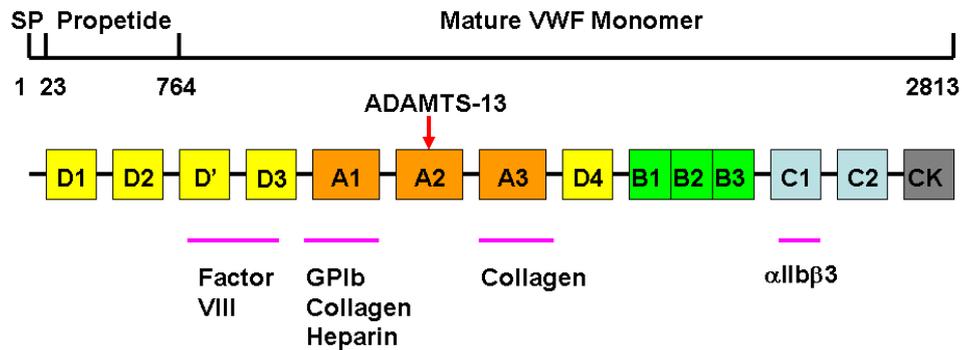
Figure 2.6 Diagram of the VWF primary structure. Signal peptide (SP): residue 1 to 22, propetide: residue 23 to 763, and mature peptide: residue 764 to 2813.

binds to platelet glycoprotein Ib (GPIb) [13, 14], collagen type VI [15], and heparin[16].

The A2 domain hosts a proteolytic site for a plasma protease [17], <u>A</u> <u>D</u>isintegrin <u>A</u>nd

<u>M</u>etalloprotease with a <u>T</u>hrombo<u>S</u>pondin type 1 motifs 13 (ADAMTS-13) [18], and the

A3 domain contains a binding site for collagen type I or type III [19].

Each of the three tandem A domains in VWF (Fig. 2.7A) adopts a Rossmann fold

with a central β-sheet flanked by α-helices as shown by the crystal structures [20-22]

(Fig. 2.7 B and D). The central β-sheet consists of six β-strands (β1-β6) while the

number of α-helices varies. The A1 domain contains six α-helices (α1, α3-α7) (Fig.

2.7B). Compared to homologous integrin α_M or α_L αA domain [23, 24], the A1 domain

lacks both the α2 helix and the metal ion-dependent adhesion site (MIDAS). Similarly,

the A3 domain does not have the α2 helix and the MIDAS but has an α8 helix right after

the α7 helix (Fig. 2.7D). A homology model [25] of the A2 domain shows that the α5

helix becomes a loop and the proteolytic site on the β4 strand is completely buried (Fig.

2.7C). The A1 and A3 domains contain disulfide bonds linking their N- and C-termini,

which the A2 domain lacks.

8
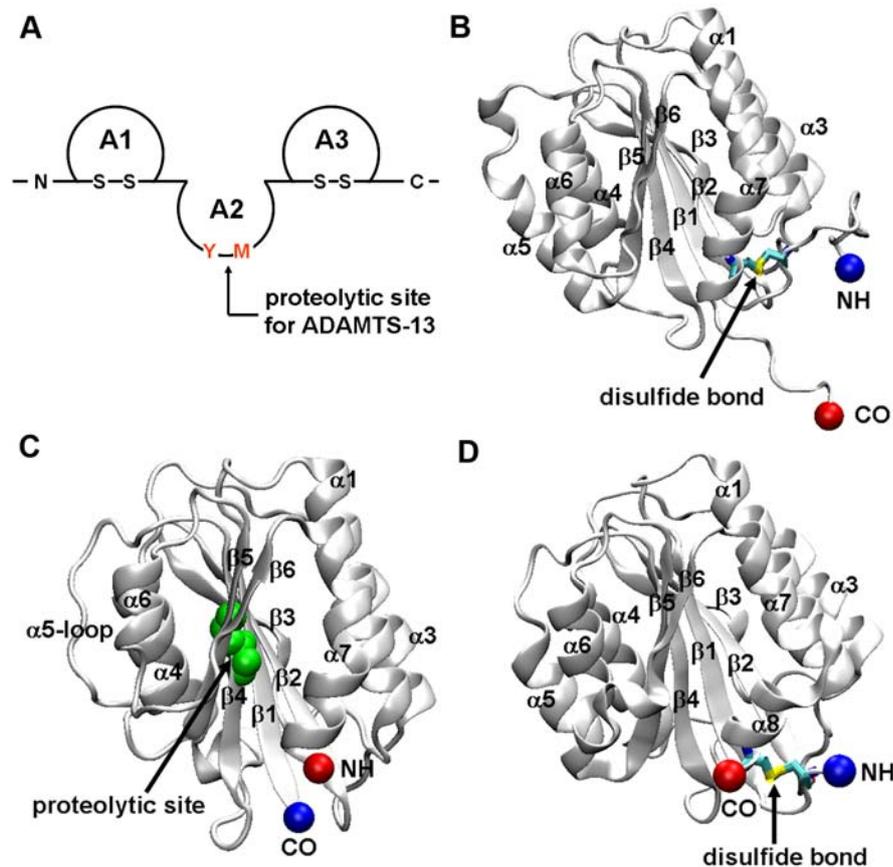
Figure 2.7 Structures of the VWF A domains. A. Schematic of VWF tandem A domains. S-S represents a disulfide bond. The ADAMTS-13 proteolytic site, which locates at the peptide bond between Tyr1605 and Met1606, is indicated. **B.** Crystal structure of the A1 domain (pdb 1auq) [22]. α-helices are shown as coiled ribbons, β-strands as ribbons with arrows, and loops as tubes. Two spheres represent, respectively, the N- and C-terminal Cα atoms. The disulfide bond is indicated. The A1 domain consists of 6 α-helices (α1, α3-α7) and 6 β-strands (β1-β6). **C.** Homology model structure of the A2 domain [25]. The backbone atoms of Tyr1605 and Met1606 adjacent to the proteolytic site are shown as spheres in the middle. The A2 domain includes 5 α-helices (α1, α3, α4, α6, α7) and 6 β-strands (β1-β6). α5 is a loop. **D.** Crystal structure of the A3 domain (pdb 1atz) [21]. The A1 domain has 7 α-helices (α1, α3-α8) and 6 β-strands (β1-β6).

VWF Cleavage by ADAMTS-13 and Unfolding of A domains

Upon stimulation, ultralarge VWF (ULVWF) multimers, which are stored in the

Weibel-Palade bodies in endothelial cells or the α-granules in megakaryocytes, are

secreted into blood [9, 26]. These ULVWF multimers bind GPIb more efficiently than

plasma VWF [27]. ADAMTS-13 rapidly cleaves ULVWF on the endothelial surface [26,

28] at the peptide bond between amino acid residues Tyr1605 and Met1606 in the A2

9
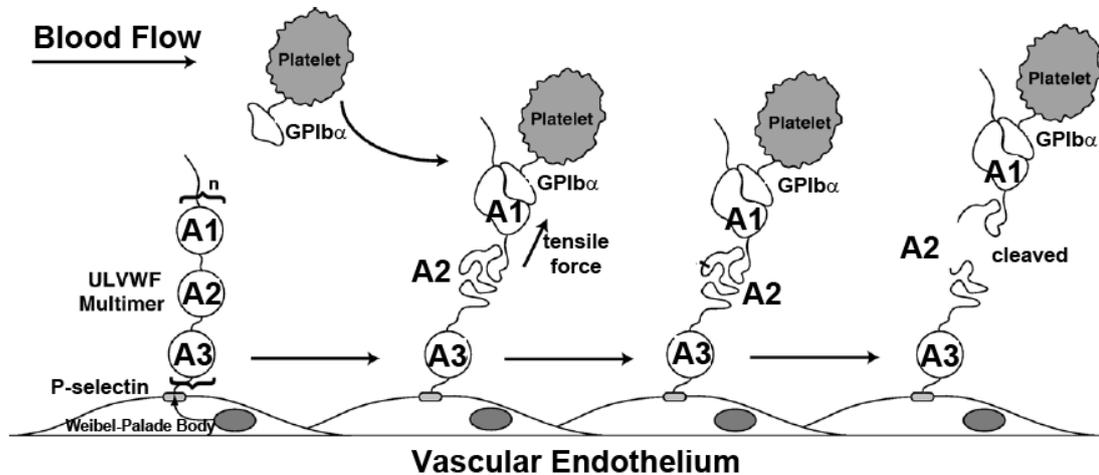
Figure 2.8 Model of VWF cleavage by ADAMTS-13. In blood flow, platelets bind to the A1 domain, generating tensile force. The force unfolds the A2 domain and exposes the proteolytic site for ADAMTS-13 cleavage. N: native state, I: intermediate state, and D: denatured state. Modified figure from Auton et al. [29]

domain [17], which disassembles ULVWF multimers and creats the full spectrum of circulating plasma VWF species, ranging from a single dimer to about 20 dimers in a multimer [9, 26]. Dysfunction of ADAMTS-13 results in systemic microvascular thrombosis in thrombotic thrombocytopenia purpura (TTP) [18, 30]. Mutations in the A2 domain, such as those in patients with type 2A von Willebrand disease (VWD), result in excessive proteolysis of VWF and thus the absence of high-molecular-weight VWF multimers, causing bleeding [31, 32].

ULVWF multimers attached to the endothelial surface are under the shear stress of flowing blood and the tensile force of attached platelets, which assist the rapid proteolysis of ULVWF by ADAMTS-13 [26, 29]. Tsai et al. showed that the proteolysis of plasma VWF by ADAMTS-13, while does not occur under static conditions *in vitro*, was enhanced under shear flow [33]. Yago et al. found that ADAMTS-13 proteolysis of an A1A2A3 tridomain linking a microsphere to a platelet in the shear flow was increased by the shear rate [34]. Also, the addition of denaturants, such as urea and guanidinium

hydrochloride, increased the ADAMTS-13 proteolysis of plasma VWF [35, 36]. Why do shear flow and denaturants facilitate the proteolysis of VWF by ADAMTS-13? Auton et al. [29] proposed a model that tensile force might unfold the A2 domain and expose the proteolytic site, thus facilitate the proteolysis of VWF by ADAMTS-13 (Fig. 2.7). Therefore, it is of interest to study the unfolding mechanism of VWF A domains to understand the regulation of the proteolysis of VWF by ADAMTS-13. Experimentally, it is very difficult to study the dynamic processes of the unfolding of the VWF A domains in atomic details. In chapter 5, molecular dynamics (MD) was used to simulate the unfolding of the VWF A1 and A2 domains by tensile force or under high temperature.

2.3 Integrins

Integrin Functions and Structures

Integrins are adhesion molecules that mediate cell-cell, cell-extracellular matrix (ECM), and cell-pathogen interactions [37]. They can regulate force-resistant adhesion, polarization in response to extracellular cues, and cell migration by integrating the cytoskeleton of cells with points of attachment in extracellular environments. Integrins are of vital importance to humans and many other organisms because they are related to important physiological processes such as tissue morphogenesis, inflammation, wound healing, and the regulation of cell growth and differentiation.

An integrin molecule is a hetero-dimer consisting of two noncovalently associated, type I transmembrane glycoproteins: α- and β-subunits. To date, 18 α- and 8 β-subunits have been identified in mammalia, forming 24 known integrins (Fig. 2.9A)
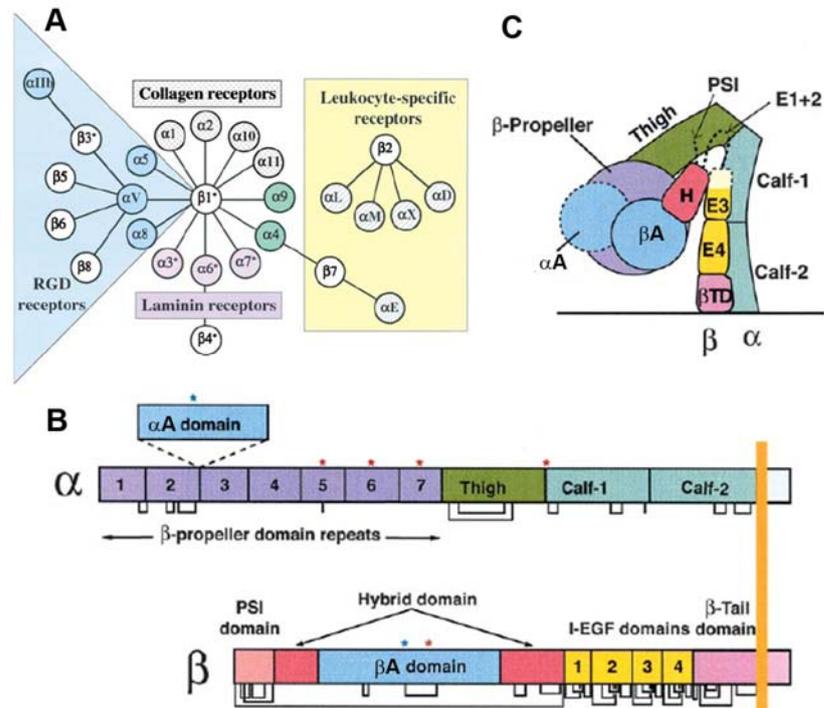
Figure 2.9 Integrin family and structure. A. Integrin family. In mammalian, 8 β and 18 α subunits form 24 distinct integrins. They can be considered in several subfamilies based on evolutionary relationships (coloring of α subunits), ligand specificity, and restricted expression on leukocytes. α subunits with gray hatching or stippling have inserted αA domains. Figure adapted from Hynes [38]. **B.** Domain organization in the primary structure of integrin. Asterisks show Mg^{2+} (blue) and Ca^{2+} (red) binding sites. Lines below the stick diagrams indicate disulfide bonds. **C.** Arrangement of domains within three-dimensional crystal structure of $\alpha_V\beta_3$ with an αA domain added. Figures B and C adapted from Takagi et al. [39]

[38]. The α- and β-subunits contain respective extracellular domains of 940 and 640

residues, single-span transmembrane domains, and short cytoplasmic domains of 20-50

amino acids, except that β_4 has a long cytoplasmic tail. The extracelluar domains form a

globular headpiece that binds ligands and two long stalk legs that connect the headpiece

to the transmembrane and the C-terminal cytoplasmic domains (Fig. 2.9C) [37, 39]. The

α-subunit extracellular portion consists of a seven-bladed β-propeller domain (60 amino

acids in each blade), a Thigh domain, and two Calf (1 and 2) domains. Nine α-subunits

include an extra αA domain inserted between blades 2 and 3 of the β-propeller domain,

12

which adopts the Rossmann fold homologous to the VWF A domains. The β-subunit extracellular portion is composed of a PSI domain (named for plexins, semaphorins and integrins), a βA domain homologous to the αA domain, a Hybrid domain with a β-sandwich fold, four cysteine-rich EGF domains, and a β-tail domain (βTD) (Fig. 2.9B) [37, 39].

Bi-directional Signaling of Integrins through Global Conformational Changes

Integrins can transmit signals bi-directionally across the plasma membrane [40, 41]. On one hand, ligand binding of integrins can be dynamically regulated through inside-out signaling, where stimuli received by cell surface receptors for chemokine, cytokines, and so on initiate intracellular signals that impinge on integrin cytoplasmic domains and alter binding affinity for extracellular ligands. On the other hand, integrins bound to soluble or immobilized physiologic ligands form clusters and transduce mechanochemical signals from the extracellular domain to the cytoplasm, which is called outside-in signaling.

During the bi-directional signaling, integrins undergo conformational changes that were revealed by a large amount of experimental data. The first crystal structure of the whole integrin ectodomain was solved for $\alpha_V\beta_3$ by Xiong et al. [42-44]. In the crystal, $\alpha_V\beta_3$ integrin adopts a bent conformation (Fig. 2.10A). The twelve domains of $\alpha_V\beta_3$ assemble into an ovoid "head" and two "legs". The headpiece bends over the tailpiece to such an extent that it is extremely close to the C-terminal, membrane-proximal end of the two legs. In the headpiece of $\alpha_V\beta_3$ crystal structure, the angle between the Hybrid domain and the βA domain is acute, which is called the closed conformation. Recent published $\alpha_V\beta_3$ integrin crystal structure also showed the same bent conformation with a closed

13
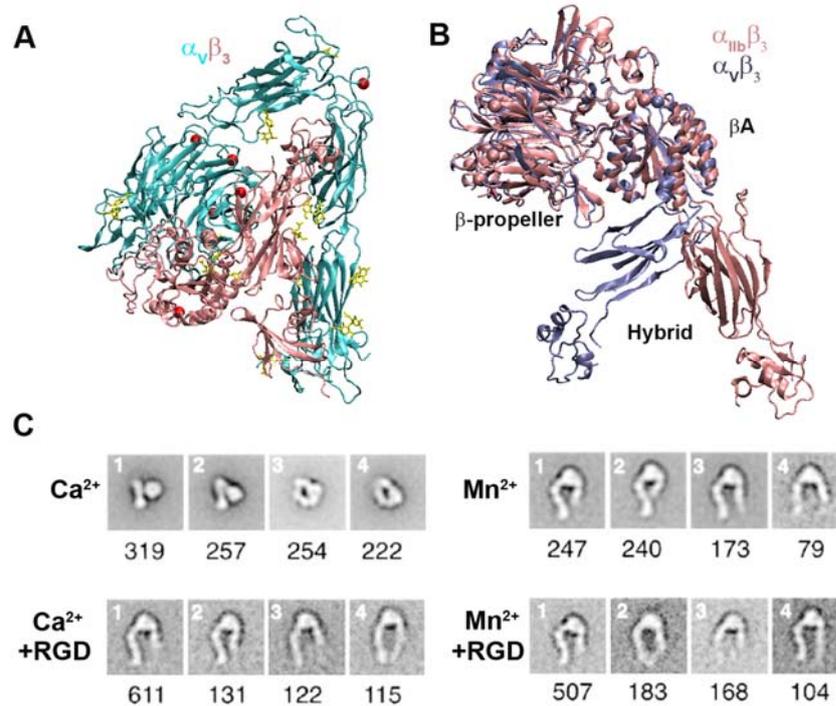
Figure 2.10 Integrin crystal structures and 2D EM images. A. The crystal structure of the ectodomain of unliganded $\alpha_V\beta_3$ integrin. α-subunit in cyan, β-subunit in blue, sugar in yellow, and Ca^{2+} ions in red. **B.** The crystal structure of the headpiece of $\alpha_{IIb}\beta_3$ integrin aligned with the headpiece of $\alpha_V\beta_3$ integrin. $\alpha_{IIb}\beta_3$ in pink and $\alpha_V\beta_3$ in blue. **C.** 2D EM images of $\alpha_V\beta_3$ integrin under conditions as labeled. Figure C adapted from Takagi et al. [45]

headpiece [46]. In contrast, in the crystal structure of a legless $\alpha_{IIb}\beta_3$ integrin with ligands or pseudo-ligands [47], Hybrid swings out for 60~90 degree, which is called the open conformation (Fig. 2.10B). Furthermore, multiple global conformations of integrins were observed in 2D negative staining electron microscopy (EM) images (Fig. 2.10C), including bent, extended with a closed headpiece, and extended with an open headpiece conformations [45, 48, 49]. In the presence of 1mM Mn^{2+}, the condition in which $\alpha_V\beta_3$ integrin was shown to have high affinity for its ligands, $\alpha_V\beta_3$ mainly adopted an extended conformation. But with 1mM Ca^{2+} and Mg^{2+}, where $\alpha_V\beta_3$ integrin was shown to have low affinity for its ligands, the majority of $\alpha_V\beta_3$ had a bent conformation. A head-to-leg disulfide bond locked $\alpha_V\beta_3$ integrin in a low-affinity state [45]. Besides, the
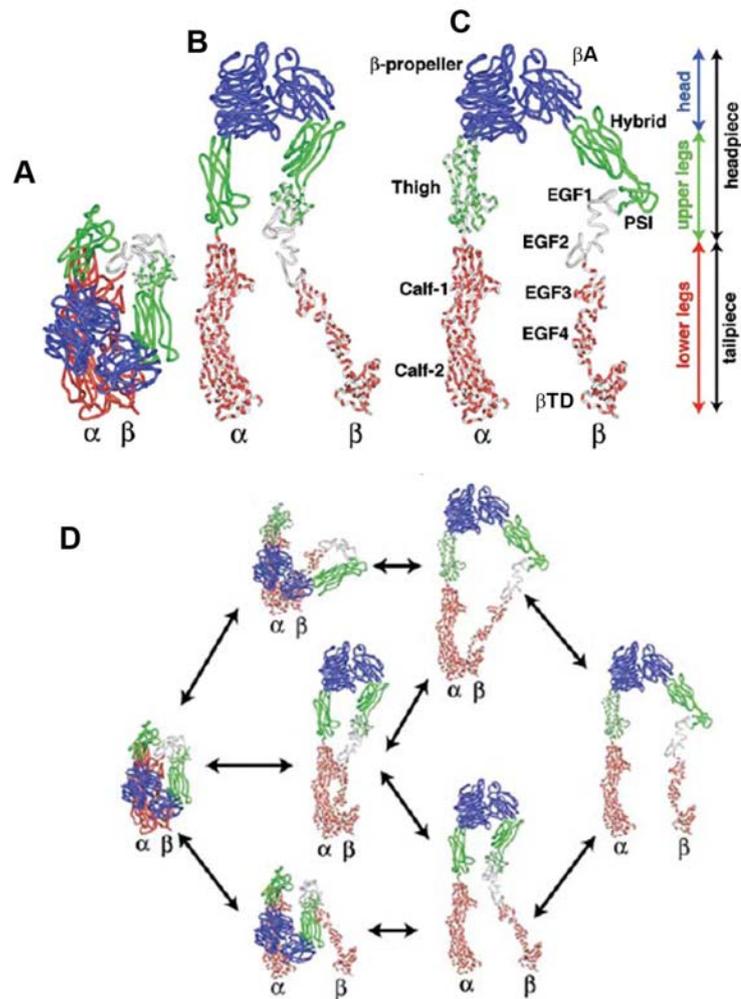
Figure 2.11 Switchblade model of integrin signaling. Three conformations: bent (**A**), extended with a closed headpiece (**B**), and extended with a open headpiece (**C**), correspond to low-affinity, intermediate-affinity, and high-affinity states, respectively. **D.** Pathways of global conformational changes during inside-out (lower pathway) or outside-in (upper pathway). Adapted from Xiao et al. [47]

epitope of the activation-sensitive antibody KIM127 in EGF2/3 of β2 subunit is masked

in the bent conformation [50]. All these data suggest that the bent conformation is the

inactive, low-affinity state while the extended conformation is the active, high-affinity

state. Therefore, transition between the inactive and active states requires the global

conformational change between the bent and extended conformations.

How does the transition between different conformations of integrin occur? A switchblade model was suggested [45, 47, 50] (Fig. 2.11). During inside-out signaling, binding of an intracellular protein (e.g., talin) to the integrin cytoplasmic tail causes the separation of α and β cytoplasmic/transmembrane domains, which induces the separation of the two legs of the integrin. As a result, the integrin stands up with an intermediate affinity and a close headpiece. Furthermore, the Hybrid domain swings out at the junction with the βA domain, which opens the headpiece and brings βA into a high affinity state for ligand binding. For outside-in signaling, ligand binding to a low affinity, bent integrin with a closed headpiece drives the outward Hybrid swing, leg separation, extension, and further separation of transmembrane and cytoplasmic segments, which exposes binding site for intracellular signaling molecules to assemble a cytosekeletal-based signaling complex.

Alternatively, a deadbolt model [51] was proposed for inside-out signaling based on the fact that a bent $\alpha_V\beta_3$ integrin binds to a soluble physiologic ligand fibronection, containing type III domains 7-10 and the EDB domain [52] and other evidences. The CD loop of the βTD domain locks the βA domain in the low-affinity state. The change of the tilting angle of the transmembrane segaments by inside-out signals leads to the disengagement of the deadbolt, which switches βA into the high-affinity state. Hybrid swing-out and integrin extension are thought to be an outside-in signaling after ligand engagement.

The role of integrin global conformational changes is still debatable, but it most likely occurs at a certain stage of integrin signaling. So studying the dynamic processes of integrin global conformational changes is very important for understanding the

16

structure-function relations of integrins. Integrin unbending from a bent to an extended conformation is a 10~20 nm large-scale conformational change. How does the unbending occur? What are the steps during the unbending process? What are important residues for the unbending? To answer these questions, steered molecular dynamics (SMD) was used to simulate in atomic details the unbending of $\alpha_V\beta_3$ integrin in response to a tensile force.

CHAPTER 3

METHODS

MD Simulations

      All topology and parameter files were generated using LEaP of the AMBER 8 suite [53] with the Cornell et al. force field [54] or Duan et al. [55] force field for proteins and Woods et al. force field [56] for carbohydrates. MD simulations were performed with the NAMD package [57]. The crystal structures for the CD16-hIgG1 Fc complex (pdb 1e4k), VWF A domains (pdb 1auq for A1 and 1atz for A3), $\alpha_V\beta_3$ integrin (pdb 1u8c) were taken from Protein Data Bank [58].

      The protein structures were placed in either spheres or rectangular periodic boxes of TIP3P water molecules. With water boxes, a 10-15 Å minimal distance between boundaries and proteins was guaranteed. Then counterions (Na^+ and Cl^-) were added to neutralize the systems, and produce 100-150 mM ionic concentration in the case of water boxes. The systems were initially energy-minimized with a conjugate gradient method for three times of 10,000 steps each: firstly all atoms of the proteins fixed, secondly only the backbone atoms fixed, and thirdly all atoms free. The particle mesh Ewald (PME) method [59] was used to treat long-range electrostatic interactions and SHAKE [60] was used to constrain bond lengths involving bonds to hydrogen atoms. The time-step for MD simulations was 2 fs with a 9 or 12 Å non-bonded cutoff. Then the systems were gradually heated up from 0 K to 300K or 310 K during 60-100 ps canonical ensemble (NVT)-MD simulations with harmonic restraints (force constants of 1 or 5 kcal mol^{-1} Å^{-2}) to all solute atoms. Constant temperature was controlled by Langevin dynamics with a damping coefficient of 5 ps^{-1}. For water boxes, subsequent isothermal isobaric ensemble

(NPT)-MD simulations were performed for 100-120 ps to adjust solvent density.

Constant pressure was controlled by Langevin piston method. After that, the harmonic

restraints on solute atoms were gradually reduced to zero during 100-120 ps NVT-MD

simulations. The equilibrated systems were used for further production runs.

The VMD program [61] was used for visualization and analysis, including

drawing of protein structures, hydrogen bond measurements (defined by a <3.5 Å

distance between donor and acceptor atoms and a >120° angle formed by donor,

hydrogen, and acceptor atoms), and calculations of root mean square deviation (RMSD),

root mean square fluctuation (RMSF), and solvent accessible surface area (SASA) with

1.4 Å probe radius. Secondary structures were determined using STRIDE [62]

incorporated in VWD.

SMD Simulations

SMD simulations were performed using NAMD [57]. A group of atoms were

chosen as SMD [63] atoms. For constant-force SMD, a given constant force was loaded

on the SMD atoms. For constant-velocity SMD, the center of mass (COM) of the SMD

atoms was linked to a spring with a given force constant, which moved at a given

constant velocity. The applied potential U is as following:

$$U(\vec{r}_1,\vec{r}_2,...,t) = \frac{1}{2}k[vt - (\vec{R}(t) - \vec{R}_0) \cdot \vec{n}]^2 \ ,$$

where k is the force constant, v is the moving speed, $\vec{R}(t)$ is the current position of COM,

$\vec{R}_0$ is the initial position of COM, and $\vec{n}$ is the force direction.

Targeted MD Simulations

Targeted MD simulations were run with AMBER8 [53]. A simulated structure was forced to move toward a reference structure. An additional term is added to the potential function based on the mass-weighted RMSD of a set of atoms in the current structure compared to a reference structure as the follows:

$$E = \frac{1}{2}k \cdot N_{atom} \cdot (RMSD - TGTRMSD)^2 ,$$

Where E is potential energy, k is a force constant, N_{atom} is the number of atoms selected to calculate current RMSD, and TGTRMSD is the target RMSD. During a simulation, TGTRMSD was decreased linearly from the initial value, equal to the RMSD of the initial simulated structure from the reference structure, to zero.

Homology Modeling

A modeled sequence was aligned with a homologous sequence, whose crystal structure is available, using the web-based tool "needle" [64] provide on the website of European Bioinformatics Institute (EMBL-EBI). The alignment was then used as an input of MODELLER [65]. Based on the crystal structure of the template, MODELLER generated a 3D model structure of the model sequence. The qualities of model structures were evaluated by (Discrete Optimized Protein Energy) DOPE scores [66]. Finally, energy minimization was run to eliminate bad interactions.

Binding Free Energy Calculation

The binding free energy of each complex was calculated using the Molecular Mechanics-Poisson Boltzmann/Generalized Born Surface Area (MM-PB/GBSA) approach [67, 68], which has been incorporated in AMBER 8 package. Isolated trajectory of each protein in the complex was obtained by extracting the corresponding portion from

the complex trajectory. For each MD trajectory, 150 snapshots were taken from the last 3 ns—one snapshot every 20 ps. All water molecules and counterions were stripped. Then the binding free energy was calculated according to:

$$\Delta G_{total} = \left\langle G^{complex}(i) - G^{protein1}(i) - G^{protein2}(i) \right\rangle_i ,$$

where $\langle \cdot \rangle$ indicates an average over all the snapshots. $G^x(i)$ is a sum of contributions of gas-phase energy, solvation free energy, and solute entropy as follows:

$$G^x(i) = H^x_{gas}(i) + G^x_{solvation}(i) - TS^x(i).$$

Gas-phase energy H^x_{gas} includes bond, angle, torsional, electrostatic, and van der Waals energies. Solvation free energy $G^x_{solvation}$ consists of polar and non-polar contributions. The polar solvation free energy was calculated using the Delphi [69] to solve Poission-Boltzmann (PB) equations or by the aoGB model [70]. Dielectric constants for the interior and the exterior of the solute were 1 and 80, respectively. The non-polar solvation free energy was estimated by solvent-accessible surface area (SASA):

$$G^x_{non-polar}(i) = \gamma \cdot SASA^x(i) + b ,$$

where $SASA^x(i)$ was calculated using the LCPO method [71] in AMBER 8 and γ, b were set to 0.00542 kcal mol^{-1} Å^{-2} and 0.92 kcal mol^{-1}, respectively, if using the Delphi; or γ, b were set to 0.005 kcal mol^{-1} Å^{-2} and 0 kcal mol^{-1}, respectively, if using the aoGB model. The solute entropy S^x was calculated with classical statistical thermodynamics. A normal mode analysis was used to obtain vibrational frequencies.

In addition, the effective energies, which do not include entropy contributions, were decomposed to the contributions of residues according to the decomposition method [68] implemented in the MM-PBSA program of AMBER 8 when using the aoGB model.

CHAPTER 4

STRUCTURAL BASIS FOR GLYCOSYLATION EFFECTS ON

FCγRIIIB-IGG BINDING

As mentioned in section 2.1, human CD16 is a highly glycosylated protein [3].

Experiments showed that the glycosylation of CD16 significantly decreased its binding

affinity to IgG (Fig. 2.4) [7]. Here, computational methods were used to reveal the

structural basis of glycosylation effects on CD16-IgG binding.

4.1 Model of a Glycosylated CD16-IgG Fc Complex
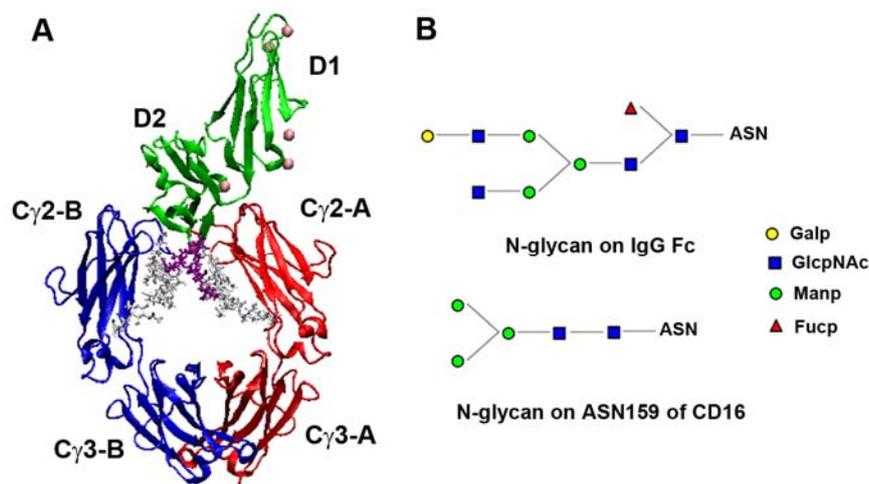
Figure 4.1 Model of a glycosylated CD16-IgG Fc complex. A. The structure of the model, which is rotated ~180 degree around a vertical axis if compared to Fig. 2.3. CD16b[NA2] in green color, chain A of IgG Fc in red, and chain B in blue. Pink spheres are N-glycosylation sites. Gray sticks represent saccharides on IgG Fc and purple sticks indicate an N-glycan core attached to ASN159 using the web-based "GlyProt" tool. **B.** Linear structures of oligosaccharide chains on IgG Fc and CD16. Galp, Galactopyranose; GlcpNAc, N-Acetyl-Glucosamine pyranose; Manp, Mannopyranose; Fucp, Fucopyranose.

Only the crystal structure of an unglycosylated CD16b[NA2]-IgG Fc complex

structure is available until now (Fig. 2.3) [5, 6]. To obtain a structure of a glycosylated

CD16-IgG Fc, I attached *in silico* an N-glycan core to residue ASN159 of CD16b[NA2] in the unglycosylated CD16b[NA2]-IgG Fc complex (Fig. 4.1), using the web-based "GlyProt" tool (http://www.glycosciences.de/glyprot) developed by Bohne-Lang et al. [72] The N-glycan oligosaccharide chain was built based on most populated N-glycan conformations in Protein Date Bank (PDB) and to avoid space conflicts with existing atoms.

4.2 Free MD Simulations of the CD16-IgG Fc Complexes

Equilibration of the Unglycosylated CD16-IgG Fc Complex
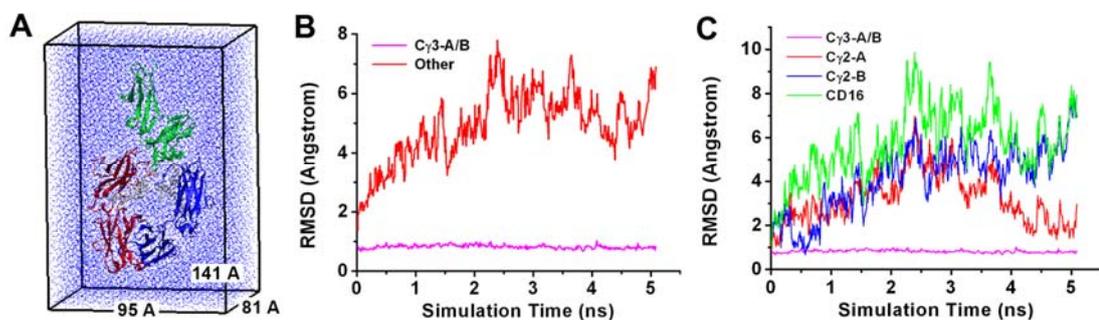
Figure 4.2 Equilibration of the unglycosylated CD16-IgG Fc complex. A. The complex was soaked in a water box. Water molecules are shown as blue dots. **B.** RMSD of Cα atoms of Cγ3-A and B domain β-sheets (purple), which were used as the alignment reference, and other Cα atoms than the reference (red). **C.** RMSD of Cα atoms of Cγ3-A and B domain β-sheets (purple), Cγ2-A domain β-sheets (red), Cγ2-B domain β-sheets (blue), and CD16 β-sheets (green).

The unglycosylated CD16-IgG Fc complex was soaked in a water box (Fig. 4.2 A), resulting in a system of ~100,000 atoms. After energy minimization, the system was heated up to 310 K. Finally, 5 ns free MD simulation was performed under constant pressure and constant temperature. The RMSD of Cα atoms, using Cγ3-A and B domain β-sheets as the alignment reference, gradually increases in the first 2 ns and then reached a plateau, indicating equilibrium (Fig. 4.2B). However, the RMSD showed quite large fluctuations. The RMSD of individual domains indicated that the Cγ2-A and B domains fluctuated a lot relative to the Cγ3-A and B domains (Fig. 4.2C). First, both Cγ2-A and B

domains moved away from their initial positions. Then Cγ2-A domain returned back while Cγ2-B domain went away further. These results demonstrated the flexibility of the connecting hinges between Cγ2 and Cγ3 domains. CD16 moved with its binding partner, the Cγ2 domains, therefore displayed large fluctuations.
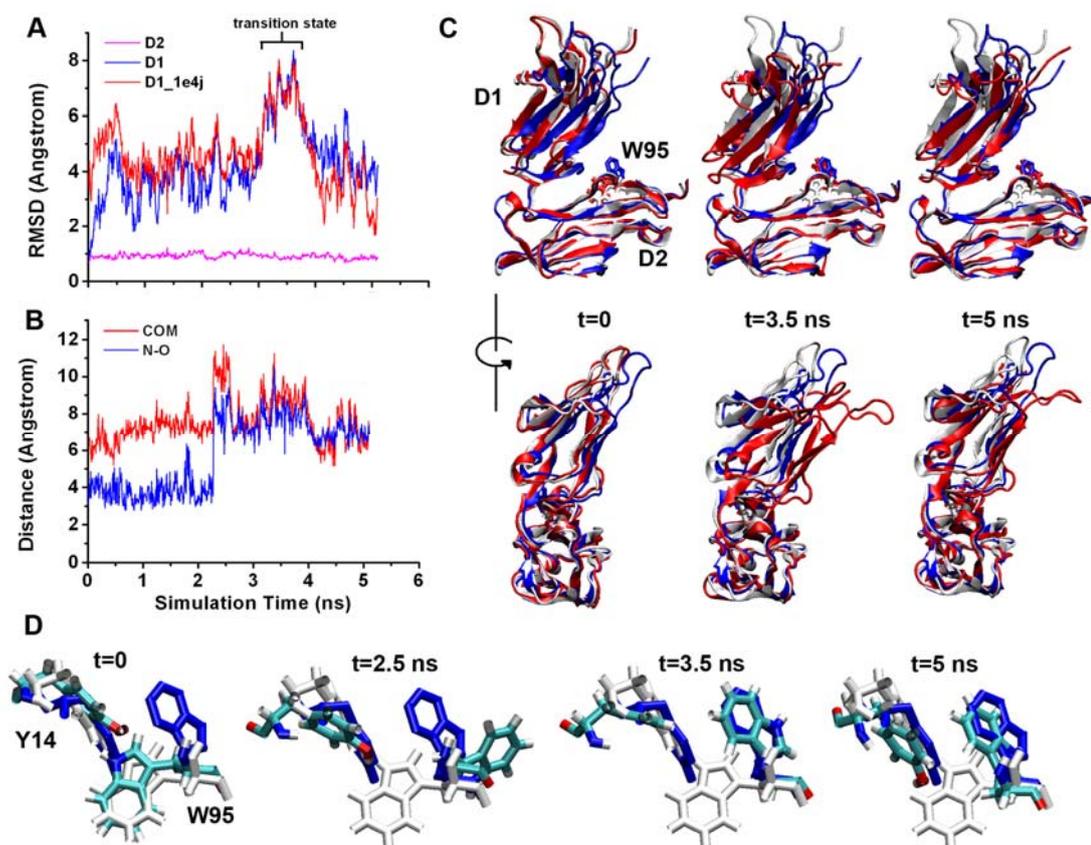
Figure 4.3 Transition of CD16 conformations. A. RMSD from the CD16-IgG Fc complex crystal structure: Cα atoms of D2 domain (purple), using as alignment reference, and D1 domain (blue). RMSD from the standalone CD16 crystal structure: Cα atoms of D1 domain (red). **B.** Distance between the center of mass (COM) of W95 and the COM of Y14 (red), and between $N_{\varepsilon 1}$ of W95 and O_{η} of Y14 (blue). **C.** Snapshots of CD16 structure at indicated time. White is crystal structure of CD16-IgG Fc complex, blue is crystal structure of standalone CD16, and red is simulated structure. Lower row is rotated along vertical axis by 90 degree relative to upper row. Residue W95 are shown as sticks. **D.** Snapshots of W95 and Y14 at indicated time. White sticks are from CD16-IgG Fc complex, blue sticks are from standalone CD16, and multiple color sticks are simulated (nitrogen in blue, oxygen atoms in red, carbon in cyan, and hydrogen in white).

Compared to its standalone crystal structure, CD16 in the complex crystal has a larger opening angle, where residue Trp95 in the D2 domain rotates ~90 degree along the Cα-Cβ bond toward the D2 domain (Fig. 4.3C). In the simulation, inter-domain motions of CD16 were observed when CD16 structures were aligned with the D2 domain (Fig. 4.3A). Interestingly, CD16 transitioned from the conformation in its complex (open conformation) to the standalone conformation (close conformation). Initially, CD16 was in the open conformation, where the RMSD of D1 from the complex structure was larger than the RMSD from the standalone structure. Then the RMSD suddenly jumped up, corresponding to a twist motion of D1 relative to D2 (Fig. 4.3B). Finally, D1 rotated back but became closer to the standalone structure. Motions of residue Trp95 seems to be correlated to the domain motions of CD16. Trp95 first pointed inward and formed a hydrogen bond with Tyr14 in the D1 domain (Fig. 4.3D). Before the domain twisting, the hydrogen bond was broken and Trp95 suddenly flipped out (Fig. 4.3C). After CD16 got into the close conformation, Trp95 kept pointing outward. These results suggest that CD16 may have bistable conformations, and the transition between them is regulated by the interaction between Tyr14 and Trp95.
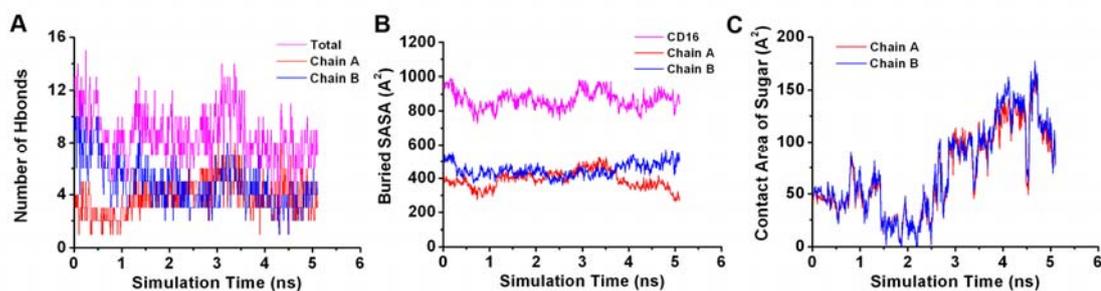
Figure 4.4 Interactions of the unglycosylated CD16-IgG Fc complex. A. Number of hydrogen bonds between CD16 and IgG Fc. A hydrogen bond is defined by a <3.5 Å distance between donor and acceptor atoms and a >120° angle formed by donor, hydrogen, and acceptor atoms. **B.** Buried SASA due to binding. **C.** Contact area of two carbohydrate chains on IgG Fc, calculated as the buried SASA due to contacts.

During the equilibration, 8-12 hydrogen bonds and 3 salt bridges existed between CD16 and IgG Fc, and almost equally distributed at chain A and B of IgG Fc (Fig. 4.4A and Table 4.1). Five residues (Leu235-Ser239) participating in hydrogen bonding locate at the lower hinge, bridging Fab to Fc, of chain A, which demonstrates the importance of this hinge to CD16 binding. The buried SASA due to binding is 800-1000 Å2. The buried SASA of chain B of Fc is slightly larger than that of chain A (Fig. 4.4B). Two carbohydrate chains on IgG Fc were interacting with each other. Their contacts were increased after 3 ns (Fig. 4.4C).

Table 4.1
Hydrogen bonds[1] between CD16 and IgG Fc for the unglycosylated CD16-IgG Fc complex

CD16	Chain A of IgG Fc	Existing Time (ns)	CD16	Chain B of IgG Fc	Existing Time (ns)
Trp110 $N_{\varepsilon1}$	Ala327 O	2.49	[2]Lys117 N_ζ	[2]Asp265 $O_{\delta1}$	1.26
			[2]Lys117 N_ζ	[2]Asp265 $O_{\delta2}$	4.11
[2]Lys111 N_ζ	[2]Glu269 $O_{\varepsilon1}$	0.6			
[2]Lys111 N_ζ	[2]Glu269 $O_{\varepsilon2}$	0.81	Asp126 $O_{\delta1}$	Ser267 O_γ	5.06
			Asp126 $O_{\delta1}$	Ser267 N	4.65
Ser157 O_γ	Leu235 O	3.11			
Ser157 N	Leu235 O	0.52	[2]Lys128 N_ζ	[2]Glu269 $O_{\varepsilon1}$	0.49
			[2]Lys128 N_ζ	[2]Glu269 $O_{\varepsilon2}$	0.89
Lys158 N_ζ	Pro238 O	3.82			
			His131 $N_{\varepsilon2}$	Ser298 O_γ	1.49
Lys158 N_ζ	Gly237 O	1.48			
			Arg152 $N_{\eta1}$	[3]NAG801 O_7	0.54
Lys158 N_ζ	Ser239 O_γ	1.7	Arg152 $N_{\eta2}$	[3]NAG801 O_7	2.24
Lys158 N_ζ	Gly236 O	3.21			

1. Only residue pairs, which formed hydrogen bonds for >1 ns, are listed.
2. Salt bridges.
3. NAG: N-Acetyl-Glucosamine pyranose.

Equilibration of the Glycosylated CD16-IgG Fc Complex

Similarly, the glycosylated CD16-IgG Fc complex structure was soaked in a water box. After energy minimization and heat-up to 310K, 5 ns free MD was run with constant temperature and pressure. When aligned with the Cγ3-A and B domain β-sheets, the

RMSD of Cα atoms showed two plateaus in the last 3 ns (Fig. 4.5A). According to the

RMSD of individual domains, the jump of RMSD was due to the motion of CD16 (Fig.

4.5B). Compared to the unglycosylated CD16-IgG Fc complex, the RMSD of CD16 was
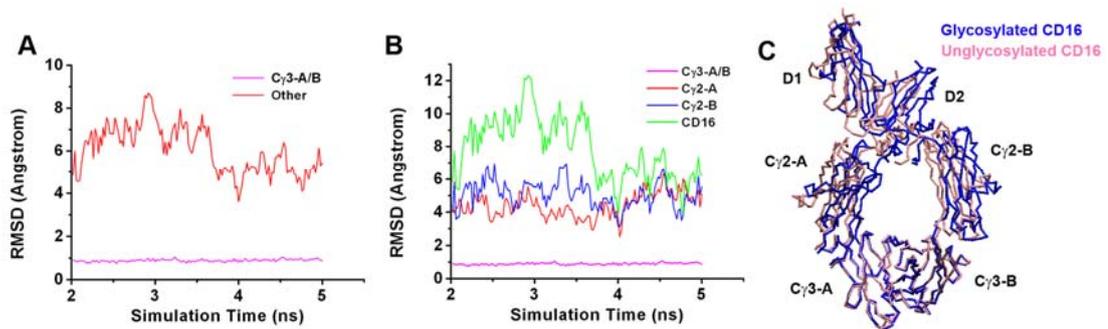
Figure 4.5 Equilibration of the glycosylated CD16-IgG Fc complex. A. RMSD of Cα atoms of Cγ3-A and B domain β-sheets (purple), which were used as the alignment reference, and other Cα atoms than the reference (red). **B.** RMSD of Cα atoms of Cγ3-A and B domain β-sheets (purple), Cγ2-A domain β-sheets (red), Cγ2-B domain β-sheets (blue), and CD16 β-sheets (green). **C.** Comparison of mean structures of unglycosylated and glycosylated CD16-IgG Fc complexes, aligned with Cγ3-A and B domain β-sheets.

much larger at the high level. The RMSD of the Cγ2-A and B domains kept at 4-6 Å

range. Interestingly, a significant conformational change was observed when the mean

structures of the two complexes over the last 3 ns simulations were compared. In the

glycosylated CD16-IgG Fc complex, the CD16 shifted more to the Cγ2-B domain of Fc,

the Cγ2-B domain opened more, and the Cγ2-A domain became more closed (Fig. 4.5C).

What caused the conformational change? In the previous simulation for the

unglycosylated CD16-IgG Fc complex, two carbohydrate chains on chains A and B of Fc

interacted with each other (Fig. 4.4C). However, for the glycosylated CD16-IgG Fc

complex, the two carbohydrate chains of Fc almost have no interactions. Instead, sugar

on chain B of Fc interacted with sugar on CD16 (Fig. 4.6A). Several hydrogen bonds

were formed between them in the last 3 ns (Fig. 4.6B). Three residues were identified to

be involved in hydrogen bonding between these two carbohydrate chains: fucose of sugar

on chain B of Fc, N-acetyl-glucosamine and mannose of sugar on CD16, although they did not interact with each other in the starting structure (Fig 4.6C).
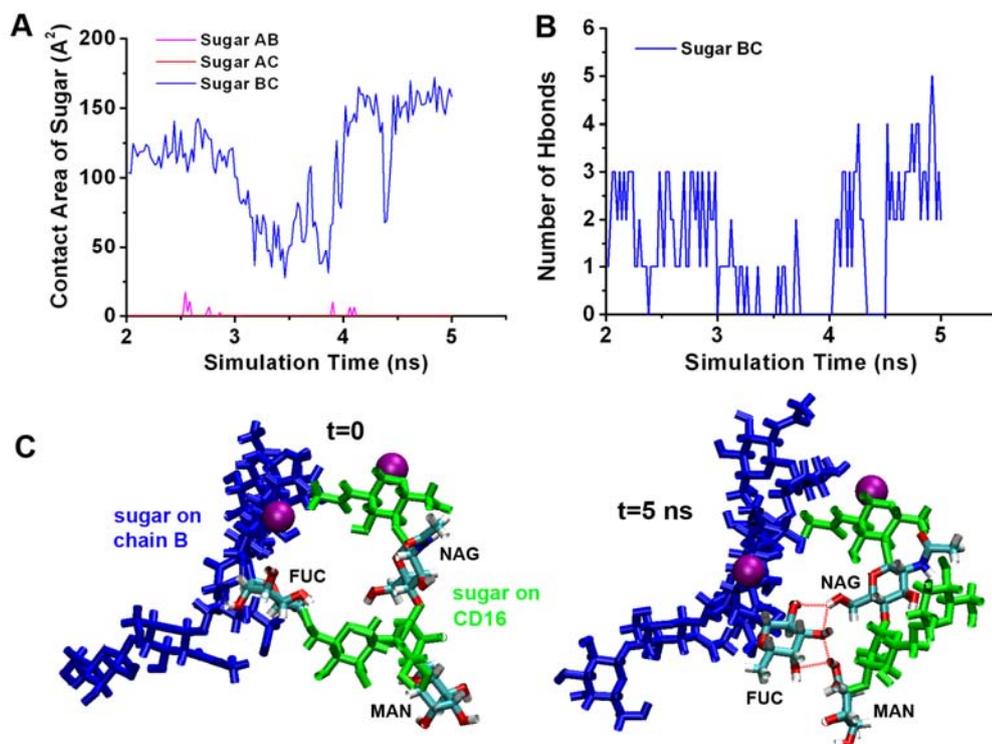
Figure 4.6 Interactions of carbohydrate chains on CD16 and IgG Fc. A. Contact area between sugar chains. Sugar AB: between sugar chains on chains A and B of IgG Fc, sugar AC: between sugar chains on chain A of Fc and CD16, sugar BC: between sugar chains on chain B of Fc and CD16. Contact area was calculated as the average of buried SASAs of two contacting partners. **B.** Number of hydrogen bonds between sugar chains on chain B of Fc and CD16. **C.** Snapshots of structures of sugar chains on chain B of Fc and CD16. Purple balls indicate $N_{\delta2}$ atoms of ASN residues that carbohydrate chains are attached at. Three interacting residues are highlighted in multicolor (cyan: carbon, red: oxygen, blue: nitrogen, white: hydrogen). FUC: fucose, NAG: N-acetyl-glucosamine, MAN: mannose.

How was the binding pocket changed due to the global conformational change? The number of hydrogen bonds on chain A of IgG Fc was largely decreased (Fig. 4.7A and Table 4.2). Also, the buried SASA of chain A was reduced about 100 $Å^2$ (Fig. 4.7B). Due to contacts between carbohydrate chains, the hydrogen bonds and buried SASA of chain B were increased. However, the total hydrogen bonds and buried SASA of CD16 were slightly decreased.
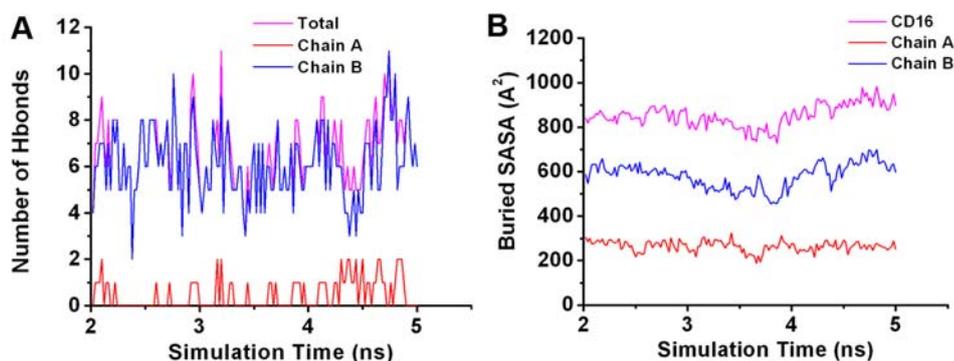
Figure 4.7 Interactions of the glycosylated CD16-IgG Fc complex. A. Number of hydrogen bonds between CD16 and IgG Fc. A hydrogen bond is defined by a <3.5 Å distance between donor and acceptor atoms and a >120° angle formed by donor, hydrogen, and acceptor atoms. **B.** Buried SASA due to binding.

Table 4.2
Hydrogen bonds[1] between CD16 and IgG Fc for the glycosylated CD16-IgG Fc complex

CD16	Chain A of IgG Fc	Existing Time (ns)	CD16	Chain B of IgG Fc	Existing Time (ns)
Lys19 N_ζ	Pro329 O	0.64	His116 $N_{\delta1}$	Gly236 N	2.58
Ser157 O_γ	Leu 235 O	0.52 (*)	[2]Lys117 N_ζ	Asp265 $O_{\delta2}$	2.96 (*)
			Lys117 N_ζ	Gly237 O	2.62
			Lys117 N_ζ	Ser239 O_γ	1.7
			His131 $N_{\delta1}$	Gly237 N	1.22
			Arg152 $N_{\eta2}$	NAG801 O_7	1.66 (*)
			Arg152 $N_{\eta1}$	NAG801 O_7	0.58 (*)
			[3]FUC802 $O_{2,3,4}$	[3]MAN4 $O_{2,5,6}$	2.74
			[3]FUC802 $O_{2,3,4}$	[3]NAG2 O_6	1.6

1. Only residue pairs, which formed hydrogen bonds for >0.5 ns, are listed.
2. Salt bridges.
3. FUC: fucose, MAN: mannose, NAG: N-Acetyl-Glucosamine pyranose.
* Existing in the unglycosylated CD16-IgG Fc complex.

In the glycosylated CD16-IgG Fc simulation, the RMSD of the D1 domain of CD16 from the complex crystal structure was always smaller than the RMSD from the standalone crystal structure (Fig. 4.8A), indicating that CD16 stayed in its complex conformation (Fig. 4.8C). Consistently, Trp95 still pointed inward and the hydrogen bond

between Tyr14 and Trp95 was formed most of time (Fig. 4.8B), although Trp95 rotated

~90 degree along the Cβ-Cγ bond (Fig. 4.8D).
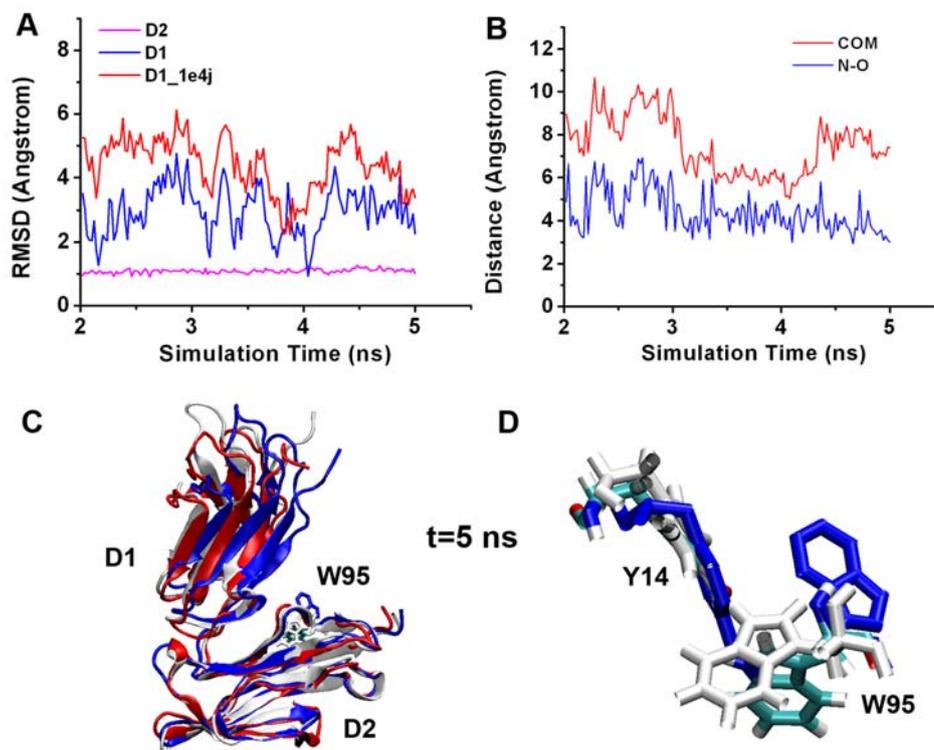
Figure 4.8 CD16 conformations in the glycosylated CD16-IgG Fc complex. A. RMSD from the CD16-IgG Fc complex crystal structure: Cα atoms of D2 domain (purple), using as alignment reference, and D1 domain (blue). RMSD from the standalone CD16 crystal structure: Cα atoms of D1 domain (red). **B.** Distance between the center of mass (COM) of W95 and the COM of Y14 (red), and between $N_{\varepsilon 1}$ of W95 and O_η of Y14 (blue). **C.** Snapshots of CD16 structure at 5 ns. White is crystal structure of CD16-IgG Fc complex, blue is crystal structure of standalone CD16, and red is simulated structure. Residue W95 are shown as sticks. **D.** Snapshots of W95 and Y14 at 5 ns. White sticks are from CD16-IgG Fc complex, blue sticks are from standalone CD16, and multiple color sticks are simulated (nitrogen in blue, oxygen atoms in red, carbon in cyan, and hydrogen in white).

4.3 Computational Binding Free Energy of the CD16-IgG Fc Complexes

To evaluate quantitatively the change of CD16-IgG Fc binding strength due to the conformational change upon CD16 glycosylation, the MM-PB(GB)SA method was used to calculate binding free energy of the CD16-IgG Fc complexes. 150 snapshots for each complex were extracted from the last 3 ns free MD simulation for MM-PBSA calculation.
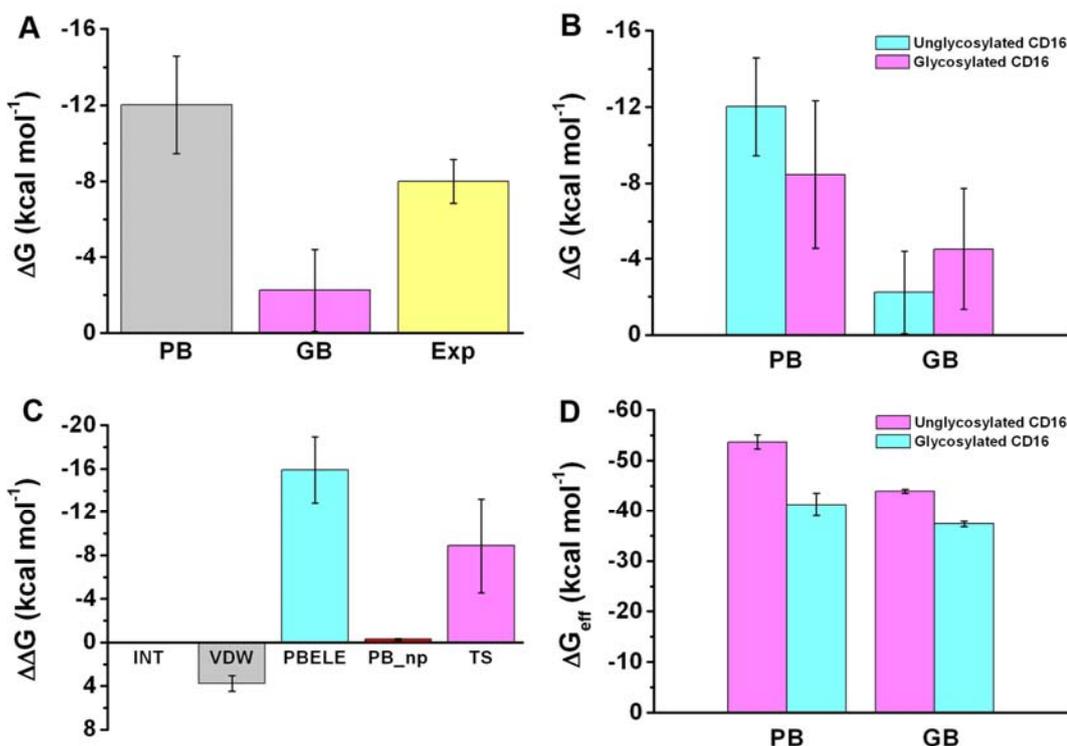
Figure 4.9 Computational binding free energy of the CD16-IgG Fc complexes. A. Comparison of computational binding free energy (ΔG) of the unglycosylated CD16-IgG Fc complex with experimental data from Maenaka et al. [73] PB: polar solvation energy calculated by Poisson-Boltzmann equation, GB: polar solvation calculated using Generalized Born model, Exp: experimental data. **B.** Comparison of computational binding free energy between the unglycosylated and glycosylated CD16-IgG Fc complexes. **C.** Components of binding free energy difference ($\Delta\Delta G$) between the two complexes. INT: internal energy including energy of bond stretching, angle bending, and torsion; VDW: van der Waals energy; PBELE: electrostatic energy in gas phase plus polar solvation energy calculated by Poisson-Boltzmann equation; PB_np: nonpolar solvation energy; TS: temperature times entropy. Total=INT+VDW+PBELE+PB_np-TS. **D.** Comparison of effective energy (ΔG_{eff}) that does not include entropy contribution. Error bars indicate standard error of mean.

Two methods were used to calculate the polar solvation energy: one is Poisson-Boltzmann equation (PB), and the other is Generalized Born model (GB). Compared to experimental data [73], PB gave a better result for the unglycosylated CD16-IgG Fc complex than GB (Fig. 4.9A). It is reasonable because GB is an approximation of PB. Comparing the binding free energy of the two complexes, PB predicted that the unglycosylated CD16-IgG Fc complex had more negative free energy while GB gave an opposite result (Fig. 4.9B). Standard errors are large, showing that it is very difficult for

Table 4.3
Binding free energy components of the CD16-IgG Fc complexes

[a]Components (kcal/mol)	Unglycosylated CD16		Glycosylated CD16	
	[b]Mean	[c]σ	[b]Mean	[c]σ
H_{elec}	-318.65	3.20	-245.09	1.67
H_{vdW}	-82.39	0.50	-86.16	0.48
H_{int}	0.01	0.00	0.01	0.00
H_{gas}	-401.04	3.18	-331.23	1.62
PB_{np}	-9.18	0.04	-8.89	0.04
PB_{polar}	356.50	3.95	298.85	2.53
PB_{sol}	347.33	3.93	289.96	2.51
PB_{ele}	37.85	1.49	53.76	2.47
PB_{eff}	-53.71	1.41	-41.27	2.24
GB_{np}	-7.62	0.04	-7.35	0.04
GB_{polar}	364.71	3.00	301.24	1.60
GB_{sol}	357.10	2.98	293.88	1.61
GB_{eff}	-43.94	0.42	-37.35	0.52
TS_{trans}	-16.88	0.00	-16.91	0.00
TS_{rot}	-16.92	0.01	-16.96	0.01
TS_{vib}	-9.74	2.13	-0.79	3.15
$TS_{trans+rot+vib}$	-43.54	2.14	-34.66	3.15
$H_{tra/rot}$	-1.85	0	-1.85	0
PB_{total}	-12.01	2.56	-8.46	3.86
GB_{total}	-2.25	2.18	-4.54	3.19
ΔG_{exp}	-8.00	1.16		

[a] H_{elec}, electrostatic energy; H_{vdW}, van der Waals energy; H_{int}, internal energy including energy of bond stretching, angle bending, and torsion; $H_{gas}= H_{elec} + H_{vdW} + H_{int}$; PB_{np} and PB_{polar}, nonpolar and polar solvation energy calculated using the PB method; solvation energy $PB_{sol} = PB_{np} + PB_{polar}$; effective energy $PB_{eff} = H_{gas} + PB_{sol}$; GB_{np} and GB_{polar}, nonpolar and polar solvation energy calculated using Generalized Born model; salvation energy $GB_{sol} = GB_{np} + GB_{polar}$; effective energy $GB_{eff} = H_{gas} + GB_{sol}$; TS_{trans}, translational entropy of solute; TS_{rot}, rotational entropy of solute; TS_{vib}, vibrational entropy of solutes; total entropy of solute $TS_{trans+rot+vib} = TS_{trans} + TS_{rot}$; total free energy $PB_{total} = PB_{gas+sol} + TS_{trans+rot+vib}$; total free energy $GB_{total} = GB_{gas+sol} + TS_{trans+rot+vib}$.
[b] Average over 150 snapshots, except that entropy are averaged over 30 snapshots.
[c] Standard error of mean.

the MM-PBSA method to predict a small free energy difference (<10 kcal/mol). On the

other hand, PB most likely gave a better prediction as demonstrated. Therefore, the more

negative free energy, calculated by PB, indicated that the unglycosylated complex had a

higher binding affinity, which is consistent with experimental data (Fig. 2.4). The

difference of ~3.55 kcal/mol between the two complexes corresponds to ~350 folds

binding affinity difference ($K_A = \exp\left(-\dfrac{\Delta G}{RT}\right)$), which is one order higher than

experimental data.

Among all components of the free energy difference between the two complexes

($\Delta\Delta G$), van der Waals energy (VDW), electrostatic energy (PBELE), and entropy portion

(TS) have the biggest contribution (Fig. 4.9C and Table 4.3). PBELE, which includes

electrostatic energy in gas phase and polar solvation energy, has a large negative value,

while VDW has a small positive value. This demonstrates that the change of electrostatic

energy is more significant and the unglycosylated complex has more favorable

electrostatic interactions, which are consistent with the observed change of hydrogen

bonds. Entropy compensates for the change of electrostatic energy. More negative

electrostatic energy, more loss of entropy. So the total free energy change was small,

although its components are big.

The vibrational entropy of solute was calculated using normal mode analysis and

very time consuming, so it was computed only for 30 snapshots. As a result, it has the

biggest standard error among all the components of free energy. If the entropy portion is

excluded, the other part of free energy (effective energy) shows the same tendency of

binding free energy difference for PB and GB computations (Fig. 4.9D). In order to

identify important residues for binding, a residue-based free energy decomposition was

performed for effective energy of GB because only GB polar solvation energy can be

decomposed to each residue. The free energy of each critical residue, which is significant

negative, is displayed like a spike on the free energy curves (Fig. 4.10). Several important

regions were identified. The lower hinges of both Cγ3-A and Cγ3-B of IgG Fc had a big

contribution. Upon CD16 glycosylation, the free energy of the lower hinge of Cγ3-A
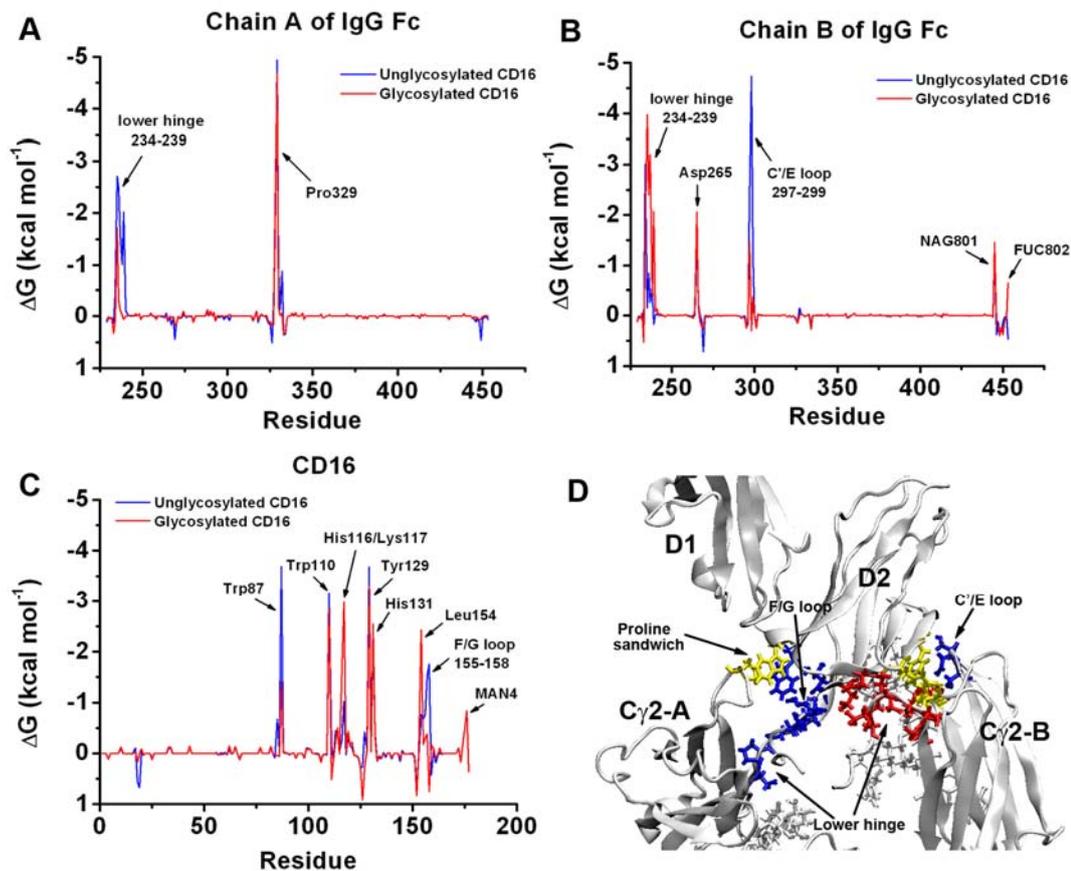
Figure 4.10 Residue-based decomposition of the effective energy of the CD16-IgG Fc complexes. Contribution of each residue of chain A of Fc (**A**), chain B of Fc (**B**), and CD16 (**C**). FUC: fucose, MAN: mannose, NAG: N-Acetyl-Glucosamine pyranose. **D.** Bind pockets of CD16-IgG complexes. Residues with biggest contribution to the effective energy are highlighted in colors: blue residues have more negative energy in the unglycosylated complex, red residues have more negative energy in the glycosylated complex, and yellow residues have almost equal energy in both complexes.

became less negative while the free energy of the lower hinge of Cγ3-B turned more

negative due to the shift of CD16 to chain B. "Proline sandwich", including Pro329 of

chain A of Fc and Trp87 and Trp110 of CD16, also contributed significantly to the

binding free energy. Upon CD16 glycosylation, the proline sandwich was weakened and

Trp87 became less negative. The C'/E loop (residue 297-299) of the Cγ3-B domain had a

large negative free energy, which became much less when CD16 was glycosylated.

Residue His116, Lys117, and Leu154 of CD16, which interacted with the lower hinge of

Cγ3-B, had much more negative free energy upon CD16 glycosylation. In contrast, the F/G loop (residue 155-158) of CD16, which interacted with the lower hinge of Cγ3-A, had much less negative free energy when CD16 was glycosylated. As observed interactions between the two carbohydrate chains of CD16 and Cγ3-B, sugar residues MAN4 of CD16 and FUC802 of chain B of Fc had a negative free energy. Other residues, including Asp265 and NAG801 of Cγ3-B and Tyr129 and His131 of CD16, also contributed largely to the binding free energy but no significant change upon CD16 glycosylation.

CHAPTER 5

UNFOLDING OF THE VWF A DOMAINS UNDER FORCE OR

HIGH TEMPERATURE

As described in section 2.2, the unfolding of VWF A2 domain facilitates its cleavage by ADAMTS-13, which assures correct length distribution of VWF multimers in blood circulation. Here, MD was used to simulate at atomic details the dynamic process of the A2 domain unfolding by tensile force or under high temperature. As a comparison, the A1 domain unfolding was also studied.

5.1 Unfolding of the A1 and A2 Domains by Tensile Force

Equilibration of A1 and A2 in Water Spheres

To prepare for the SMD simulations of the unfolding, the crystal structure of the A1 domain (pdb 1auq, Fig. 2.7B) [22] and the homology model of the A2 domain (Fig. 2.7C) [25] were soaked in water spheres with diameters of 75 Å and 70 Å, respectively (Fig. 5.1 A and B). After energy minimization and heat-up to 300 K, the two systems were equilibrated under constant temperature without restraints for 1 ns. In both cases, the RMSDs of all Cα atoms displayed plateaus (Fig. 5.1C), indicating that the systems reached equilibrium quickly. The RMSD of both systems was smaller than 2 Å, showing that the structural changes were small during equilibration. However, the RMSD of the A2 domain was larger than that of the A1 domain (the N- and C-terminal loops of the A1 domain were excluded for RMSD calculation because they are not in the core residues Cys1272-Cys1458 within the disulfide bond), indicating more changes in the equilibrated

36
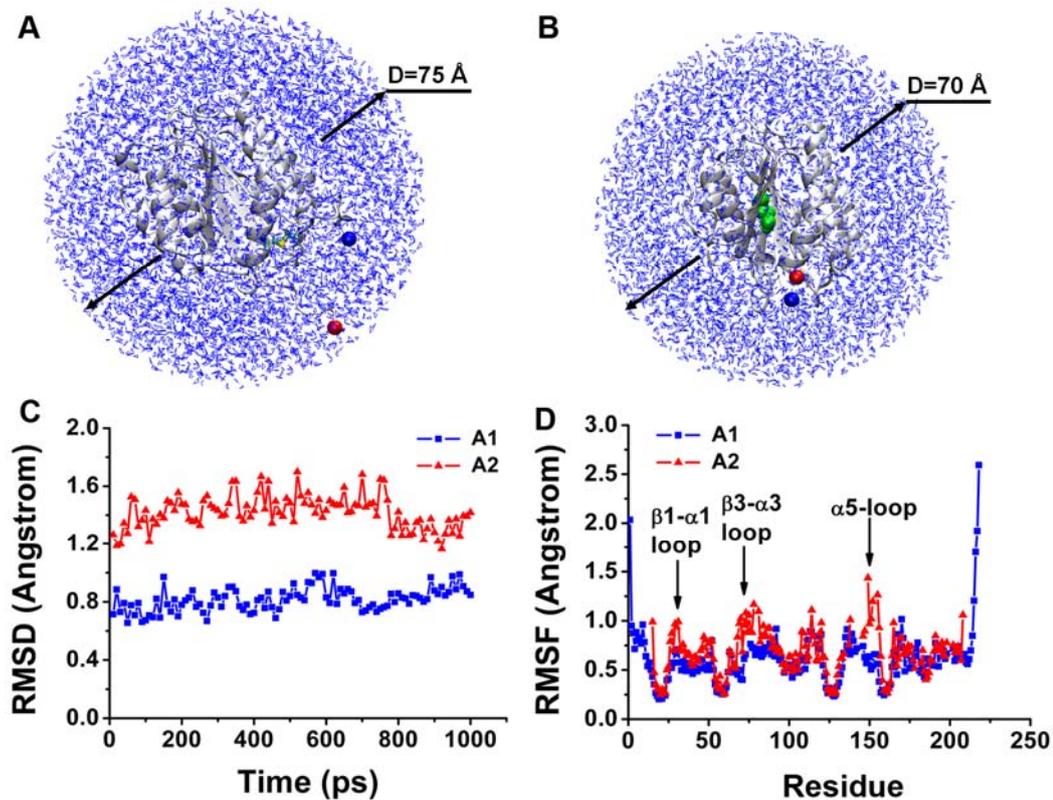
Figure 5.1 Equilibration of the A1 and A2 domains in water spheres. The A1 (**A**) and A2 (**B**) domains were soaked in water spheres. Water molecules are colored blue. **C.** RMSD of all Cα atoms using the starting structures as references. Structures at different time were aligned with the β1, β2, β4, and β5 strands of the corresponding starting structures. **D.** RMSF of each Cα atom averaged over 1 ns equilibration. Structures at different time were aligned to the starting structure via the four strands (β1, β2, β4, and β5). The residue numbers of A1 and A2 are shown according to their sequence alignment in Fig. 1 of Sutherland et al.[25] Residue 1 of A1 corresponds to residue 1261 in VWF sequence, while residue 15 of A2 corresponds to residue 1496 in VWF sequence. Loops that have large RMSF are indicated by arrows. Representations of A1 and A2 refer to Fig. 2.7.

structure of the A2 domain. This is probably because the initial structure of the A2

domain is from a homology model and may not be as close to its native structure as the

initial structure of the A1 domain, which is a crystal structure. Furthermore, the RMSF

for each Cα atom was calculated. The RMSF of loops β1-α1, β3-α3, and α5 of the A2

domain was larger than that of the corresponding parts of the A1 domain (Fig. 5.1D),

indicating the A2 domain is more flexible than the A1 domain. Outside the A1 domain
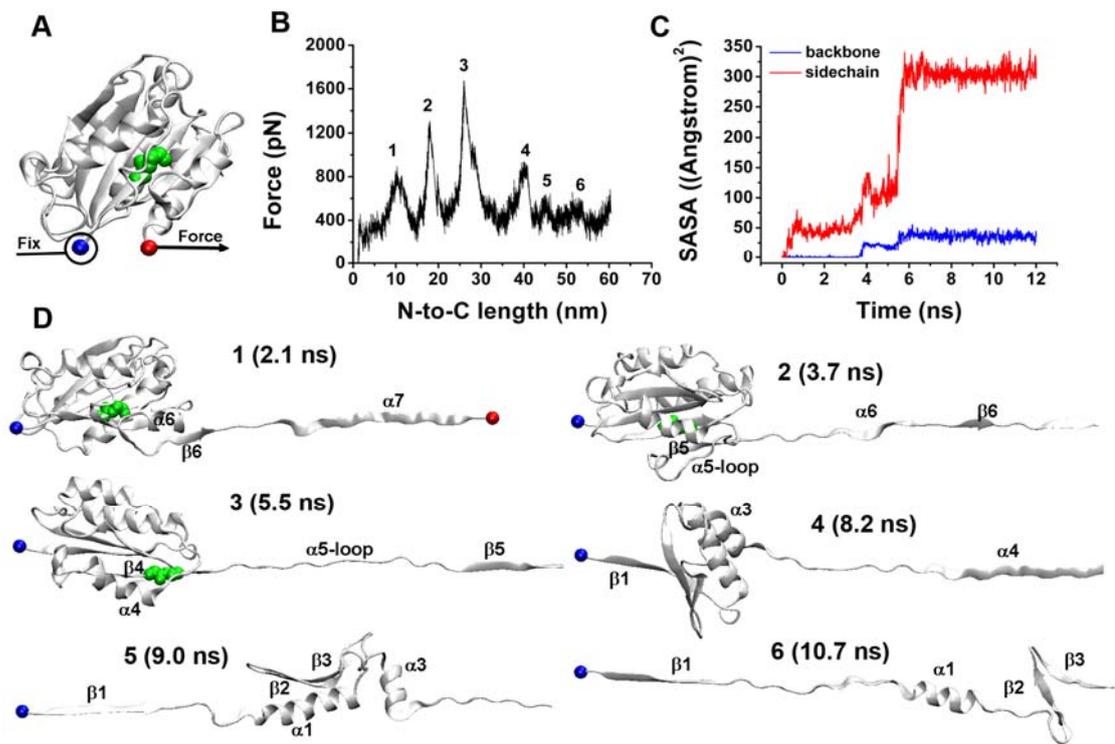
Figure 5.2 The A2 domain unfolding by pulling at the C-terminus. A. The start point (0 ns) of the SMD simulation. A force was pulling rightward on the C-terminal Cα atom (red ball) with the N-terminal Cα atom (blue ball) fixed. **B.** Plot of force vs. distance between the N-terminal Cα atom and the C-terminal Cα atom (N-to-C length) during the SMD simulation. The force peaks are labeled with numbers 1-6. **C.** SASA of the proteolytic site vs. simulation time. **D.** Snapshots of the simulated A2 domain structure. Numbers 1-6 label snapshots (taken at indicated times) that correspond to the force peaks 1-6 in B. Representations of A2 refer to Fig. 2.7C.

core, the N- and C-terminal loops were much more flexible, with large RMSF, as

expected.

Unfolding of the VWF A2 Domain by a Tensile Force

To simulate unfolding by force with SMD, the C-terminal Cα atom of the A2

domain was pulled at a constant speed of 5 nm ns^{-1} through a spring (spring constant of 1

kcal mol^{-1} Å^{-2}) attached to the C-terminal Cα atom while the N-terminal Cα atom was

fixed (Fig. 5.2A). The force-extension curve displayed six peaks (Fig. 5.2B). The first

Figure 5.3 The A2 domain unfolding by pulling at the N-terminus. A. The start point (0 ns) of the SMD simulation. A force was pulling leftward on the N-terminal Cα atom (blue ball) with the C-terminal Cα atom (red ball) fixed. **B.** Plot of force vs. distance during the SMD simulation. The force peaks ars labeled with numbers 1-6. **C.** SASA of the proteolytic site vs. simulation time. **D.** Snapshots of the simulated A2 domain structure. Numbers 1-6 label snapshots (taken at indicated times) that correspond to the force peaks 1-6 in B. Representations of A2 refer to Fig. 2.7C.

three peaks corresponded to the sequential pullouts of the β6, β5, and β4 strands of the

central β-sheet from the C-terminus (Fig. 5.2D, 1-3). Instead, the fourth peak showed the

pullout of the β1 strand from the N-terminus (Fig. 5.2D, 4). The second and third peaks

had higher peak forces than the first and fourth peaks, indicating stronger intradomain

interactions of the β5 and β4 strands. Unfolding of the remaining structure, which

included the α1 and α3 helices and the β2 and β3 strands (Fig. 5.2D, 5), resulted in two

peaks that were much smaller than the first four peaks. The fifth peak appeared when the

α3 helix was pulled out and unfolded (Fig. 5.2D, 5). Finally, the breakage of the β2-β3

hairpin gave the sixth peak (Fig. 5.2D, 6).

39

To examine how the proteolytic site would become exposed during the unfolding of the A2 domain, the SASA for the backbone and the sidechain of Tyr1605 and Met1606, two residues adjacent to the proteolytic site on the β4 strand, was calculated (Fig. 5.2C). The SASA curve of the sidechain showed three steps while that of the backbone had two steps. Initially, the two residues were completely buried, as indicated by the zero SASA of the sidechain and the backbone. After the pulling was started, the α7 helix was pulled out first to partially expose the proteolytic site, resulting in a sudden jump in the SASA of the sidechain to 50 $Å^2$. But the SASA of the backbone was still zero because the backbone of the β4 strand was buried by its neighboring β1 and β5 strands. Then, one side of the backbone was exposed when the β5 strand was pulled out (Fig. 5.2D, 2), producing a sudden jump in the SASA of both the sidechain and the backbone. Finally, the β4 strand on which the proteolytic site resides was pulled out (Fig. 5.2D, 3), giving rise to the last jump in the SASA values.

Considering that the pulling speed in SMD is several orders of magnitude higher than that in experiments such as those done with atomic force microscopy (AFM), there is always a question whether the force propagates well through an object that is pulled. To address this for the A2 domain, a second SMD simulation was run where force was loaded along the opposite direction on the N-terminal Cα atom with the C-terminal Cα atom fixed (Fig. 5.3A) and the same pulling speed and spring constant. The results showed that the A2 domain unfolding still started from the C-terminus (Fig. 5.3D). Just as the first SMD, the β6, β5, and β4 strands of the A2 domain were pulled sequentially out of the central β-sheet (Fig. 5.3D, 1-3), which resulted in the first three peaks in the force-extension curve (Fig. 5.3B). Then a much smaller peak was observed, indicating

40

the pullout and unfolding of the α4 helix (Fig. 5.3D, 4). This peak was not observed in the first SMD simulation. After the fourth peak, the β1 strand was pulled out from the N-terminus (Fig. 5.3D, 5), resulting in the fifth peak, which corresponded to the fourth peak in the first SMD simulation. The unfolding of the remaining structure (α1 and α3 helices and β2 and β3 strands) generated some small peaks (Fig. 5.3D, 6), but much less pronounced than those in the first SMD simulation. Since most features of the two simulated unfolding processes were similar with the four major peaks (corresponding to the pullouts of β6, β5, β4, and β1) appearing in both force-extension curves in the same sequence, it was concluded that the A2 domain unfolded along the same pathway in both cases. This suggests that the force actually propagated well through the whole A2 domain regardless of whether it was pulled at the N- or the C-terminus.

Just as in the first SMD simulation, the SASA curves of the sidechain and the backbone of the proteolytic site include three and two steps, respectively (Fig. 5.3C). The exposure of the proteolytic site became more gradually: the pullout of the α7 helix first exposed part of the sidechain of the proteolytic site but not the backbone, the pullout of the β5 strand then exposed one side of the backbone, the pullout of the β4 strand finally fully exposed the proteolytic site.

Unfolding of the VWF A1 Domain by a Tensile Force

As mentioned, the A1 domain has a disulfide bond that links its C- and N-termini. It is therefore expected that the A1 domain is harder to unfold than the A2 domain, especially under tensile force. When a spring (spring constant = 1 kcal mol^{-1} Å^{-2}) attached to the C-terminal Cα atom of the A1 domain was pulled at a constant speed of 5 nm ns^{-1}

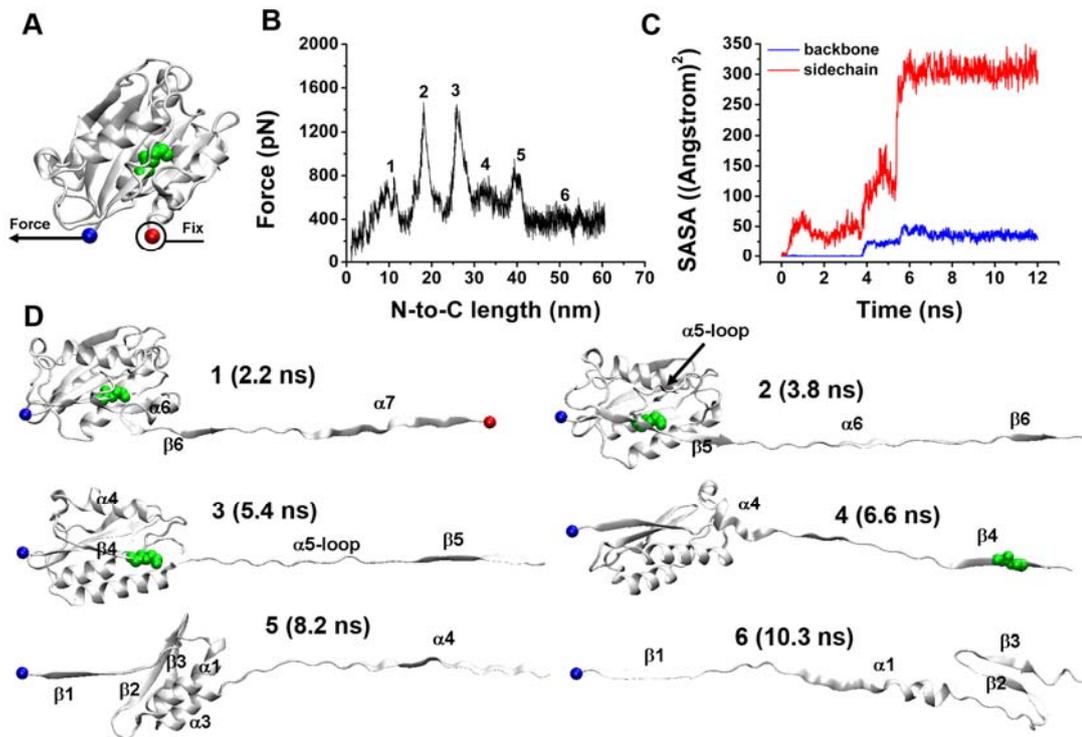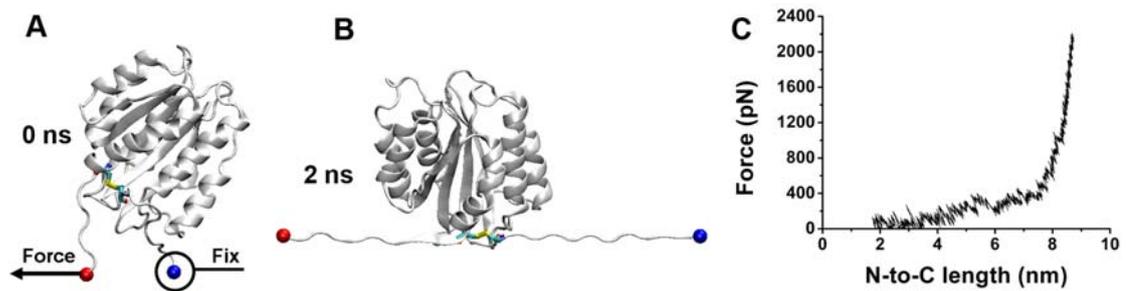**Figure 5.4 Unfolding of the A1 domain with an intact disulfide bond by pulling at the C-terminus. A.** The start point (0 ns) of the SMD simulation. A force was pulling rightward on the C-terminal Cα atom (red ball) with the N-terminal Cα atom (blue ball) fixed. **B.** The structure of the A1 domain at the end of the SMD simulation (2 ns). **C.** Plot of force vs. distance during the SMD simulation. Representations of A1 refer to Fig. 2.7B.

(the N-terminal Cα atom fixed, Fig. 5.4A), the force increased gradually to a value higher than 2,000 pN (Fig. 5.4C). The force could have increased to infinite because the disulfide bond is not allowed to break in a classic MD simulation. At 2 ns, the end of the SMD simulation, the N- and C-terminal loops of the A1 domain were stretched tautly (Fig. 5.4B), but the core of the A1 domain was not affected at all due to the presence of the disulfide bond.

In reality, the disulfide bond can be broken by reducing agents such as DTT [74, 75], enzymes that catalyze disulfide bond reduction such as Thioredoxin [76], or even a very large force. Since the disulfide bond in the A1 domain is exposed, it is susceptible to reduction. In order to simulate the unfolding of the A1 domain after reduction of its disulfide bond, we manually broke the disulfide bond at the beginning of the SMD simulation while keeping other setup of the simulation the same (Fig. 5.5A). During the pulling, the α7 helix was pulled out at first (Fig. 5.5C, 1). The unfolding of the α7 helix resulted in small force peaks in the force-extension curve (Fig. 5.5B). Then the β6 strand was pulled out, generating a small force peak (Fig. 5.5C, 2). After that, the β5 strand was pulled out (Fig. 5.5C, 3), causing the third force peak, followed by the pullout of the β4

42
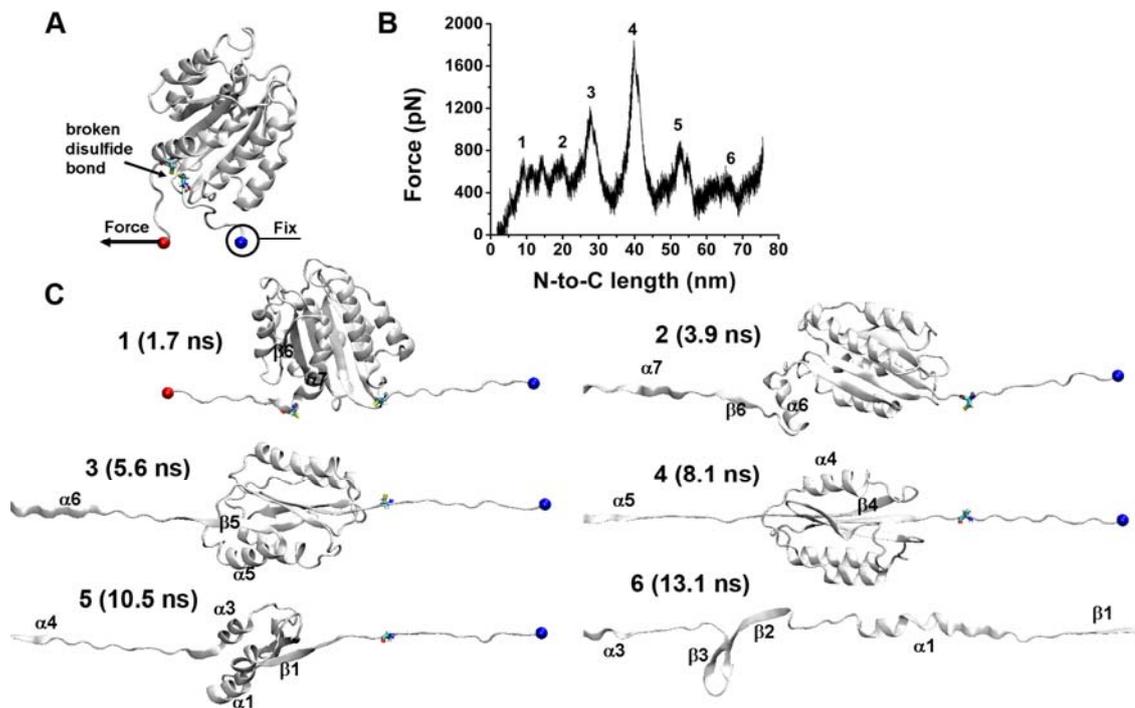
Figure 5.5 Unfolding of the A1 domain with a broken disulfide bond by pulling at the C-terminus. A. The start point (0 ns) of the SMD simulation. The disulfide bond was broken *in silico*. A force was pulling rightward on the C-terminal Cα atom (red ball) with the N-terminal Cα atom (blue ball) fixed. **B.** Plot of force vs. distance during the SMD simulation. The force peaks are labeled with numbers 1-6. **C.** Snapshots of the simulated A1 domain. Numbers 1-6 label snapshots (taken at indicated times) that correspond to the force peaks 1-6 in B. Representations of A1 refer to Fig. 2.7B.

strand, producing the highest force peak (Fig. 5.5C, 4). The β1 was then pulled out from

the N-terminus, yielding the fifth force peak with only α1, α2, β1, and β2 in the

remaining structure (Fig. 5.5C, 5). Finally, the remaining structure was unfolded (Fig.

5.5C, 6), giving rise to the sixth force peak.

Hydrogen Bond Network of β Sheet Determines Unfolding Resistance to Tensile

Force

As shown above, the A2 domain unfolding always started from the C-terminus

with the same sequence of β strand pullouts regardless of whether the A2 domain was

pulled at the C- or N-terminus. When it was pulled at the C-terminus, the A1 domain
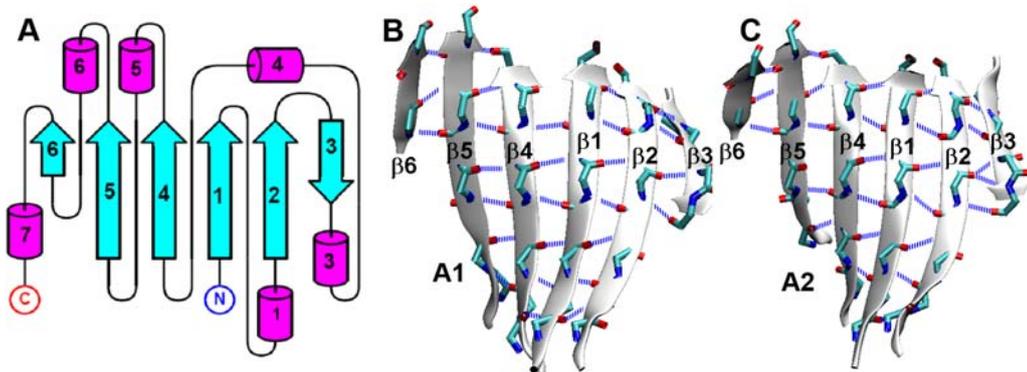
Figure 5.6 Topology and hydrogen bond networks of the central β-sheet of the A domains. A. The topology of the A domains. Cyan arrows indicate β strands and purple cylinders represent α helices (α5 is a loop in A2). Generated by TopDraw [77]. A1 (**B**) and A2 (**C**) hydrogen bond (blue dash line) networks in the central β-sheet.

unfolding with a broken disulfide bond also started from the C-terminus. Since the

sequences of the A1 and A2 domains are quite different (only ~20% identity), the

similarity of their unfolding pathways may be attributed to their specific topology (Fig.

5.6A). Since the C-terminus of the A domains started at the β6 strand on one edge of the

central β-sheet, which forms only 3 hydrogen bonds with the β5 strand (Fig. 5.6 B and

C), the β6 strand can be easily pulled out. In contrast, the N-terminus of the A domains

started at the β1 strand right in the middle of the central β-sheet, which has more than 10

hydrogen bonds with the neighboring β4 and β2 strands (Fig. 5.6 B and C), which is

harder to break. Only after one of the neighboring β strands is pulled out to expose the β1

strand can it be pulled out. This is exactly what was observed in the forced unfolding

simulations.

The force peaks in the force-extension curves are attributed as the forces required

to pull out the β strands (i.e., pullout forces) because the force peaks were observed to

follow the β strand pullouts during the unfolding processes. To identify the atomic-level

interactions that contribute to the pullout forces, the changes of the hydrogen bonds of the
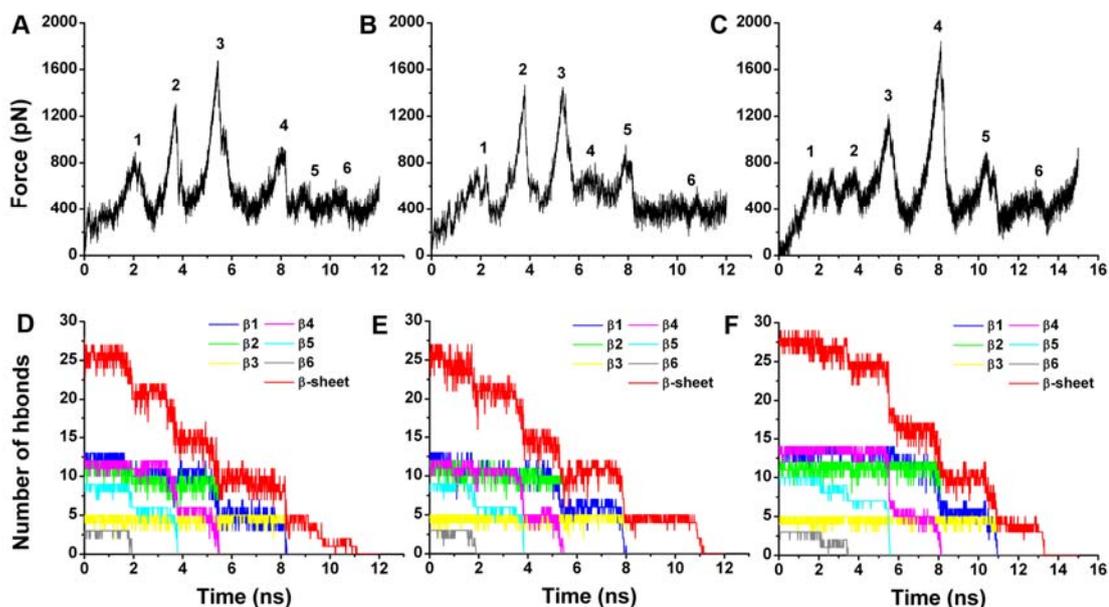
Figure 5.7 Coincidence of pullout forces and changes in numbers of hydrogen bonds. Figs. 5.2B, 5.3B, and 5.5B are replotted in **A**, **B**. and **C**. Numbers of hydrogen bonds of the β-sheet and each strand (indicated) are plotted vs. time for A2 pulled at C- (**D.**) or N- (**E**) terminus, or A1 without the disulfide bond pulled at C-terminus (**F**). A hydrogen bond is defined by a <3.5 Å distance between donor and acceptor atoms and a >120° angle formed by donor, hydrogen, and acceptor atoms.

central β-sheet were monitored. It is evident from Fig. 5.7 that the force peaks coincided

with the step decrease of the hydrogen bond number, except for the 4th peak in Fig. 5.7C,

which corresponded to the unfolding of the α4 helix. Therefore, most resistance to

unfolding by force came from the hydrogen bond network of the central β-sheet.

Different β strands required different magnitudes of pullout forces. The pullout

forces for the β5 and β4 strands (peak 2 and 3 for A2 or peak 3 and 4 for A1) were larger

than the others because more than five hydrogen bonds have to be broken in order to pull

out the β5 and β4 strands (Fig. 5.6). To pull out the β1 strand also needs to break more

than five hydrogen bonds; but it required a much lower pullout force (peak 4 for pulling

A2 at the C-terminus, peak 5 for pulling A2 at the N-terminus, or peak 5 for pulling A1 at

the C-terminus). The reason for this difference lies in the different unfolding pathways.

The pullouts of the β5 and β4 strands followed a sliding pathway with two β strands
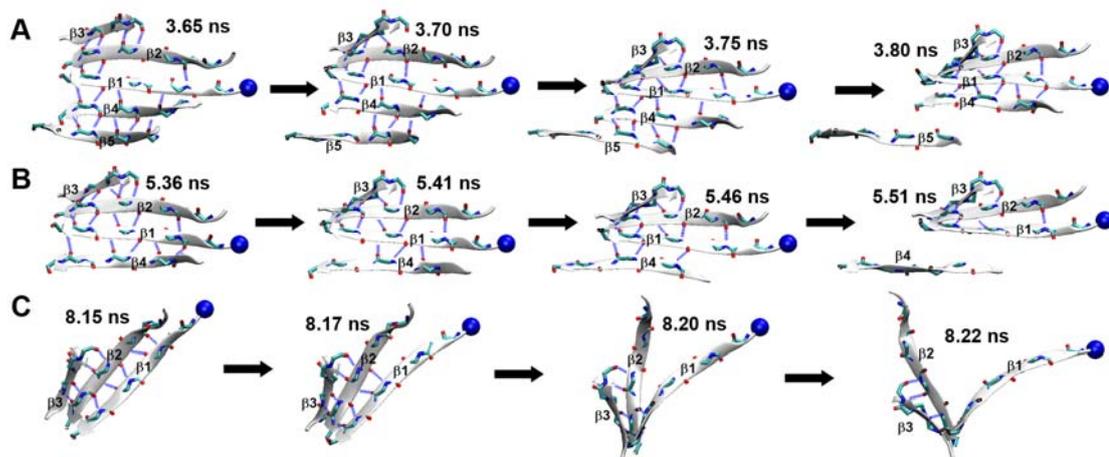
Figure 5.8 Two pathways of β strand pullouts. The β5 (**A**) and β4 (**B**) were pulled out along a sliding pathway. **C.** The β1 pullout followed an unzipping pathway. β strands are shown as ribbons with arrows, backbones are shown as sticks, hydrogen bonds are represented by blue dash lines, and the N-terminal is marked by a blue sphere.

sliding relative to each other in the direction along the long axis (Fig. 5.8 A and B). The old hydrogen bonds broke, and then the β strand started to slide. During sliding, new hydrogen bonds formed and then broke. For sliding to occur, all hydrogen bonds had to be broken at the same time, resulting in a large force. In contrast, the pullout of the β1 strand was along an unzipping pathway (Fig. 5.8C). The hydrogen bonds started to break from one end, and then propagated to the other end. Only one hydrogen bond broke each time, therefore requiring a small force.

Helices Unfolding is Different between A1 and A2

As mentioned, the A1 and A2 domains unfolded quite similarly in terms of the β strand pullouts. To further quantify the changes of secondary structures during unfolding, mean residue molar ellipticity in circular dichroism of the A domain structures was calculated as a linear combination of basic spectra of secondary structures (helix, β-sheet, turn, and coil) [78]. First, calculated spectra of the A1 and A2 domain were compared with experimental measurements [29]. The calculated spectra were comparable to the
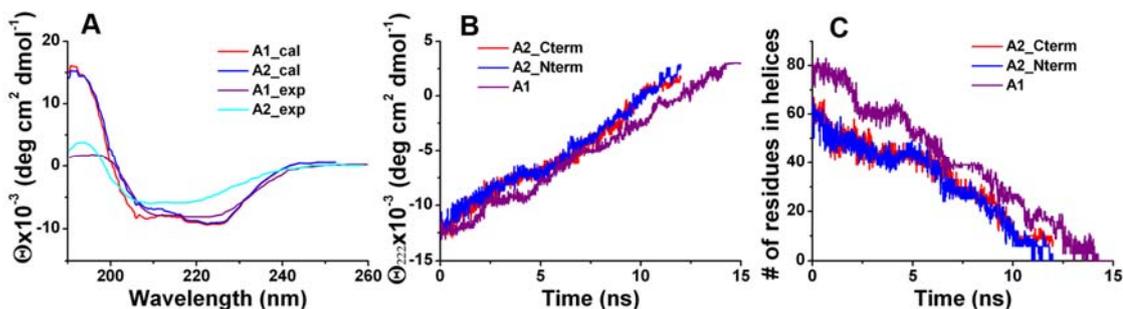
Figure 5.9 Mean residue molar ellipticity of the A domain structures. A. Calculated mean residue molar ellipticity (Θ) spectra of A1 and A2 (A1_cal and A2_cal) are compared with experimental values (A1_exp and A2_exp) measured in circular dichroism (from Ref. 18). **B.** The simulation time dependent ellipticities at 222 nm (Θ_{222}) of A2 pulled at C- (A1_Cterm) or N- (A1_Nterm) terminus, or A1 with a broken disulfide bond (A1). **C.** Number of residues in the helices of A2 pulled at C- (A1_Cterm) or N- (A1_Nterm) terminus, or A1 with a broken disulfide bond (A1) are plotted vs. simulation time.

experimental values of the A1 domain in wavelengths ranging from 200-260 nm, but differed in the 190-200 nm wavelength range (Fig. 5.9A). The experimental data showed significant differences between the A1 and A2 domains. But the computational results did not show much difference. This may indicate that the homology model of the A2 domain is different from its native structure. Furthermore, the ellipticity at 222 nm (where the helix spectrum are dominant) during unfolding was calculated. The curves of the A2 domain pulled at the C- or N-terminus almost overlapped (Fig. 5.9B), indicating the unfolding of secondary structures was quite similar. However, the curves of the A1 domain were different and had clear stepwise increase. This is because the helices of the A1 domain were unfolded in steps while the unfolding of the helices of the A2 domain was more continuous (Fig. 5.9C).

5.2 Thermal Unfolding of the A1 and A2 Domains under High Temperature

Equilibration of A1 and A2 in Water Boxes
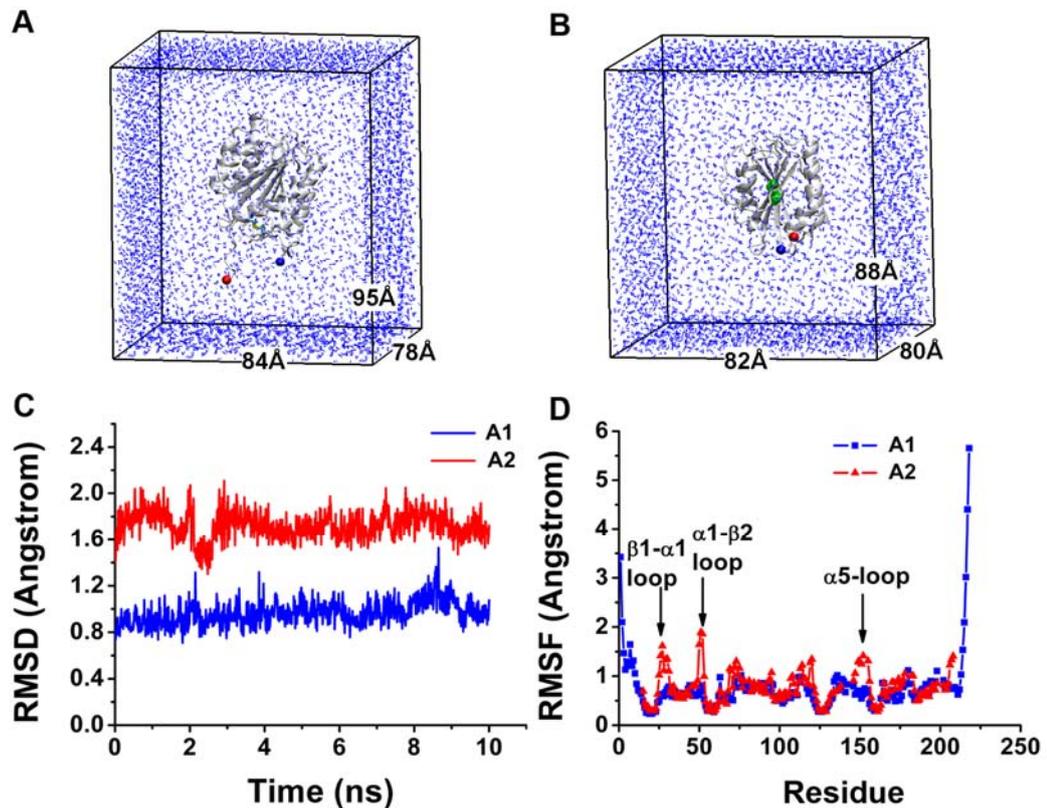
Figure 5.10 Equilibration of the A1 and A2 domains in water boxes. The A1 (**A**) and A2 (**B**) domains were soaked in water boxes. Water molecules are colored blue. **C.** RMSD of all Cα atoms using the starting structures as references. Structures at different time were aligned with the β1, β2, β4, and β5 strands of the corresponding starting structures. **D.** RMSF of each Cα atom averaged over 10 ns equilibration. Structures at different time were aligned to the starting structure via the four strands (β1, β2, β4, and β5). The residue numbers of A1 and A2 are shown according to their sequence alignment in Fig. 1 of Sutherland et al.[25] Residue 1 of A1 corresponds to residue 1261 in VWF sequence, while residue 15 of A2 corresponds to residue 1496 in VWF sequence. Loops that have large RMSF are indicated by arrows. Representations of A1 and A2 refer to Fig. 2.7.

How would the SMD-simulated unfolding compare with unfolding by denaturants (e.g., urea)? The pathways of unfolding by denaturants and by tensile force were expected to be different. For example, the A1 domain can be unfolded upon addition of urea even with its intact disulfide bond [29]. To mimic the unfolding of the A domains

48

with denaturants, MD simulations under high temperature were performed because the thermal forces caused by Brownian motions act on all the part of the A domains as denaturants do. Indeed, MD under high temperature has been widely used to study protein folding and unfolding [79-83].

The A1 and A2 domains were soaked in water boxes (Fig. 5.10 A and B) instead of water spheres because water boxes were more convenient to adjust the pressure to avoid artifacts caused by high pressure when temperature was increased. After energy minimization and heat-up to 300 K, the two systems were equilibrated under constant temperature and constant pressure without restraint for 10 ns. The RMSD of all Cα atoms rapidly arrived at plateaus (Fig. 5.10C), indicating that equilibrium was reached. Similar to water spheres, the RMSD of the A2 domain was higher than that of the A1 domain probably because the structure of the A1 domain is closer to the native structure. The A2 domain displayed higher RMSF in loops β1-α1, α1-β2, and α5 than the corresponding parts of the A1 domain (Fig. 5.10D), suggesting that the A2 domain was more flexible than the A1 domain. The N- and C-terminal loops of the A1 domain out of the disulfide bond Cys1272-Cys1458 showed the largest RMSF because they were most flexible.

Thermal Unfolding of the VWF A2 Domain at 500 K

After equilibration at 300 K, the temperature of the A2 system was increased to 500 K in 100 ps with constant volume, so the pressure was elevated to 4 kilobar. Then the pressure was adjusted to 1 atm (1.01325 bar) by changing the system volume (Fig. 5.11A). As a result, the density of the system was lowered to ~0.7 g ml^{-1} (Fig. 5.11B), similar to published simulations with SPC water model [84]. During the heat-up and the pressure adjustment, the A2 domains were restrained. In preliminary simulations, higher
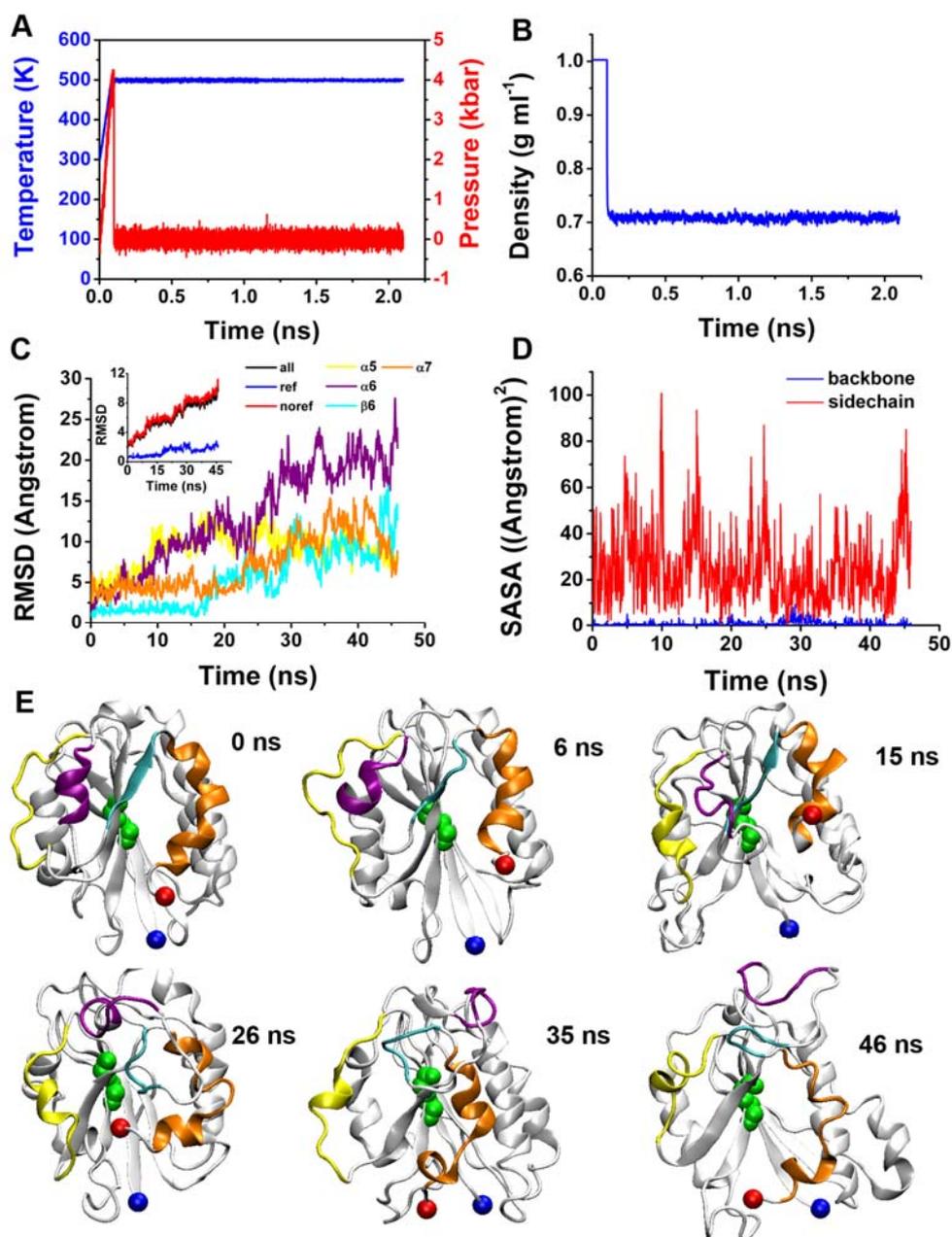
Figure 5.11 Thermal unfolding of the A2 domain at 500 K. A. The temperature of the system (blue lines) was heated up from 300 K to 500 K in 100 ps. Then the pressure (red lines) was adjusted to 1 atm (1.01325 bar). **B.** During the pressure adjustment, the density of the system decreased from ~1 g ml^{-1} to ~0.7 g ml^{-1}. **C.** RMSD of Cα atoms of α5, α6, β6, and α7 using the starting structure as reference is plotted versus simulation time. Structures at different time were aligned to the starting structure via the four strands (β1, β2, β4, and β5). Inset include RMSDs of all Cα atoms (all), Cα atoms of β1, β2, β4, and β5 used for alignment (ref), or Cα atoms of other than the four strands (noref). **D.** SASA of the proteolytic site is plotted versus simulation time. The SASA of the backbone of Tyr1605 and Met1606 at the proteolytic site is shown as blue lines while the SASA of the sidechain is indicated as red lines. **E.** Snapshots of the structure of the A2 domain during the thermal unfolding

simulation. The α7 helix is colored in orange, the β6 strand in cyan, the α6 helix in purple, and the α5-loop in yellow. The time of each snapshot is indicated in the parenthesis. 0 ns is the moment when the restraints on the A2 domain were turned off. Representations of A2 refer to Fig. 2.7C.

temperature was tried in order to accelerate the unfolding process. But water was rapidly evaporated to gas phase with higher temperature such as 600 K. Finally, all the restraints on the A2 domain were released and its unfolding was observed under constant temperature (500K) and constant pressure (1 atm).

The RMSD of all Cα atoms of the A2 domain displayed stepwise increase during its unfolding process (inset of Fig. 5.11C). The most significant changes occurred at the C-terminal portion of the A2 domain. The β6 strand, the α5, α6, and α7 helices displayed the largest RMSDs (Fig. 5.11C). First, the α6 and α7 helices moved away from the central β-sheet, which pushed the α5-loop aside and opened a space for exposure of the proteolytic site (Fig. 5.11E, 6 ns). Second, the α6 helix was unfolded into a loop (α6-loop) while part of the α5-loop became a helix (Fig. 8E, 15 ns). Third, the α6-loop moved upward, the β6 strand was disconnected from the central β-sheet and converted into a loop (β6-loop), and the α7 helix was split into two helices (Fig. 5.11E, 26 ns). Fourth, the α6-loop moved rightward while the β6-loop moved upward (Fig. 5.11E, 35 ns). At the end of the simulation, the α6-loop and β6-loop moved further away from the protein core while the α7 helix was mostly unfolded (Fig. 5.11E, 46 ns), so the RMSD kept increasing.

During the unfolding of the A2 domain, the SASA of the sidechain of the proteolytic site fluctuated around 20 Å^2 (Fig. 5.11D), showing partial exposure of the sidechain. In contrast, the SASA of the backbone was almost zero because the

interactions between the β4 strand, where the proteolytic site is, and its neighbour strands (β1 and β5) were still intact.

Thermal Unfolding of the VWF A1 Domain at 500 K
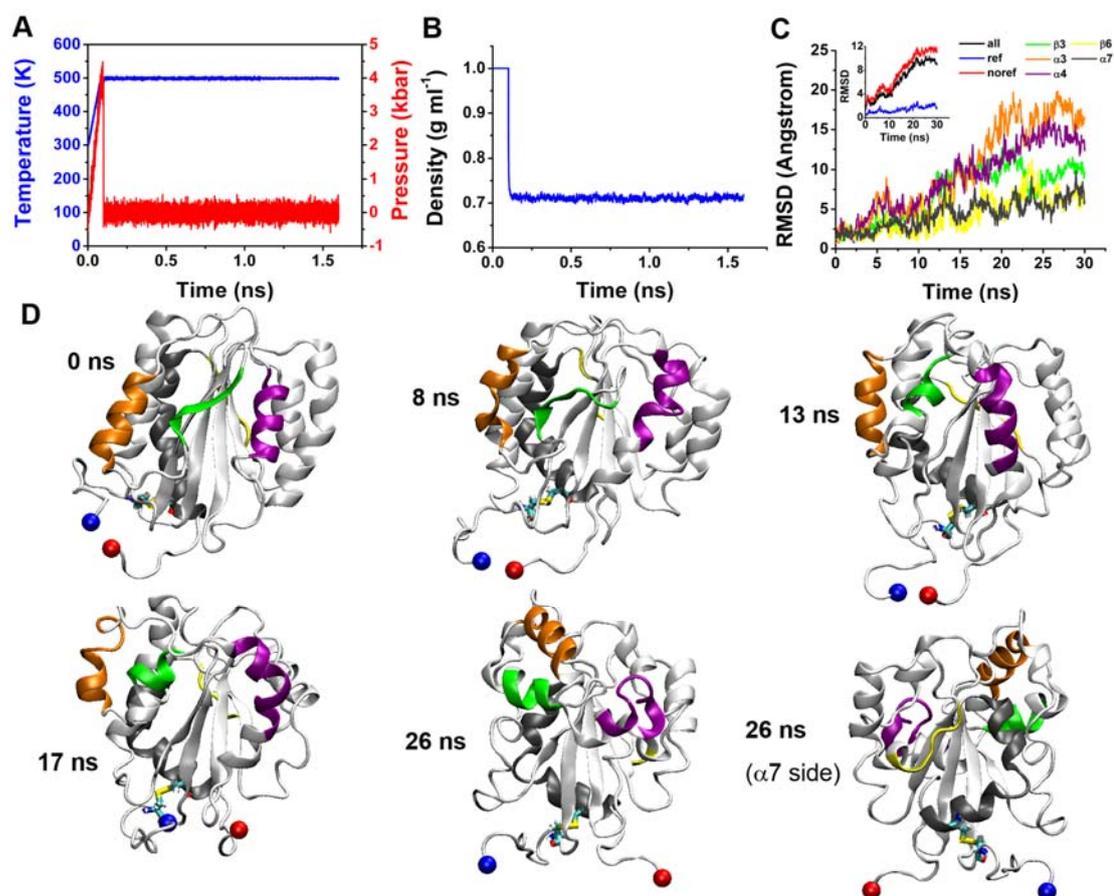
Figure 5.12 Thermal unfolding of the A1 domain at 500 K. A. The temperature of the system (blue lines) was heated up from 300 K to 500 K in 100 ps. Then the pressure (red lines) was adjusted to 1 atm (1.01325 bar). **B.** During the pressure adjustment, the density of the system decreased from ~1 g ml^{-1} to ~0.7 g ml^{-1}. **C.** RMSD of Cα atoms of β3, α3, α4, β6, and α7 using the starting structure as reference is plotted versus simulation time. Structures at different time were aligned to the starting structure via the four strands (β1, β2, β4, and β5). Inset include RMSDs of all Cα atoms (all)**,** Cα atoms of β1, β2, β4, and β5 used for alignment (ref), or Cα atoms of other than the four strands (noref). **D.** Snapshots of the structure of the A1 domain during the thermal unfolding simulation. The α3 helix is colored in orange, the β3 strand in green, the α4 helix in purple, the α7 helix in black, and the β6 strand in yellow. The time of each snapshot is indicated in the parenthesis. 0 ns is the moment when the restraints on the A1 domain were turned off. Representations of A1 refer to Fig. 2.7B.

52

Similar to the A2 domain, the temperature of the A1 system was increased to 500 K in 100 ps with constant volume while the pressure went up to 4 kilobar (Fig. 5.12A). By changing the system volume, the pressure was adjusted back to 1 atm while the density of the system decreased to ~0.7 g ml^{-1} (Fig. 5.12B). The RMSD of all Cα atoms of the A1 domain firstly jumped to 4 Å and reached a plateau until 10 ns (inset of Fig. 5.12C). In contrast to the A2 domain, the major changes occurred at the N-terminal portion of the A1 domain. The $\beta3$ strand, the $\alpha3$, and $\alpha4$ helices showed the largest RMSDs (Fig. 5.12C). During the first jump of the RMSD of all Cα atoms, the $\alpha3$ helix and the $\beta3$ strand started to unfold while the $\alpha4$ helix was elongated (Fig. 5.12D, 8ns). From 10 to 20 ns, the RMSD of all Cα atoms increased continuously. During this period, the $\beta3$ strand turned into a helix ($\beta3$ helix) (Fig. 5.12D, 13ns), and then the $\beta3$ helix moved up- and leftward (Fig. 5.12D, 17ns). Finally, the RMSD of all Cα atoms reached a plateau with 10 Å. Both the $\alpha3$ and $\beta3$ helices flipped upward while the $\alpha4$ helix was unfolded and also moved upward (Fig. 5.12D, 26 ns). Meanwhile, on the other side of the A1 domain, the $\beta6$ strand was converted to a helix ($\beta6$ helix) and the $\alpha7$ helix was divided into two helices (Fig. 5.12D, 26 ns ($\alpha7$ side)).

Comparing the thermal unfolding of the A2 and the A1 domains, it was found that the most significant variations occurred at different locations: the C-terminal portion in the A2 domain, but the N-terminal portion in the A1 domain. This reflected their sequence difference, which affects the structure packing and thus the stability.

CHAPTER 6

UNBENDING OF $\alpha_V\beta_3$ INTEGRIN UNDER FORCE

As mentioned in section 2.3, experimental data suggest that integrin undergoes large-scale conformational changes during its bi-directional, inside-out or outside-in, signaling processes. In this chapter, SMD simulations were applied to study the unbending of $\alpha_V\beta_3$ integrin from the bent to the extended conformation induced by a tensile force.

6.1 Homology Models of the Unresolved EGF1/2 Domains of $\alpha_V\beta_3$ Integrin

In the crystal structure of the full ectodomain of $\alpha_V\beta_3$ integrin (pdb 1u8c), the EGF1/2 domains at the knee of β subunit were unresolved. The unresolved portion includes ~90 residues, which is ~6% of the whole ectodomain. In order to get a complete ectodomain structure, MODELLER[65] was used to build the unresolved EGF1/2 domains. A crystal structure of β_2 subunit (pdb 2p28) [85], which includes Hybrid, PSI, and EGF1/2/3 domains, was used as the template. β_2 and β_3 subunits have ~40% sequence identity and ~55% similarity in the modeling region of PSI, and EGF1/2 domains (Fig. 6.1A), so β_2 crystal structure is a really good template for homology modeling of β_3. First, the Hybrid and EGF3 domains of the $\alpha_V\beta_3$ crystal structure (pdb 1u8c) were aligned with the Hybrid and EGF3 domains of the β_2 crystal structure, respectively (Fig. 6.1B). Fixing the Hybrid and EGF3 domains, the PSI and EGF1/2 domains were then built. The resulted model had a similar quality as the template according to DOPE scores [66] (Fig. 6.1C).

```
A   # Length: 158
    # Identity:      65/158 (41.1%)
    # Similarity:    86/158 (54.4%)
    # Gaps:          17/158 (10.8%)
    # Score: 317.5
    #
    #
    #========================================

    2P28  β2      1  QECTKFKVSSCRECIESGPGCTWC--QKLNFTGPGDPDSIRCDTRPQL   46
                     ..||...||||::|:...|.|.||  :.|....|      |||.:..|
    1U8C  β3      1  GPNICTTRGVSSCQQCLAVSPMCAWCSDEALPLGSP------RCDLKENL  44
                             PSI  ←─────→  EGF1/2
    2P28  β2     47  LMRGCAADDIMDPTSECRCRDQSRDRS-LC-HGKGFLECGICRCDTGYIG  94
                     |...||.:.|..|.||:|.|:.|:...|  .|  :|.|..|||:|||..|::|
    1U8C  β3     45  LKDNCAPESIEFPVSDCACQAQAEPNSHRCNNGNGTFECGVCRCGPGWLG  94

    2P28  β2     95  KNCECQTQG-RSSQELEGSCRKDNNSIICSGLGDCVCGQCLCHTSDVPGK  143
                     ..|||..:. |.||:.|  |.......:||..|:|:||||:||:||. ||
    1U8C  β3     95  SQCECSEEDYRPSQQDE--CSPREGQPVCSQRGECLCGQCVCHSSDF-GK  141

    2P28  β2    144  LIYGQYCE   151
                     |.|:|||
    1U8C  β3    142  -ITGKYCE   148
```

Figure 6.1 Homology models of the unresolved EGF1/2 domains. A. Sequence alignment of PSI, EGF1, and EGF2 domains between β_2 and β_3. Generated by "needle" [64] of EMBL-EBI. B. Structural alignment between template (pdb 2p28) and model. Left is before model building and right is after. C. DOPE scores of template and model structures.

Because the template of the β_2 crystal structure is in an extended conformation, the model of β_3 subunit ended up in an extended conformation. To fit the extended β_3 model in the bent $\alpha_V\beta_3$ crystal structure, a targeted MD simulation with implicit water was run to bend the extended model. First, the EGF3 domain was restrained and Cα atoms of the β-sheet of the Hybrid domain were used as the target. Gradually, the β_3 model bent over toward its target (Fig. 6.2A). Second, all the heavy atoms in the Hybrid, PSI, and EGF3 domains were used as the target. The model became closer to the target.

Figure 6.2 Targeted MD to fit the extended model in the bent crystal structure. A. Process of bending of β3 model (red). Crystal structure of $\alpha_V\beta_3$ (pdb 1u8c) was used as target (white). **B.** Final model in the bent crystal structure. **C.** RMSD of all heavy atoms of PSI and EGF1/2 domains from the targeted crystal structure.

Finally, force constant was increased to 1 kcal mol^{-1}, so the RMSD of the model from the target further decreased to ~3 Å (Fig. 6.2C). The final model did not conflict with α_V subunit in the crystal structure. It also kept similar DOPE score as initial model, indicating that targeted MD did not distort the model badly.

6.2 Equilibration of Bent $\alpha_V\beta_3$ Integrin in a Water Box

The complete structure of $\alpha_V\beta_3$ ectodomain, except for the unresolved Calf2 EE' loop, was soaked in a water box (Fig. 6.3A). The protein included 1,617 residues and the systems had ~250,000 atoms in total. After system energy minimization and heat-up to 300 K, a free MD simulation was run for ~40 ns. The RMSD from the initial structure of the protein reached a plateau after 5 ns. However, a small jump was seen at 15 ns. The
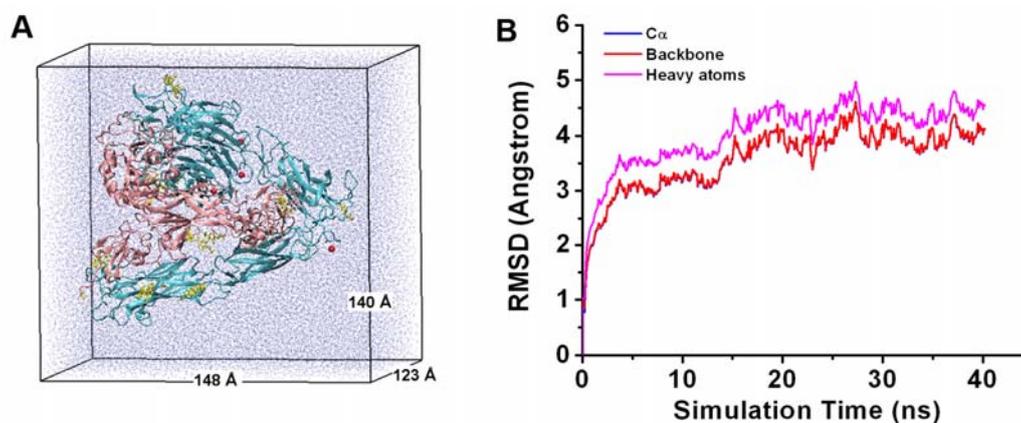
Figure 6.3 Equilibration of bent $\alpha_V\beta_3$ in a water box. A. The complete ectodomain of $\alpha_V\beta_3$ was soaked in a water box. α-subunit is in cyan, β-subunit is in pink, and water in blue. Red balls represent calcium ions and yellow sticks indicate sugar residues attached on the protein core. **B.** RMSD from the initial structure of Cα atoms, backbone, or heavy atoms during the equilibration process.

following analysis will show that it corresponds to the relaxation of the modeled part at the knee.

During the equilibration, extensive interactions existed between α- and β-subunits as shown in Fig. 6.4. The total buried SASA of each subunit was ~4,000 Å^2, half of which was between the βA and β-propeller domains. Besides, a protruding β hairpin of the β-propeller domain had a little contact (<100 Å^2 buried SASA) with the Hybrid domain of β-subunit. Similarly, among ~30 hydrogen bonds in total, half was between the βA and β-propeller domains. The contacts between the β hairpin of the β-propeller domain and the Hybrid domain were two salt bridges (Asp306/Lys384 and Lys308/Glu358) (Table 6.1). Around the knee region, Thigh and Calf1 of α-subunit and EGF2 and EGF3 of β-subunit contacted with each other. Initially, their contacts were decreased due to the relaxation of the built homology models. Then the buried SASAs became stable at 200-500 Å^2. About five hydrogen bonds formed for each domain. In the lower leg region, Calf2 of α-subunit had ~800 Å^2 contact area and ~7 hydrogen bonds

57
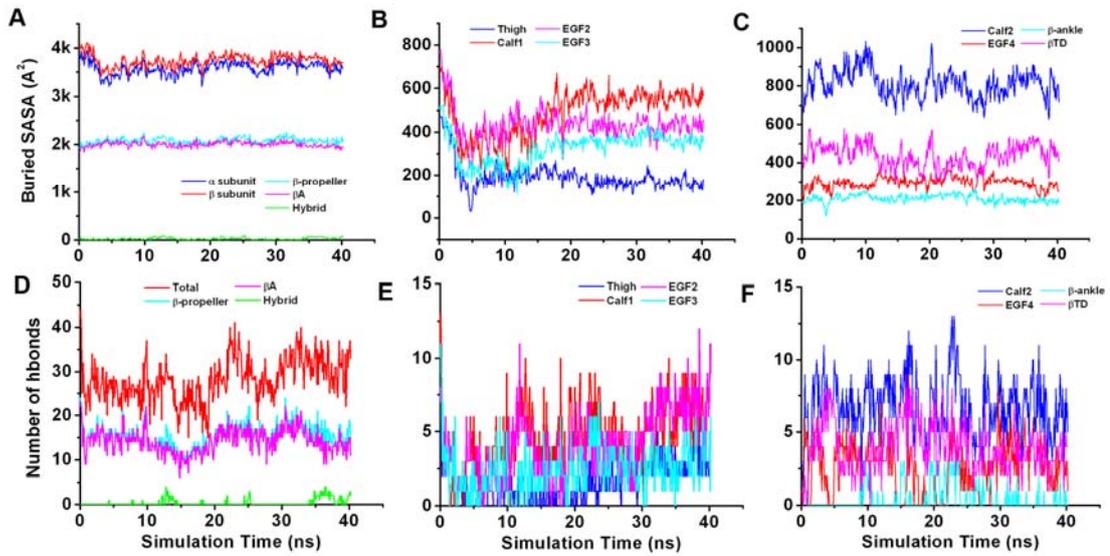
Figure 6.4 Interactions between α- and β-subunits of α$_V$β$_3$. Buried SASA in total and in the headpiece region (**A**), in the knee region (**B**), or in the lower legs (**C**). Hydrogen bonds formed in total and in the headpiece region (**D**), in the knee region (**E**), or in the lower legs (**F**).

Table 6.1
Major hydrogen bonds between α- and β-subunits of α$_V$β$_3$

α-subunit	β-subunit	Existing Time (ns)	α-subunit	β-subunit	Existing Time (ns)
*Asp306 O$_{\delta 1}$ (β-propeller)	*Lys384 N$_\zeta$ (Hybrid)	3.5	*Arg303 N$_{\eta 1}$ (β-propeller)	*Asp565 O$_{\delta 1}$ (EGF4)	7.8
*Asp306 O$_{\delta 2}$ (β-propeller)	*Lys384 N$_\zeta$ (Hybrid)	4.5	Arg783 N$_{\eta 1}$ (Calf2)	Ser593 O (EGF4)	13.8
*Lys308 N$_\zeta$ (β-propeller)	*Glu358 O$_{\varepsilon 1}$ (Hybrid)	2.5	Arg783 N$_{\eta 2}$ (Calf2)	Ser593 O (EGF4)	6.7
*Lys308 N$_\zeta$ (β-propeller)	*Glu358 O$_{\varepsilon 2}$ (Hybrid)	2.3	Arg783 N$_{\eta 1}$ (Calf2)	Ser589 O (EGF4)	9.2
Lys503 N$_\zeta$ (Thigh)	Ser510 O (EGF2)	21.2	Arg783 N$_{\eta 2}$ (Calf2)	Ser589 O (EGF4)	11
*Glu661 O$_{\varepsilon 1}$ (Calf1)	*Arg489 N$_{\eta 2}$ (EGF2)	17.6	Arg783 N$_{\eta 1}$ (Calf2)	Gln590 O (EGF4)	8.5
*Glu661 O$_{\varepsilon 2}$ (Calf1)	*Arg489 N$_{\eta 2}$ (EGF2)	12.9	Arg745 N$_\varepsilon$ (Calf2)	Thr603 O$_{\gamma 1}$ (β-ankle)	6.4
*Glu661 O$_{\varepsilon 2}$ (Calf1)	*Arg489 N$_{\eta 1}$ (EGF2)	9.6	Arg745 N$_{\eta 2}$ (Calf2)	Thr603 O (β-ankle)	4.5
*Glu661 O$_{\varepsilon 2}$ (Calf1)	*Arg489 N$_{\eta 1}$ (EGF2)	7.4	Arg745 N$_{\eta 2}$ (Calf2)	Thr603 O$_{\gamma 1}$ (β-ankle)	3.7
Arg665 N$_{\eta 1}$ (Calf1)	Glu522 O (EGF2)	9	*Glu769 O$_{\varepsilon 1}$ (Calf2)	*Lys650 N$_\varepsilon$ (βTD)	13.3
*Asp652 O$_{\delta 1}$ (Calf1)	*Lys532 N$_\zeta$ (EGF3)	15.8	*Glu769 O$_{\varepsilon 2}$ (Calf2)	*Lys650 N$_\varepsilon$ (βTD)	13.8
*Asp652 O$_{\delta 2}$ (Calf1)	*Lys532 N$_\zeta$ (EGF3)	18.9	*Glu770 O$_{\varepsilon 2}$ (Calf2)	*Lys650 N$_\varepsilon$ (βTD)	12.2

Arg698 $N_{\eta 1}$ (Calf1)	Tyr557 O_{η} (EGF3)	9.2	Ser749 O (Calf2)	Thr609 $O_{\gamma 1}$ (βTD)	11.1
Ser305 O_{γ} (β-propeller)	Trp553 $N_{\varepsilon 1}$ (EGF3)	10.8	Gly925 O (Calf2)	Lys658 N_{ε} (βTD)	10.3
			Ser749 O (Calf2)	Asp606 N (βTD)	9.9

Domains are indicated in parenthesis. Hydrogen bonds between β-propeller and βA are not included.
* Salt bridges.

with EGF4, β-ankle (a small loop between EGF4 and βTD constrained by a disulfide

bond), and βTD.

The simulated $\alpha_V \beta_3$ was kept in a bent conformation during the equilibration

because of strong interactions between the headpiece and tailpiece, especially for β-

subunit (Fig. 6.5). The total contact area between the headpiece and tailpiece of β-subunit

was ~1400 Å^2 while it was only ~500 Å^2 for α-subunit. Similarly, β-subunit had ~20

hydrogen bonds while α-subunit only 5. A sudden drop of buried SASA and hydrogen

bonds was observed for β-subunit at 15 ns because of the adjustment of the structure at

the knee region, which also caused the jump of RMSD as mentioned. The interactions
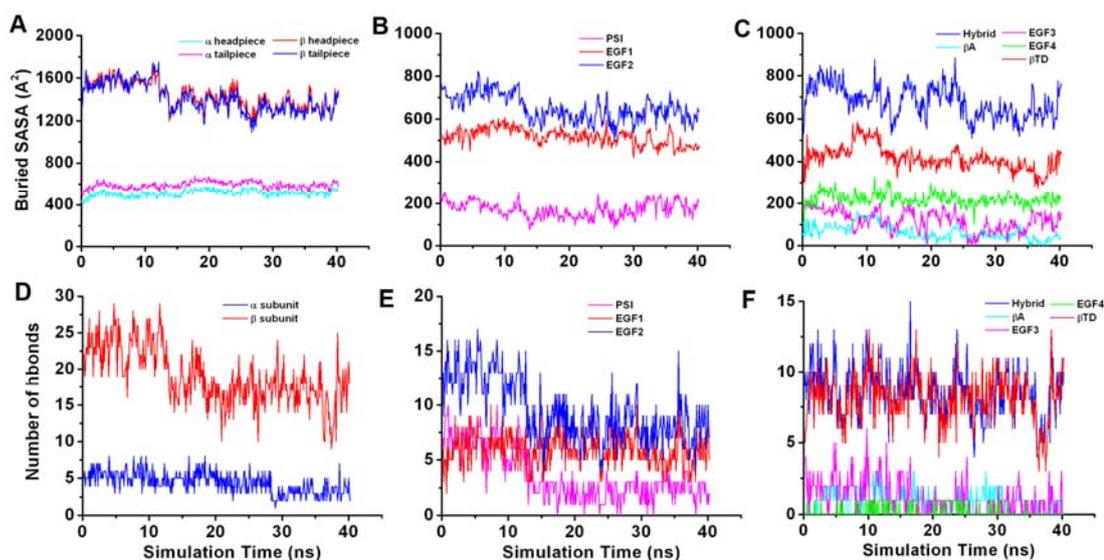
Figure 6.5 Interactions between the headpiece and tailpiece of $\alpha_V \beta_3$. Buried SASA in total (**A**), in the knee region (**B**), or in other regions (**C**). Hydrogen bonds formed in total (**D**), in the knee region (**E**), or in other regions (**F**).

Table 6.2
Major hydrogen bonds between the headpiece and tailpiece of $\alpha_V\beta_3$

Headpiece	Tailpiece	Existing Time (ns)	Headpiece	Tailpiece	Existing Time (ns)
His591 $N_{\varepsilon 2}$ (Thigh)	Glu640 $O_{\varepsilon 2}$ (Calf1)	36.1	*Asp346 $O_{\delta 1}$ (βA)	*Lys676 N_ζ (βTD)	15.3
Ile592 O (Thigh)	Tyr639 O_η (Calf1)	29.9	*Asp346 $O_{\delta 2}$ (βA)	*Lys676 N_ζ (βTD)	13
Gln589 $O_{\varepsilon 2}$ (Thigh)	Asn685 $N_{\delta 2}$ (Calf1)	1.91	Cys374 O (Hybrid)	Arg633 N_ε (βTD)	35.4
*Arg8 $N_{\eta 2}$ (PSI)	*Glu490 $O_{\varepsilon 1}$ (EGF2)	31.2	Cys374 O (Hybrid)	Arg633 $N_{\eta 2}$ (βTD)	15.5
*Arg8 N_ε (PSI)	*Glu490 $O_{\varepsilon 1}$ (EGF2)	22.5	Gly388 O (Hybrid)	Arg633 $N_{\eta 1}$ (βTD)	33.4
*Arg8 N_ε (PSI)	*Glu490 $O_{\varepsilon 2}$ (EGF2)	15.4	*Asp393 $O_{\delta 2}$ (Hybrid)	*Arg633 $N_{\eta 1}$ (βTD)	30.3
*Arg8 $N_{\eta 2}$ (PSI)	*Glu490 $O_{\varepsilon 2}$ (EGF2)	12.9	*Asp393 $O_{\delta 2}$ (Hybrid)	*Arg633 $N_{\eta 2}$ (βTD)	26.1
Gly465 N (EGF1)	Asp484 O (EGF2)	31.8	*Asp393 $O_{\delta 1}$ (Hybrid)	*Arg633 $N_{\eta 2}$ (βTD)	24.6
Asn452 $O_{\delta 1}$ (EGF1)	Gln482 $N_{\varepsilon 2}$ (EGF2)	30.2	*Asp393 $O_{\delta 1}$ (Hybrid)	*Arg633 $N_{\eta 1}$ (βTD)	12.6
Asn452 $N_{\delta 2}$ (EGF1)	Gln482 $O_{\varepsilon 1}$ (EGF2)	26	Asp393 $O_{\delta 1}$ (Hybrid)	Asn632 $N_{\delta 2}$ (βTD)	29.3
Asn450 $N_{\delta 2}$ (EGF1)	Cys473 O (EGF2)	26.4	Asp393 $O_{\delta 2}$ (Hybrid)	Asn632 $N_{\delta 2}$ (βTD)	22.7
Asn452 $N_{\delta 2}$ (EGF1)	Cys473 O (EGF2)	21.6			

Domains are indicated in parenthesis.
* Salt bridges.

between PSI and EGF2 were unfavorable initially due to errors in structure building.

After adjustment, the interactions became stable. Major contacts at knee were between

EGF1 and EGF2. However, a few contacts between PSI and EGF2 including a salt bridge

(Arg8/Glu490, Table 6.2) may be more important because all of them need to be broken

for extension while only part of contacts between EGF1 and EGF2 need to. Other than

the knee, the Hybrid domain contacted with EGF3, EGF4, and βTD. Most contacts were

between Hybrid and βTD, including a salt bridge (Asp393/Arg633) that might be

important to regulate extension. EGF4 had ~200 $Å^2$ contacts with Hybrid, most of which

were hydrophobic. EGF3 had much less contacts (~100 $Å^2$) with Hybrid. Interestingly, a

salt bridge (Asp346/Lys676) was formed between the α7 helix of βA and the CD loop of

βTD during equilibration, which might play a role in regulating the binding affinity of

βA as suggested in the "dead-bolt" model [51].

6.3 Unbending of $\alpha_V\beta_3$ Integrin by Force
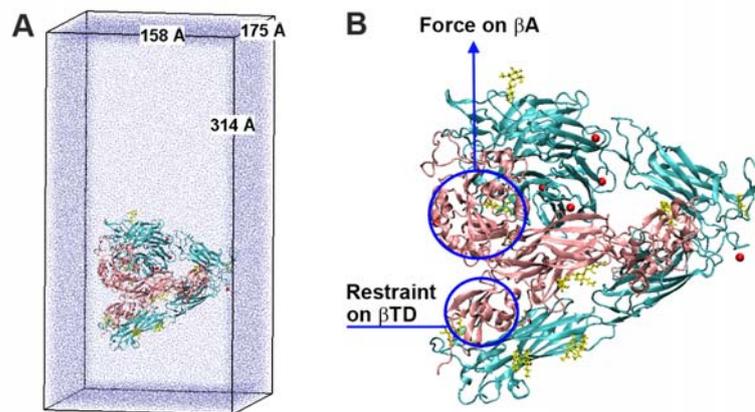
Figure 6.6 Set-up of SMD simulations. A. The water box was enlarged along the vertical direction to provide space for $\alpha_V\beta_3$ extension. Representations are the same as in Fig. 6.3. **B.** The pulling force is loaded on the COM of Cα atoms of the central β sheet of βA while the COM of Cα atoms of βTD, except for loops, was restrained by a harmonic potential.

To facilitate the transition from a bent to an extended conformation, SMD

simulations were performed. First, the previous water box used for equilibration was

enlarged along the vertical direction in order to provide space for $\alpha_V\beta_3$ integrin extension

(Fig. 6.6A), resulting in a system of ~850,000 atoms. In the following SMD simulations,

a force was loaded on the COM of Cα atoms of the central β sheet of βA along the

vertical direction, while the COM of Cα atoms of βTD, except for loops, was restrained

by a harmonic potential. In the SMD simulations, the springs attached on pulled atoms

moved at constant velocities, so-called constant-velocity SMD. In a fast pulling SMD

simulation at 2 nm ns^{-1} with a spring constant of 1 kcal mol^{-1} Å^{-2} (695 pN nm^{-1}), a big

force peak was observed at ~2 nm (Fig. 6.7A). In comparison, a slow pulling SMD
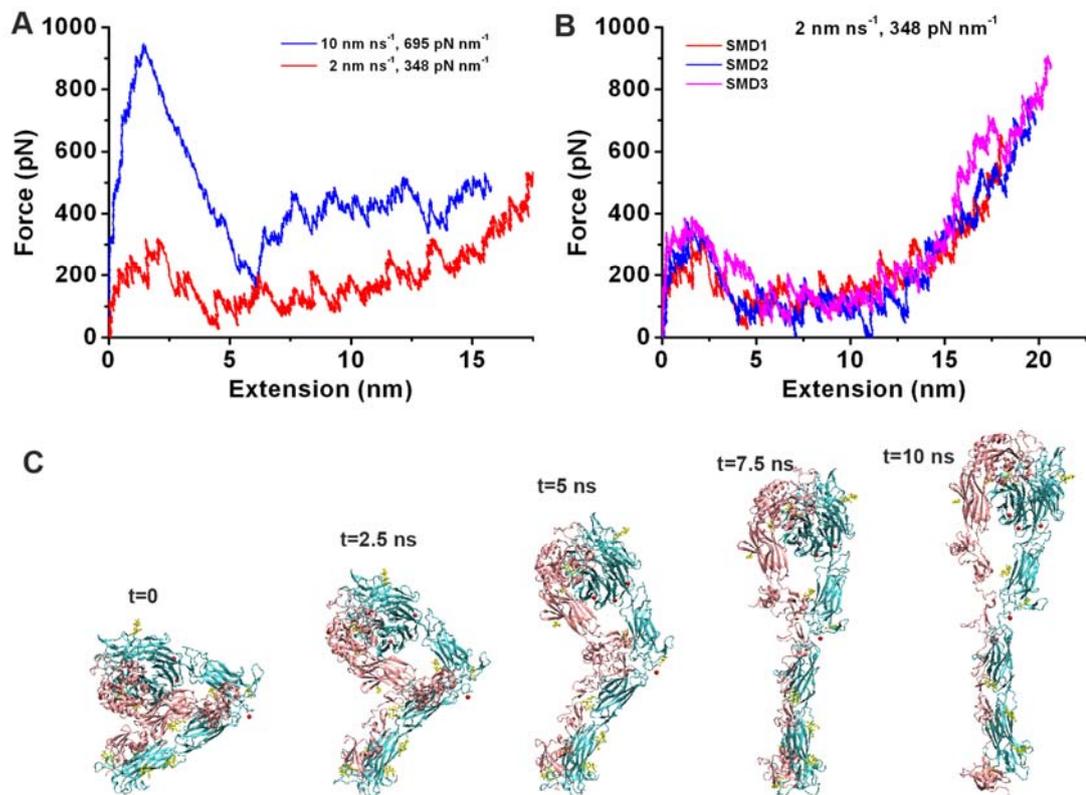
Figure 6.7 Constant-velocity SMD simulations of $\alpha_V\beta_3$ integrin unbending. A. Plot of force versus extension for fast pulling with hard spring or slow pulling with soft spring. **B.** Plot of force versus extension for three slowing pulling with soft spring starting from different equilibrated conformations. **C.** Snapshots at indicated time during the first slow pulling SMD simulation.

simulation was run at 2 nm ns^{-1} with a spring constant of 0.5 kcal mol^{-1} Å^{-2} (348 pN nm^{-1}).

A force peak occurred at the same location, but it is much smaller probably due to lower

viscosity resistance. After the first peak, the force dropped to near zero, followed by

multiple less pronounced peaks. During the pulling, simulated $\alpha_V\beta_3$ integrin was

gradually extended as shown in Fig. 6.7C. Because slow pulling would give more time

for the simulated structure to equilibrate so that the structure would be less distorted, two

more constant-velocity SMD simulations were run at 2 nm ns^{-1}. Overall, the three force-

extension curves overlapped quite well (Fig. 6.7B). A force peak of ~300 pN occurred at

~2 nm, and then force dropped to below 100 pN. After that, multiple small peaks were
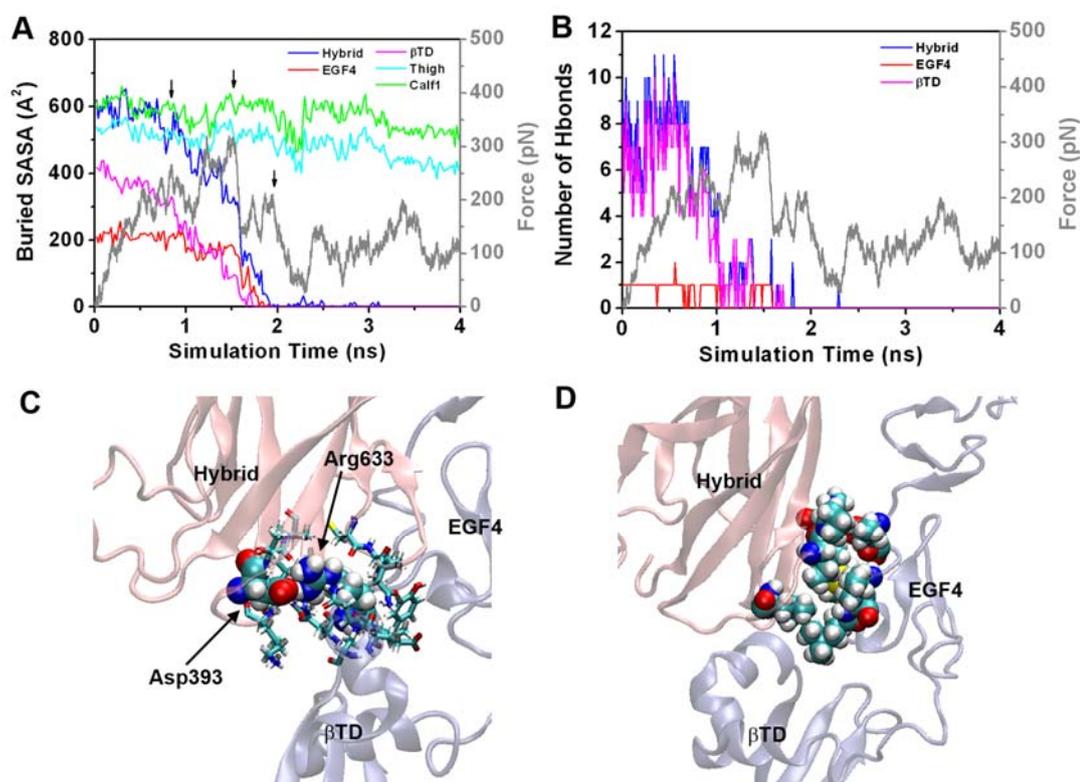
Figure 6.8 Critical interactions for $\alpha_V\beta_3$ integrin unbending. A. Changes of buried SASA (left y axis) compared with pulling force (right y axis) in the 1st slow pulling SMD. Arrows indicate force drops. **B.** Changes of number of hydrogen bonds compared with pulling force in the 1st slow pulling SMD. **C.** Interactions between Hybrid and βTD. The two residues forming a salt bridge are shown as spheres. Other surrounding residues are shown as sticks. **D.** Interactions between Hybrid and EGF4. Nitrogen atoms in blue, oxygen atoms in red, carbon atoms in cyan, and hydrogen atoms in white.

observed. Beyond 10 nm extension, the force increased as a worm-like chain model

because initial relaxed coiled structures were stretched after extension. During pulling,

inter-domain contacts, especially between EGF1-4 and βTD domains of β-subunit, were

broken. But overall, the simulated structure was not distorted badly. At end, an extended

$\alpha_V\beta_3$ integrin with closed headpiece and closed legs were obtained.

The first major force peak indicated a major energy barrier that need to be

overcome to extend $\alpha_V\beta_3$ integrin. If zoom in the first peak of the 1st slow pulling SMD,

three force drops were observed (Fig. 6.8A). At the first drop, contact area between
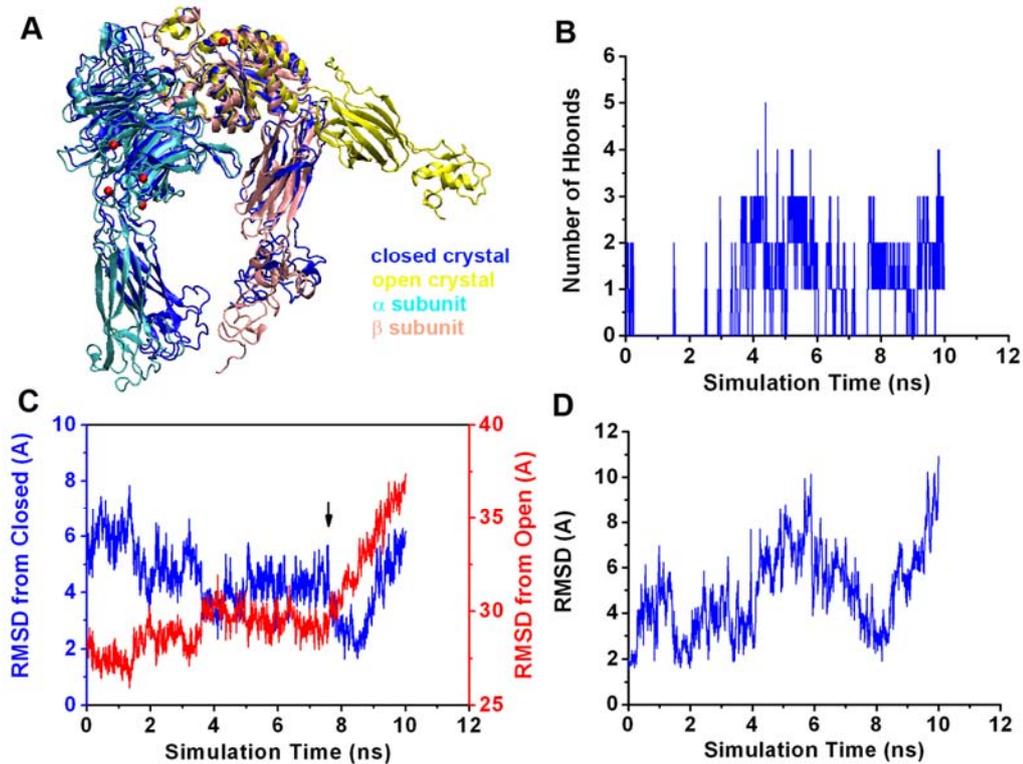
Figure 6.9 Headpiece conformations during 1st slow pulling SMD simulation. A. Comparison of simulated structure at 10 ns with closed and open crystal structure. All structures were aligned with Cα atoms of the central β sheet of the βA domain. **B.** Number of hydrogen bonds between the protruding β-hairpin of the β-propeller domain and Hybrid domain. **C.** RMSD of the Hybrid domain from the closed or open crystal structures. **D.** RMSD of the Thigh domain from closed the crystal structure after aligned with the β-propeller domain.

Hybrid and βTD decreased, indicating the breakage of their interactions. The second drop

corresponds to the disruption of interactions between Hybrid and EGF4, shown by the

decrease of their contact area. The third drop was mainly due to the breakage of

interactions between Thigh and Calf1 of α-subunit. Before the second drop, force

increased to the maximal, showing the strongest interactions to be broken during pulling.

These interactions were between Hybrid and EGF4/βTD. Interestingly, the number of

hydrogen bonds between Hybrid and βTD decreased from 8 to 0, but only one hydrogen

bond existed between Hybrid and βTD before interaction breakage (Fig. 6.8B). Indeed,

the interactions between Hybrid and βTD are mostly hydrophilic, where a salt bridge

(Asp393/Arg633) is surrounded by a few hydrogen bonds (Fig. 6.8C). In contrast, the interactions between Hybrid and EGF4 are mainly hydrophobic, involving Leu375, Ile380, and Leu383 on Hybrid, and Met568, Leu573, and Leu574 on EGF4 (Fig. 6.8D).

In the SMD simulations, the headpiece of $\alpha_V\beta_3$ integrin was still in a closed conformation. The two salt bridges between the β hairpin of β-propeller and hybrid were kept during the simulations (Fig. 6.9B), indicating a closed conformation. At ~8 ns of 1st slow pulling SMD, RMSD of Hybrid from the closed crystal structure decreased while RMSD from the open crystal structure increased (Fig. 6.9C), showing Hybrid further swung inward. In contrast, RMSD of Thigh increased because it swung out. The opposite movement of Hybrid and Thigh was because that β-subunit had longer length than α-subunit.

6.4 Free MD Simulations of Conformations along Unbending Pathway

During the unbending processes of $\alpha_V\beta_3$ integrin, different conformations were observed from partially to fully extended. Are these conformations stable? Are there any stable intermediate states? To answer these questions, free MD simulations were used to test the stabilities of different conformations. First, two partially extended structures were extracted from the two slow constant-velocity SMD simulations (Fig. 6.10). They were immediately after the first force peaks when the interactions between Hybrid and EGF4/βTD were just broken. Then, the loading force was turned off and the simulated integrin was allowed to move freely. Interestingly, the first partially extended structure (P1) from 1st SMD gradually bended back as indicated by decreasing RMSD from the equilibrated bent structure during the free MD simulation (Figs. 6.11 A and B). The RMSD went down to ~4 Å, showing that P1 structure almost returned to the bent
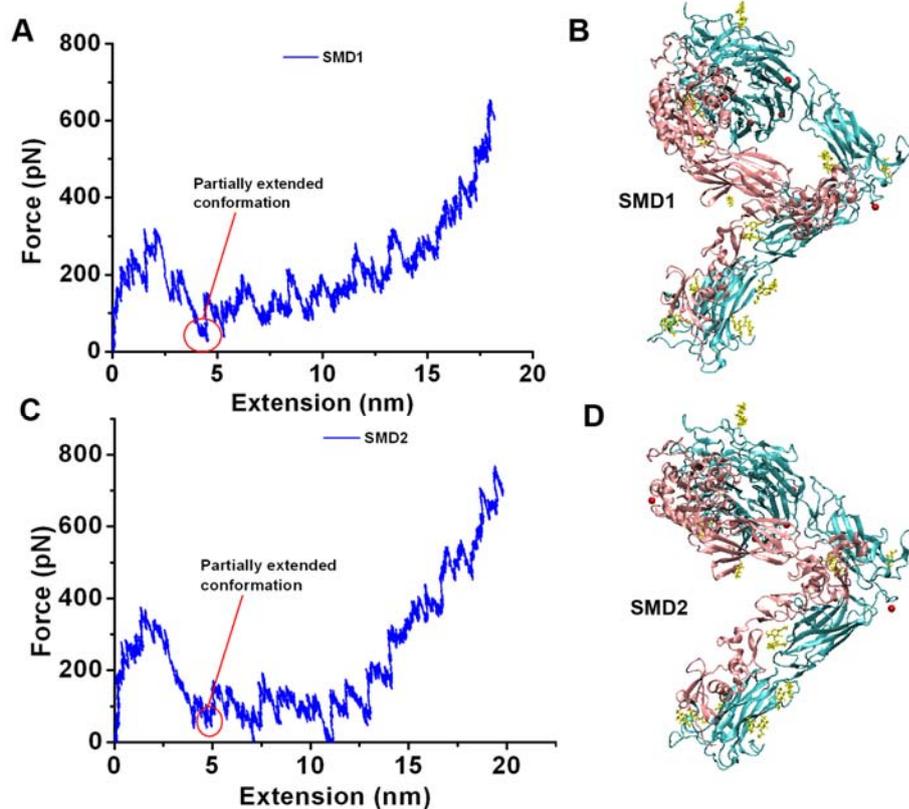
Figure 6.10 Partially extended structures for free MD simulations. Two structures were chosen right after the first force peak from 1st (**A** and **B**) and 2nd (**C** and **D**) slow constant-velocity SMD trajectories as indicated.

conformation. At residue level, part of interactions was restored. Initially, the two groups

of polar residues at Hybrid and βTD contacted with each other first, probably because

theses residues have long sidechains and the electrostatic force are long distance

interaction. But their contacts are not stable and frequently on and off. Compared to the

bent structure, the distance between the COMs of sidechains of the two polar groups was

still far away (> 10 Å) (Fig. 6.11C). Nevertheless, these interactions brought the

headpiece and the tailpiece of β3 subunit closer, which gave chances for interactions

between two groups of hydrophobic residues at Hybrid and EGF4. At ~7 ns, the two

hydrophobic groups were close enough for contacts. Strong hydrophobic interactions

brought them much closer (distance of COMs of sidechains ~5 Å). These interactions
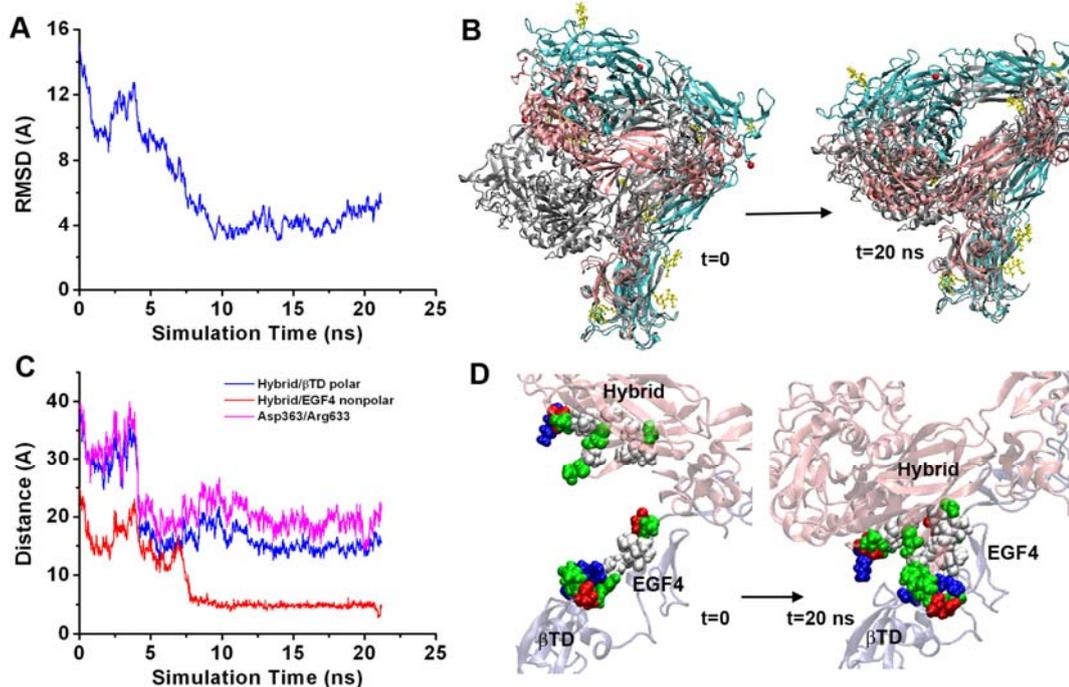
Figure 6.11 Free MD simulation of the P1 structure. A. RMSD of all Cα atoms from the equilibrated bent structure. **B.** Beginning and ending P1 structures compared with the equilibrated structure. Structures were aligned with EGF4, βTD, and Calf2. **C.** Distances between sidechain COMs of groups of interacting residues: polar residues at Hybrid and βTD, nonpolar residues at Hybrid and EGF4, and a salt bridge (Asp363/Arg633) between Hybrid and βTD. **D.** Beginning and ending structures at Hybrid/EGF4/ βTD interface. Basic residues are in blue, acidic residues in red, polar residues in green, and nonpolar in white.

were stable during the later 10 ns simulation and helped stabilize the bent conformation.

A similar free MD simulation was performed for the second partially extended structure (P2) from 2nd SMD. The headpiece of P2 also moved toward the tailpiece, indicated by decreasing RMSD from the equilibrated bent structure (Fig. 6.12A). However, the RMSD fluctuated between 6 and 8 Å, larger than that of P1, showing it was less close to the bent structure. Indeed, neither polar nor nonpolar interactions were restored as shown by large distances between COMs of sidechains (> 10 Å) (Fig. 6.12C). Although some polar residues at Hybrid and βTD were observed to contact frequently with each other, the distance between the two hydrophobic groups was still too far to get contacts. This was probably because a twist of the headpiece relative to the tailpiece,

67
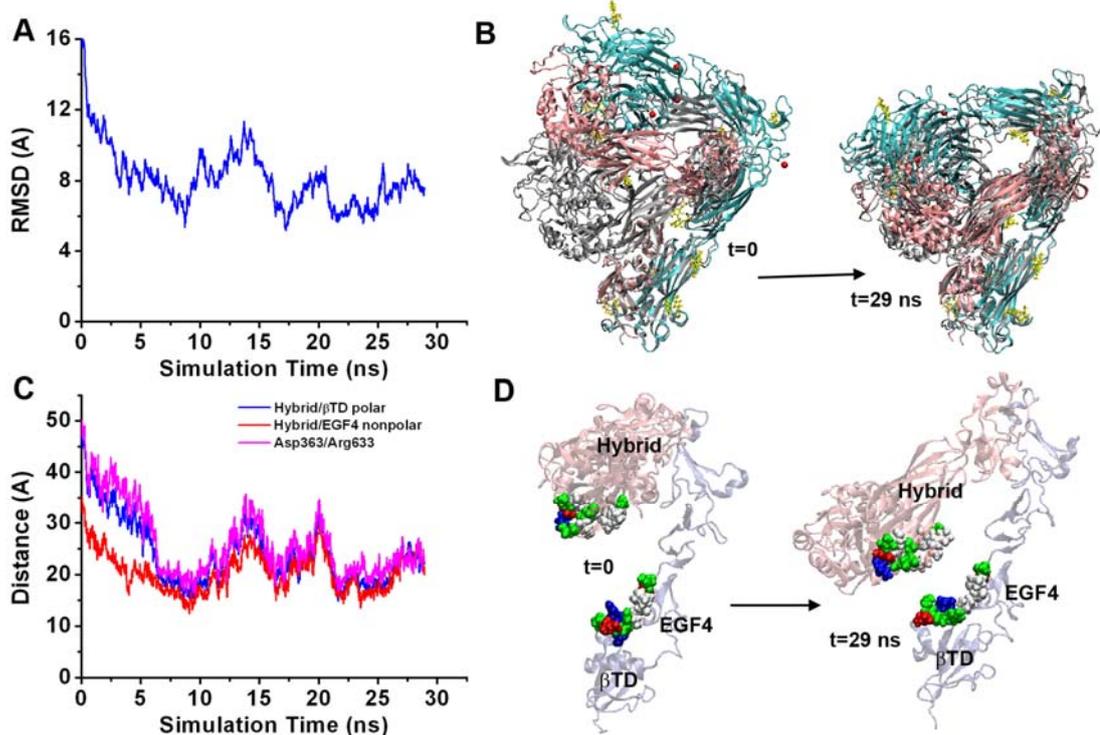
Figure 6.12 Free MD simulation of the P2 structure. A. RMSD of all Cα atoms from the equilibrated bent structure. **B.** Beginning and ending P1 structures compared with the equilibrated structure. Structures were aligned with EGF4, βTD, and Calf2. **C.** Distances between sidechain COMs of groups of interacting residues: polar residues at Hybrid and βTD, nonpolar residues at Hybrid and EGF4, and a salt bridge (Asp363/Arg633) between Hybrid and βTD. **D.** Beginning and ending structures at Hybrid/EGF4/ βTD interface. Basic residues are in blue, acidic residues in red, polar residues in green, and nonpolar in white.

induced by either different starting conditions or restraint on βTD that was not used for

P1.

According to the above free MD simulations, a partially extended conformation

tends to bend back because further extension costs more energy. Will a fully extended

conformation bend back or be stable? Two fully extended structures were extracted from

1st and 2nd slow constant-velocity SMD simulations as shown in Fig. 6.13. As mentioned

before, the fully extended structures still had a closed headpiece and closed legs. The first

fully extended structure (E1) from 1st SMD became relaxed after the loading force was

released. The RMSD from the initial structure first decreased and then reached a plateau
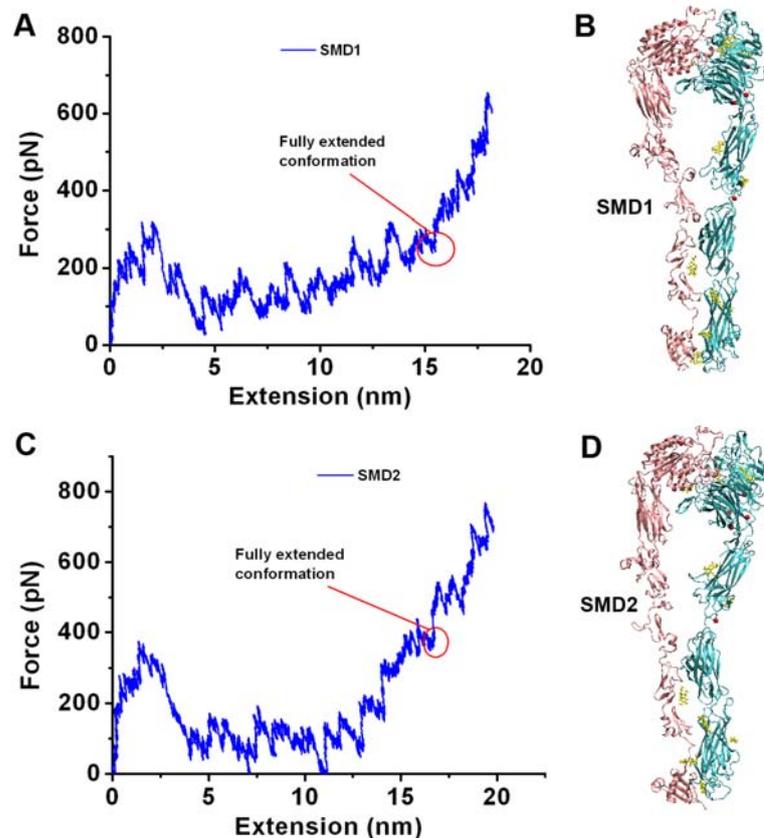
Figure 6.13 Fully extended structures for free MD simulations. Two fully extended structures were chosen at indicated time from 1st (**A** and **B**) and 2nd (**C** and **D**) slow constant-velocity SMD trajectories.

 (Fig. 6.14A), indicating the conformation was stable in the 20 ns free MD simulation. At the end of the simulation, E1 was still extended but curved (Fig. 6.14A). What are interactions that contributed to stabilize the extended conformation? Some new interactions at α-genu, which did not exist in the bent structure, were formed for the extended structure. Asp457 of Thigh became close to and coordinated with Ca^{2+} at α-genu (Fig. 6.14D). This coordination was very stable and kept in the last 15 ns, indicated by a small distance (< 4 Å) between O_δ atoms of Asp457 and Ca^{2+} (Fig. 6.14C). Because the Ca^{2+} also coordinated with Glu636 of Calf1 domain, Asp457-Ca^{2+}-Glu636 formed a bridge between Thigh and Calf1 to stabilize the extended conformation. Besides, salt bridges between Lys688 of Calf1 and Glu545 of Thigh or between Lys688 of Calf1 and

69
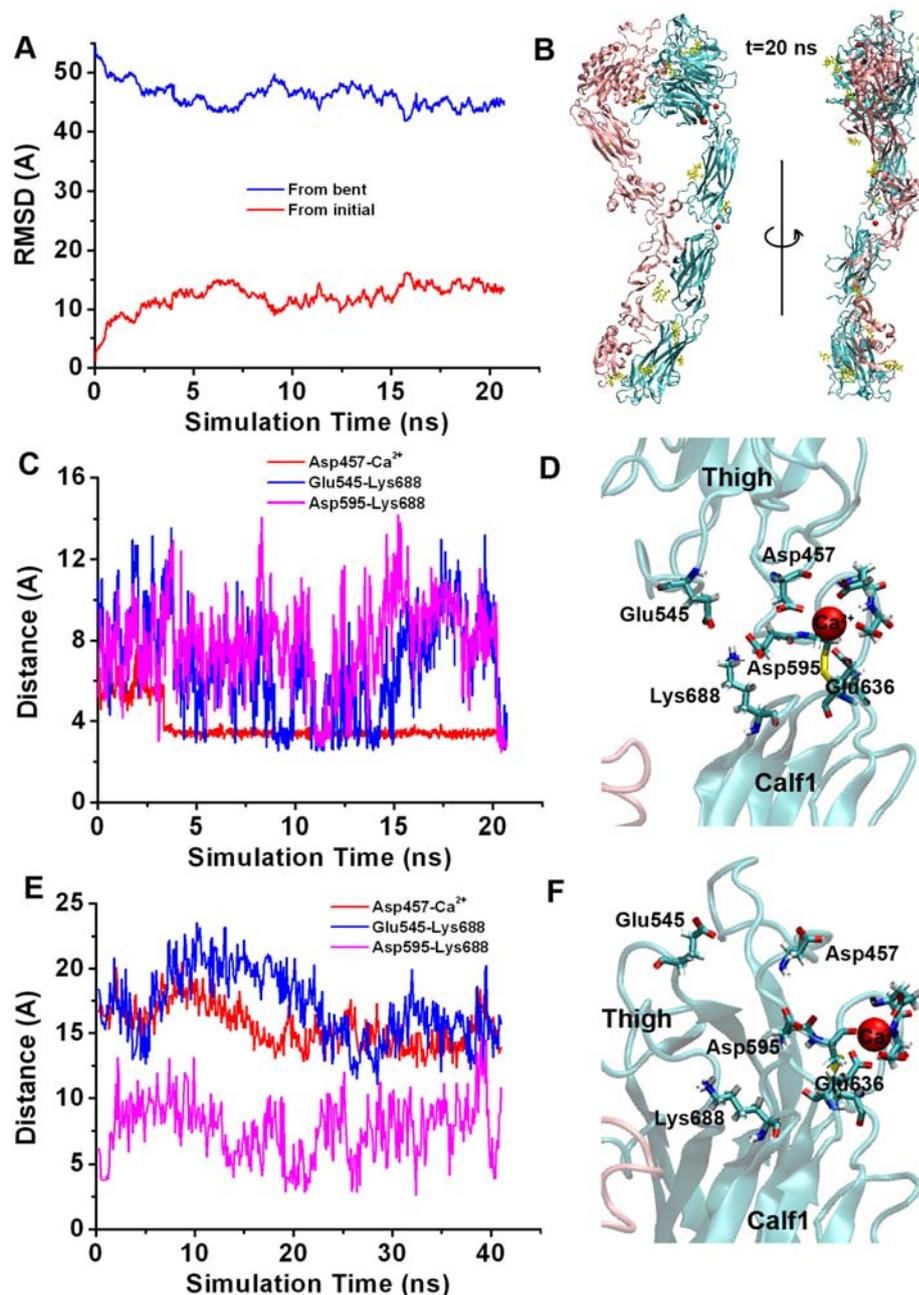
Figure 6.14 Free MD simulation of the E1 structure. A. RMSD of all Cα atoms from the equilibrated bent structure or the initial extended structure. **B.** Ending E1 structure. Right one was rotated 90 degree around the vertical axis. **C.** Distances between three pairs of residues during free MD of E1: O$_\delta$ atoms of Asp457 and Ca^{2+} at α-genu, O$_\varepsilon$ atoms of Glu545 and N$_\zeta$ atom of Lys688, and O$_\delta$ atoms of Asp595 and N$_\zeta$ atom of Lys688. **D.** Ending structure of free MD of E1 at Thigh/Calf1 interface. Nitrogen atoms in blue, oxygen atoms in red, carbon atoms in cyan, and hydrogen atoms in white. **E.** Distances between three pairs of residues during equilibration of the bent $\alpha_V\beta_3$. **F.** Ending structure of equilibration of the bent $\alpha_V\beta_3$ at Thigh/Calf1 interface.

70

Glu595 at the linker were formed, although they were not so stable and fluctuated. The salt bridges might also stabilize the extended conformation. These new interactions did not exist in the equilibrated bent structure (Fig. 6.14F). Indeed, Asp457 was quite far away from Ca^{2+} (> 10 Å) (Fig. 6.14E). The distance between O_ε atoms of Glu545 and N_ζ atom of Lys688 was also too large to form salt bridge during equilibration of the bent structure. Only Asp595 sometimes got close to Lys688 to form a salt bridge, but much less frequently.
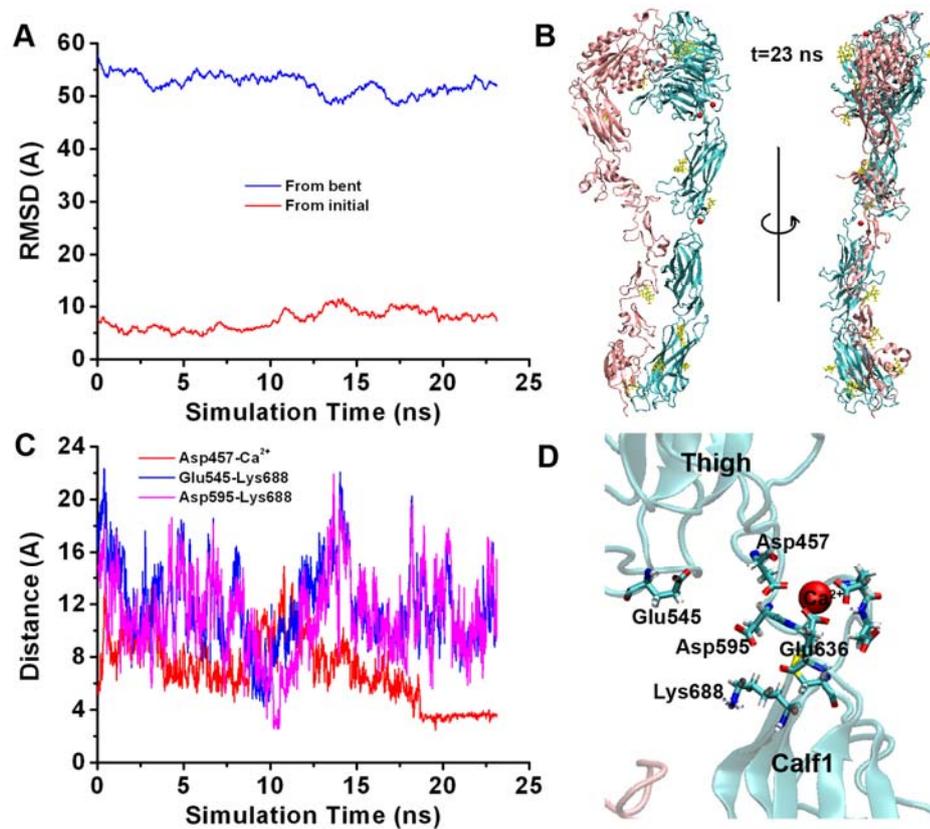
Figure 6.15 Free MD simulation of the E2 structure. A. RMSD of all Cα atoms from the equilibrated bent structure or the initial extended structure. **B.** Ending E2 structure. Right one was rotated 90 degree around the vertical axis. **C.** Distances between three pairs of residues during free MD of E2: O_δ atoms of Asp457 and Ca^{2+} at α-genu, O_ε atoms of Glu545 and N_ζ atom of Lys688, and O_δ atoms of Asp595 and N_ζ atom of Lys688. **D.** Ending structure of free MD of E1 at Thigh/Calf1 interface. Nitrogen atoms in blue, oxygen atoms in red, carbon atoms in cyan, and hydrogen atoms in white.

Similarly, a free MD was performed for the second fully extended structure (E2) from 2nd SMD. E2 was also stable and kept in an extended conformation in 20 ns. At end, E2 was extended but curved (Fig. 6.15B). Differently, Asp457 initially fluctuated around and did not have stable coordination with Ca^{2+} at α-genu. Until 20 ns, the distance between O$_\delta$ atoms of Asp457 and Ca^{2+} decreased to below 4 Å and then the coordination became stable (Fig. 6.15C). The interactions between Lys688 and Glu545/Asp595 were seldom observed during the simulation. The difference between free MD simulation of E1 and E2 demonstrated the effects of initial conditions. Nevertheless, both simulations showed the probably important role of Asp457 in stabilizing the extended conformation.

6.5 Separation of the Lower Legs of α$_V$β$_3$ Integrin by Force

In the switchblade model [50], it was suggested that separation of lower legs of integrins would induce integrin extension and further swing-out of the Hybrid domain to bring the βA domain into a high affinity state for ligands during an inside-out signaling. To test this model, the lower legs of simulated α$_V$β$_3$ integrin were separated by a force in constant-velocity SMD simulations. A force was loaded on the COM of Cα atoms of βTD, except for loops, and pointed along the horizontal direction (Fig. 6.16A). At the same time, the COM of Cα atoms of β-sheet of Calf2 was restrained by a harmonic potential. Two SMD simulations were run with different pulling speed and spring constant. A big force peak was observed at ~2 nm extension in both simulations (Fig. 6.16B). Compared to the unbending SMD in section 6.3, the peak force of leg separation was almost two folds of that of unbending at the same pulling speed of 2 nm ns^{-1}, which indicated that leg separation was much harder than unbending. After the big force peak, two lower legs of the simulated integrin gradually separated as shown in Fig. 6.16C. At
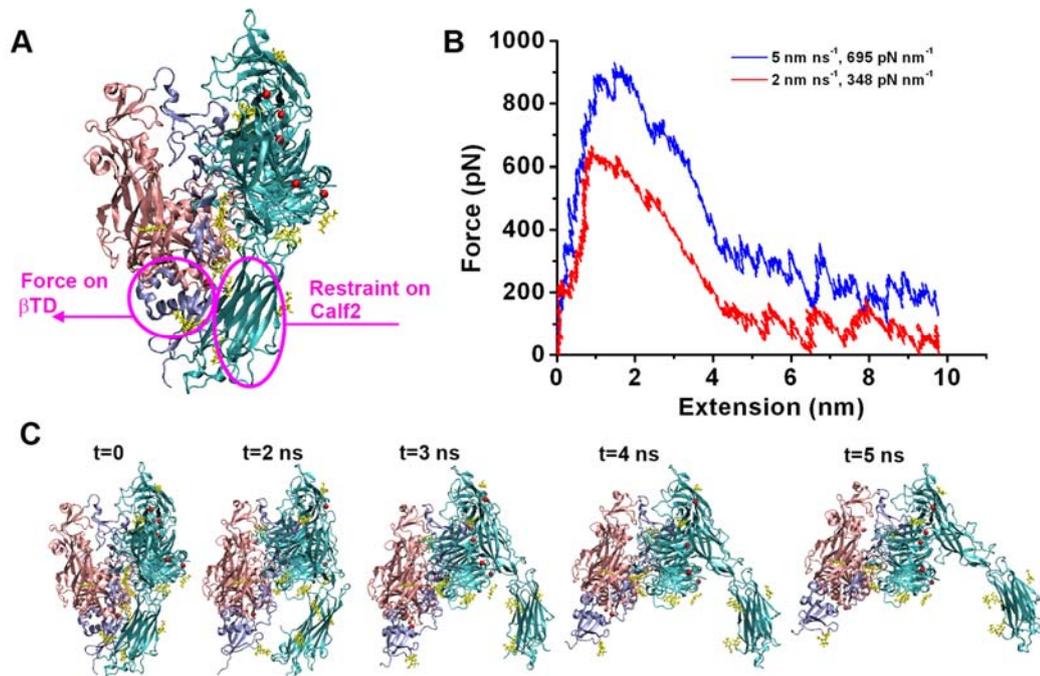
Figure 6.16 SMD simulations for separation of lower legs of $\alpha_V\beta_3$ integrin. A. A force was loaded on βTD while Calf2 was restrained. α-subunit in cyan, headpiece of β-subunit in pink, and tailpiece of β-subunit in iceblue. **B.** Force vs. extension curves for two SMD simulations with different pulling velocity and spring constant. **C.** Snapshots of simulated structures at indicated time during the simulation with 2 nm ns^{-1} pulling velocity.

end, the lower legs were fully separated to an extension of ~10 nm, but part of

interactions between α- and β-subunit at the genu and most of interactions between the

headpiece and the tailpiece of β-subunit still existed.

The big force peak corresponded to the disruption of interactions between

domains of lower legs as indicated by sudden drops of their contact areas (Figs. 6.17 A

and B). Interactions between Calf2 and EGF4/β-ankle/βTD were first broken so that

force started to drop. But interactions between Calf2 and EGF3 were still held and kept

the force at a high level. Until these interactions are broken, the force decreased greatly.

In contrast, Thigh/EGF2 contact area initially increased when the force dropped. When

the legs separated quite far away, the Thigh/EGF2 contact area had a sudden drop. The
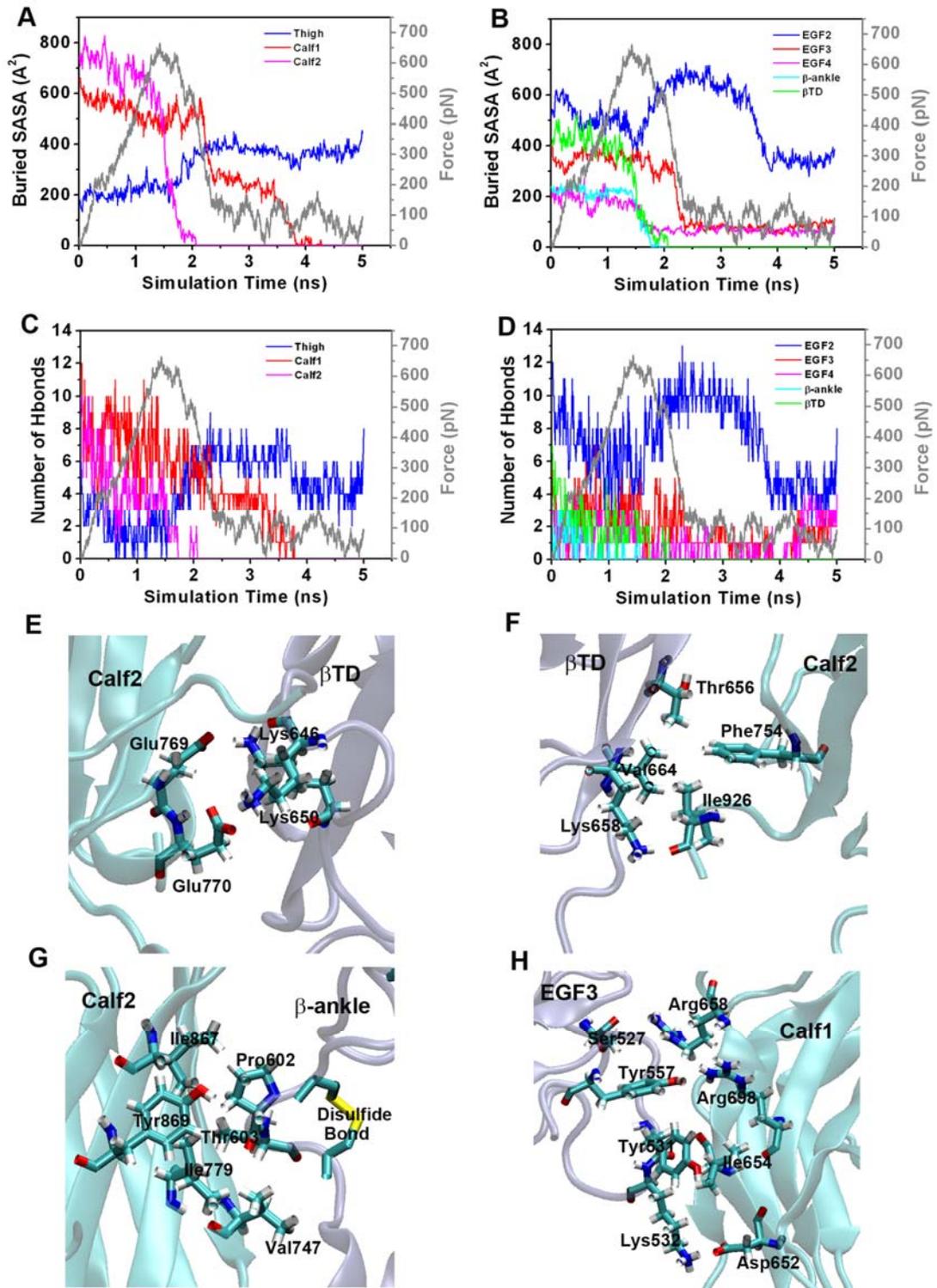
Figure 6.17 Changes of interactions during leg separation with a pulling speed of 2 nm ns⁻¹.
Buried SASAs between α- and β-subunits for each domains of α-subunit (**A**) or β-subunit (**B**) were plotted versus simulation time. Number of hydrogen bonds that formed between α- and β-subunits for

each domains of α-subunit (**C**) or β-subunit (**D**) were plotted versus simulation time. Structures of critical salt bridges (**E**) or hydrophobic interactions (**F**) between Calf2 and βTD. **G.** Hydrophobic interactions between Calf2 and β-ankle. **H.** Interactions between Calf1 and EGF3.

numbers of hydrogen bonds that formed between α- and β-subunits changed similarly as buried SASAs (Figs. 6.17 C and D). However, the changes of hydrogen bond number were not as big as expected, showing a large portion of interactions was probably hydrophobic. Indeed, both hydrophilic and hydrophobic interactions were found between the lower legs. Two salt bridges (Glu770/Lys650 and Glu769/Lys646) existed between Calf2 and βTD (Fig. 6.17E). At the same time, several hydrophobic residues of Calf2 and βTD also contacted with each other (Fig. 6.17F). Similarly, Pro602 and Thr603 of β-ankle inserted into a hydrophobic core formed by several hydrophobic residues of Calf2 (Fig. 6.17G). Between Calf1 and EGF3, salt bridges, hydrogen bonds, and hydrophobic interactions all existed (Fig. 6.17H). Compared to the interactions between the headpiece and the tailpiece, the interactions between the lower legs are much more extensive, therefore leg separation required higher force (or more energy) than unbending.

Before simulations of unbending of integrins with separated legs, the final structure obtained from the separation SMD simulation was equilibrated for 10 ns in order to reduce distortion due to nonequilibrium force pulling. The RMSD of all Cα atoms from the initial structure increased and then reach a plateau (Fig. 6.18A), indicating the system arrived in equilibrium. During equilibration, Thigh/Calf1 contacts decreased (Fig. 6.18D), but hydrogen bonds between them did not decrease (Fig. 6.18B), showing most of lost contacts were hydrophobic. Interactions between PSI and EGF2 increased (Fig. 6.18E), the same did contacts between Hybrid and EGF3/EGF4 (Fig. 6.18F). The changes of number of hydrogen bonds were small, showing at least part of
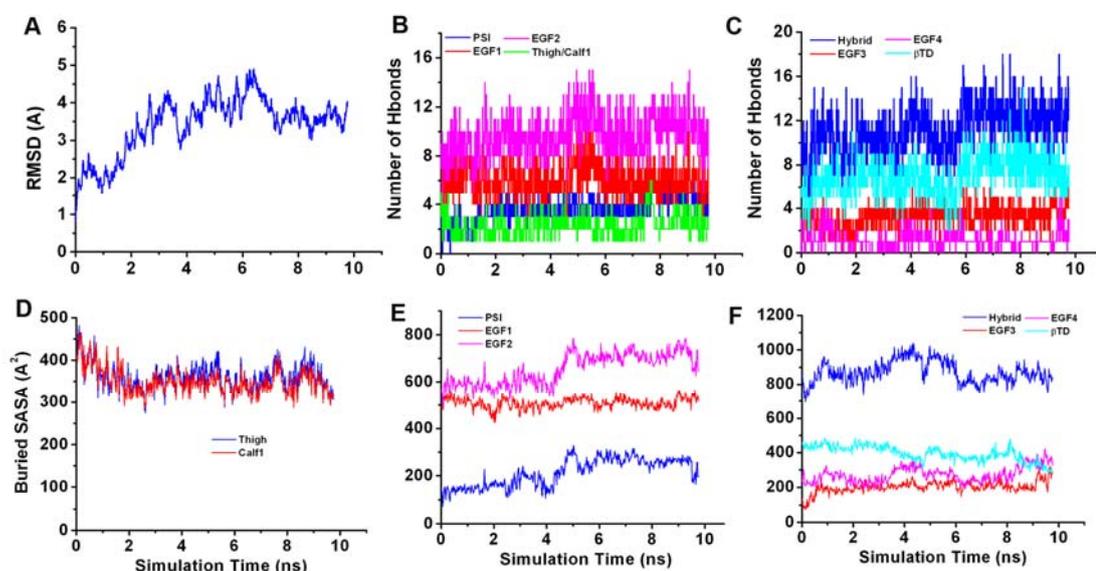
Figure 6.18 Equilibration of the bent $\alpha_V\beta_3$ integrin with separated lower legs. A. RMSD of all Cα atoms from the initial structure. Number of hydrogen bonds were plotted versus equilibration time for the knee region (**B**) or other regions (**C**). Buried SASAs between the headpiece and the tailpiece were plotted versus equilibration time for domains in α-subunit (**D**), the knee of β-subunit (**E**), or the head and the leg of β-subunit (**F**).

changes of contacts was hydrophobic.

6.6 Unbending of $\alpha_V\beta_3$ Integrin with Separated Lower Legs

Starting from different equilibrated structures of $\alpha_V\beta_3$ Integrin with separated lower legs, three SMD simulations (SMD4-6) were performed. A force was loaded on Cα atoms of the βA domain while Cα atoms of βTD and Calf2, except for loops, were restrained (Fig. 6.19A). The attached spring moved along the vertical direction at a speed of 2 nm ns^{-1} with a force constant of 0.5 kcal mol^{-1} A^{-2}. The force-extension curves of the three SMD simulations overlapped well (Fig. 6.19B). Similar to previous unbending of $\alpha_V\beta_3$ integrin with closed legs, a big force peak occurred at ~2 nm. Differently, two other peaks were observed, indicating more resistance during unbending. As shown in Fig. 6.19C, $\alpha_V\beta_3$ integrin gradually stood up. The interactions at the knee were broken, but the Hybrid domain still kept closed.
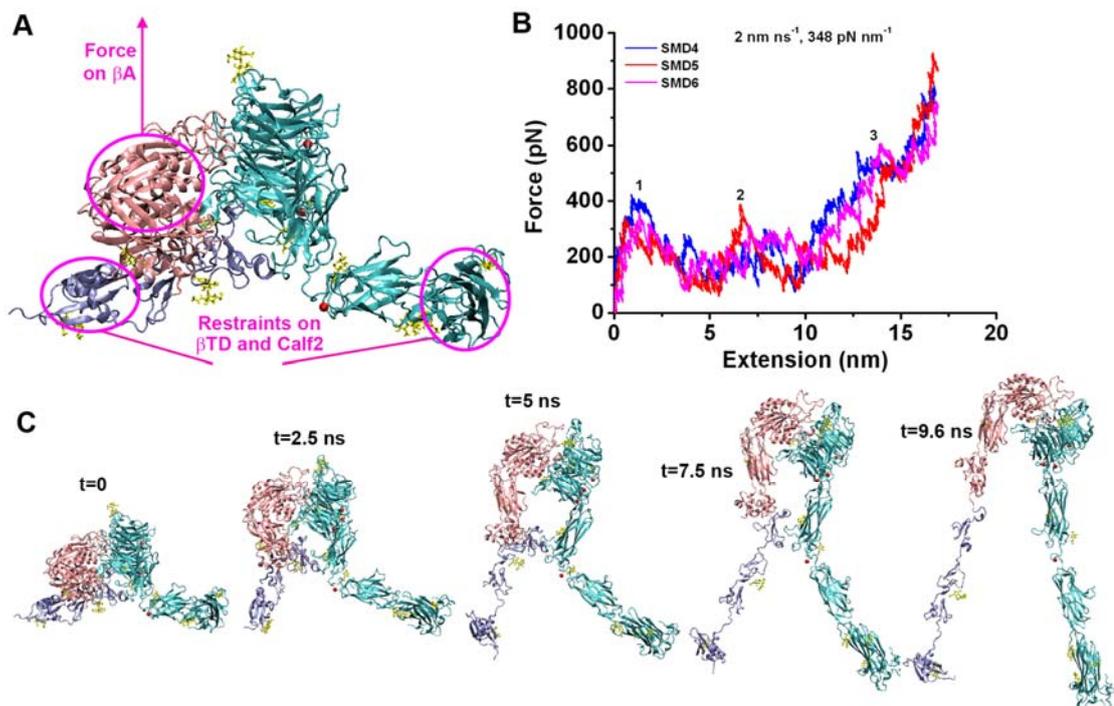
Figure 6.19 SMD simulations of unbending of $\alpha_V\beta_3$ integrin with separated lower legs. A. Set-up for SMD simulations. Force was loaded on βA while restraints were on βTD and Calf2. **B.** Force vs. extension curves for three SMD simulations (SMD4-6) with a pulling speed of 2 nm ns^{-1}. Three major force peaks were labeled with numbers. **C.** Snapshots at indicated time during SMD4.

Comparison between force and contact areas showed that each force peak corresponded to a breakage of contacts between domains of β-subunit (Fig. 6.20A). The breakage of interactions at Hybrid/βTD/EGF4 interfaces caused the first force drop. Then, the loss of contacts at Hybrid/EGF3 interface resulted in the second force drop. After that, EGF3 formed some interactions with PSI, indicated by increased buried SASA. Finally, the disruption of contacts at PSI/EGF1/EGF2 interfaces induced the third force drop. Part of broken interactions included hydrogen bonding as shown by drops of number of hydrogen bonds (Fig. 6.20B), except for EGF4/Hybrid contacts, which are mainly hydrophobic. The last two force peaks were not seen in the unbending with closed legs but pronounced with separated legs because the leg separation released constraints on the lower leg of β-subunit and thus resulted in more contacts between the headpiece and the

77
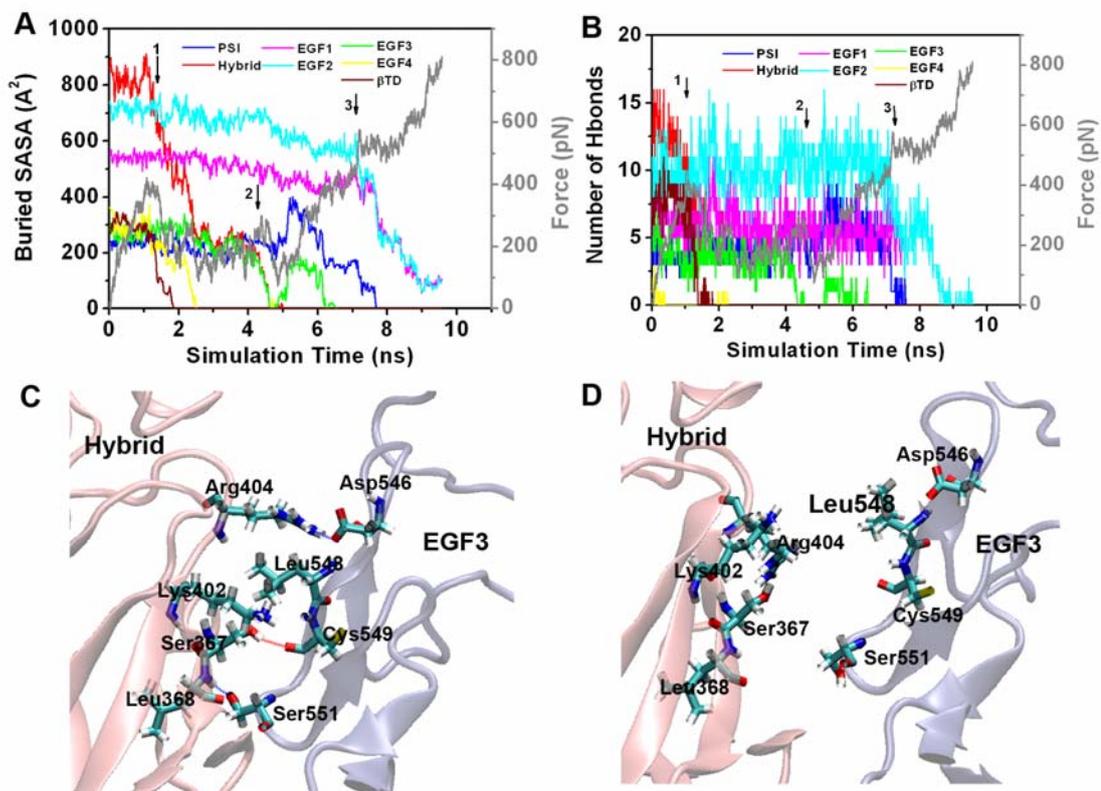
Figure 6.20 Critical interactions for $\alpha_V\beta_3$ integrin unbending with separated lower legs. A. Buried SASAs between the headpiece and the tailpiece of β-subunit for each domain. Arrows indicate sudden drops of contact areas. **B.** Changes of numbers of hydrogen bonds between the headpiece and the tailpiece of β-subunit for each domain. Arrows indicate sudden drops. Interacting residues between Hybrid and EGF3 of $\alpha_V\beta_3$ integrin with separated legs (**C**), which did not contact with each other in $\alpha_V\beta_3$ integrin with closed legs (**D**).

tailpiece of β-subunit. For example, several new interactions between Hybrid and EGF3, including a salt bridge, hydrogen bonds, and hydrophobic interactions, were formed when the legs were separated, but did not exist when the legs were closed (Figs. 6.20 C and D).

As mentioned above, the Hybrid domain did not swing out although the interactions at the knee were broken (Fig. 6.21A). The hydrogen bonding between the protruding β-hairpin of the β-propeller domain and Hybrid domain existed during the pulling (Fig. 6.21B), indicating that Hybrid kept closed. The RMSD of Hybrid from the closed or the open crystal structure, aligned with the central β-sheet of βA, did not change much, which also showed Hybrid was still closed (Fig. 6.21C). However, Hybrid
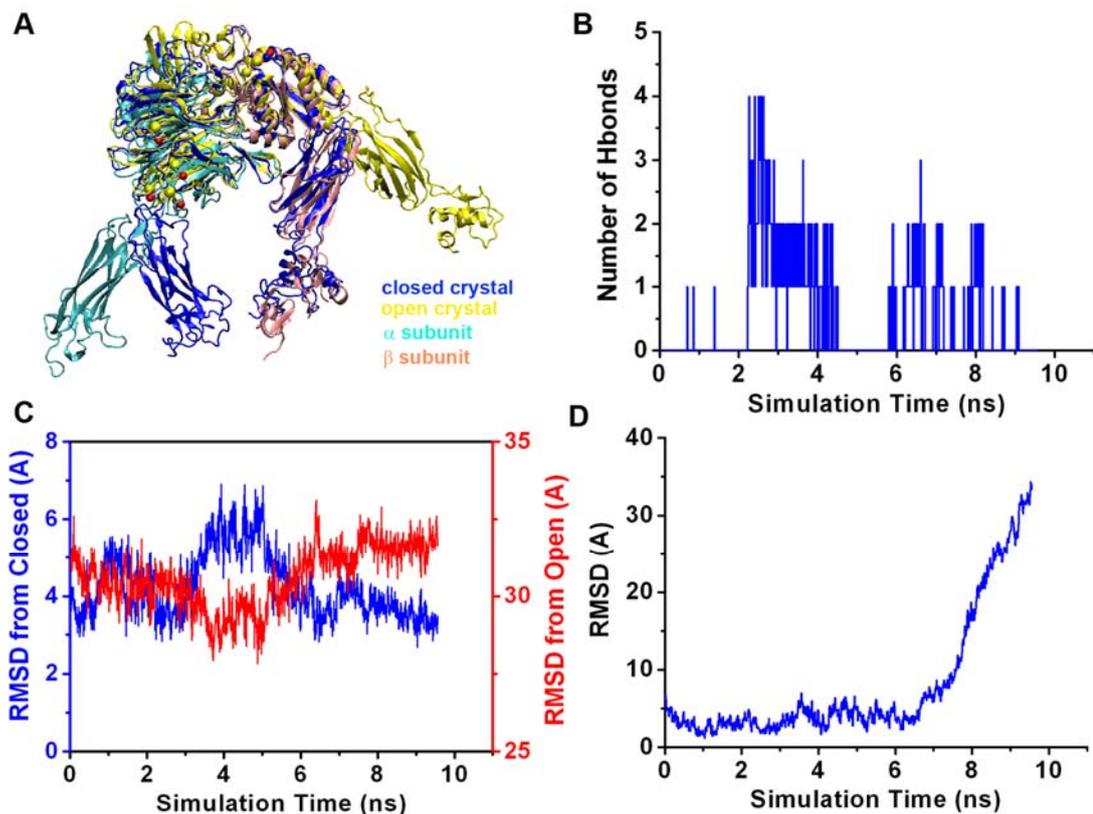
Figure 6.21 Headpiece conformations during the SMD4 simulation with separated lower legs. A. Comparison of simulated structure at 10 ns with closed and open crystal structure. All structures were aligned with Cα atoms of the central β-sheet of the βA domain. **B.** Number of hydrogen bonds between the protruding β-hairpin of the β-propeller domain and Hybrid domain. **C.** RMSD of the Hybrid domain from the closed or open crystal structures. **D.** RMSD of the Thigh domain from closed the crystal structure after aligned with the β-propeller domain.

did not swing further in as seen in previous SMD simulations with closed legs probably

because leg separation indeed changed force exerted on two subunits. This was also

demonstrated by much larger swing-out of the Thigh domain (Fig. 6.21D).

6.7 Free MD Simulations of Conformations with Separated Legs

Here, different conformations of $\alpha_V\beta_3$ integrin with separated legs along its

unbending pathway were tested for their stabilities. First, a partially extended structure

(P3) immediately after the first force drop was used for a free MD simulation (Fig.

6.22B). In the starting structure of P3, contacts between Hybrid and EGF4/βTD were
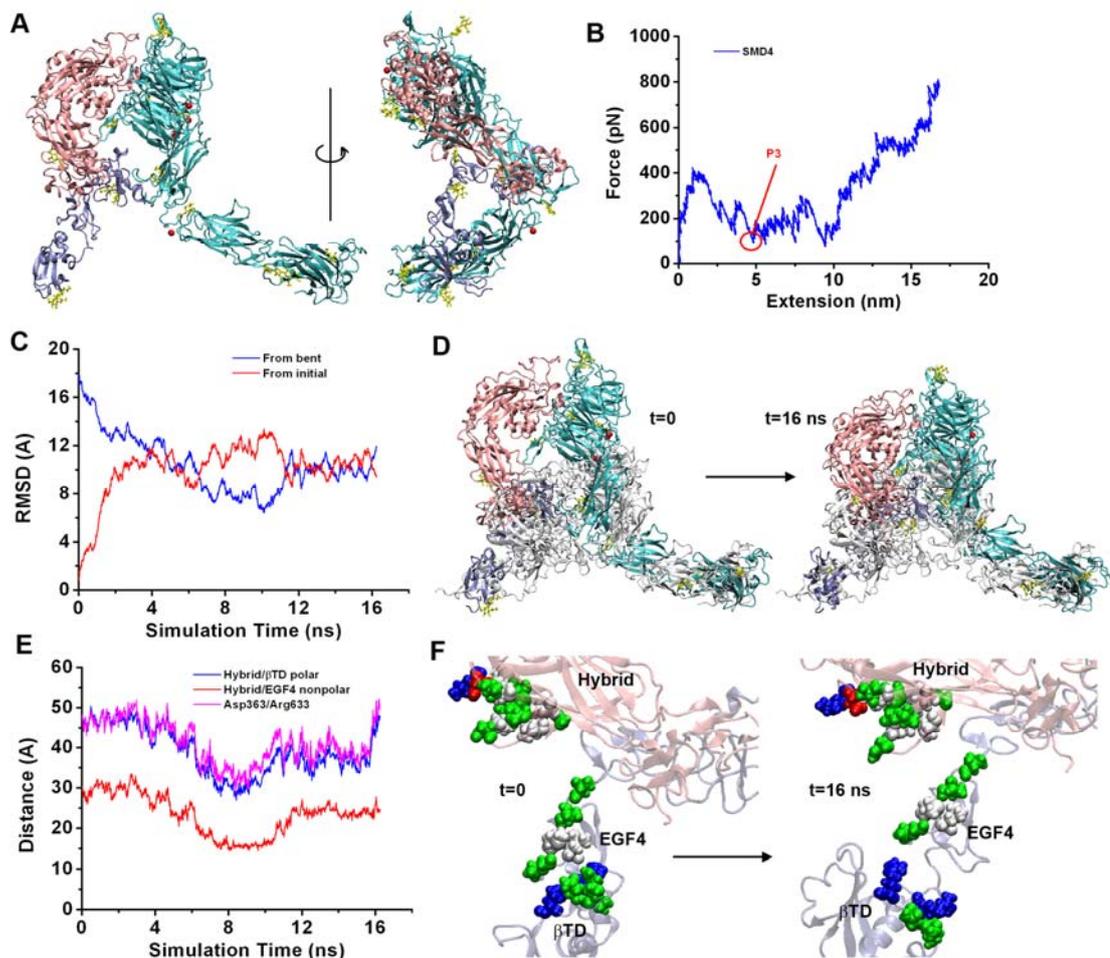
Figure 6.22 Free MD simulation of the P3 structure. A. Starting P3 structure. Right one was rotated 90 degree around the vertical axis. **B.** The starting P3 structure was chosen at indicate location right after the first force drop. **C.** RMSD of all Cα atoms from the equilibrated bent structure or the initial P3 structure. **D.** Comparison between P3 and bent structures at starting and ending time of the free MD simulation. **E.** Distances between sidechain COMs of groups of interacting residues in the bent structure: polar residues at Hybrid and βTD, nonpolar residues at Hybrid and EGF4, and a salt bridge (Asp363/Arg633) between Hybrid and βTD. **F.** Beginning and ending structures at Hybrid/EGF4/ βTD interface. Basic residues are in blue, acidic residues in red, polar residues in green, and nonpolar in white.

already broken, but EGF3 still interacted with Hybrid (Fig. 6.22A). During the free MD,

RMSD from the equilibrated bent structure with separated legs decreased (Fig. 6.22C).

But it was still large in the end (~ 10 Å), indicating P3 did not return to the bent structure

as seen in close leg simulations. Indeed, P3 was still in a partially extended conformation

(Fig. 6.22D). The interactions at Hybrid/EGF4/βTD, which existed in the bent structure,
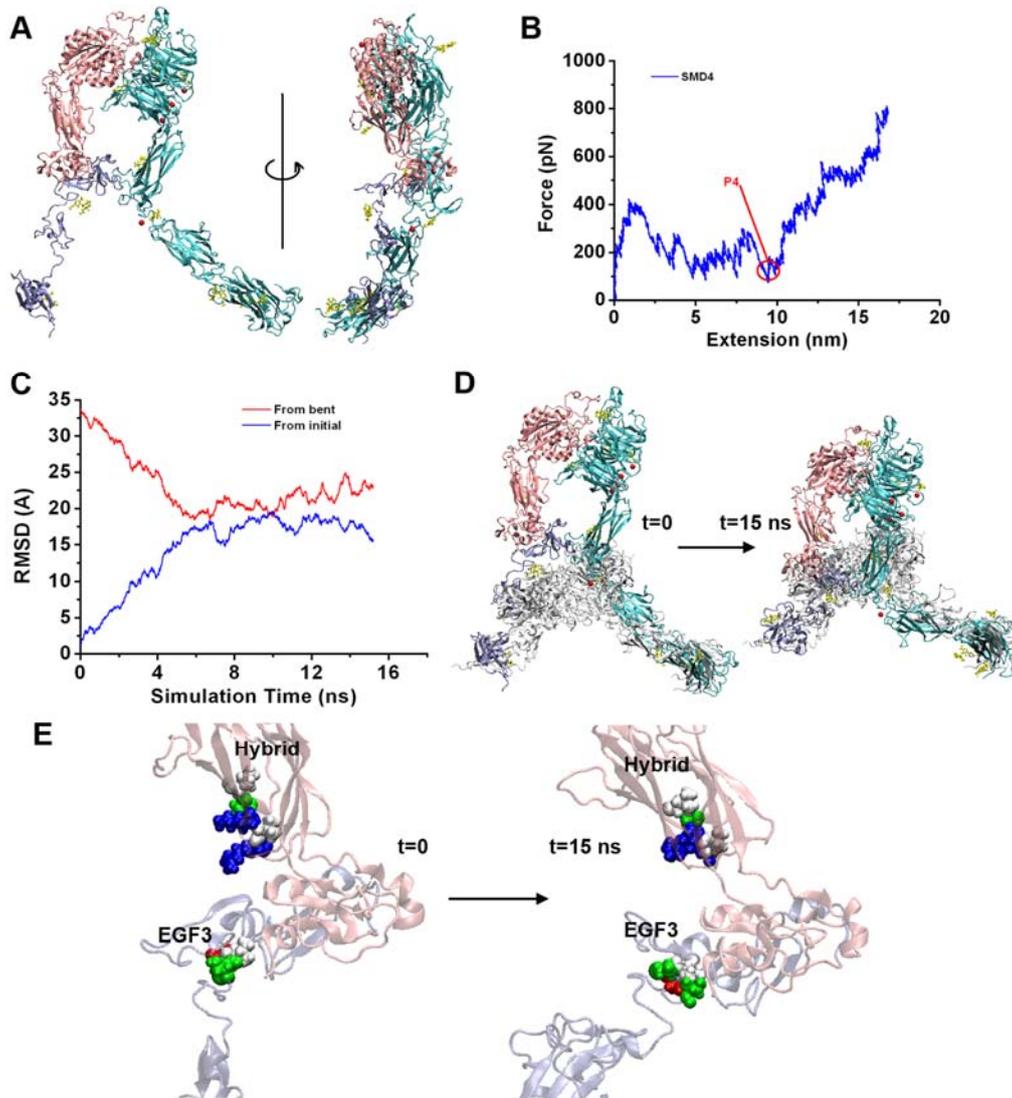
Figure 6.23 Free MD simulation of the P4 structure. A. Starting P4 structure. Right one was rotated 90 degree around the vertical axis. **B.** The starting P4 structure was chosen at indicate location right after the second force drop. **C.** RMSD of all Cα atoms from the equilibrated bent structure or the initial P4 structure. **D.** Comparison between P4 and bent structures at starting and ending time of the free MD simulation. **E.** Beginning and ending structures at Hybrid/EGF3 interface. Basic residues are in blue, acidic residues in red, polar residues in green, and nonpolar in white.

were not restored at all (Figs. 6.22 E and F).

A second partially extended conformation with separated legs (P4) was selected

right after the second force drop (Fig. 6.23B). In the beginning structure of P4,

interactions at the Hybrid/EGF3 interface were also disrupted (Fig. 6.23A). After force

release, P4 tended to bend back, indicated by decreasing RMSD from the bent structure
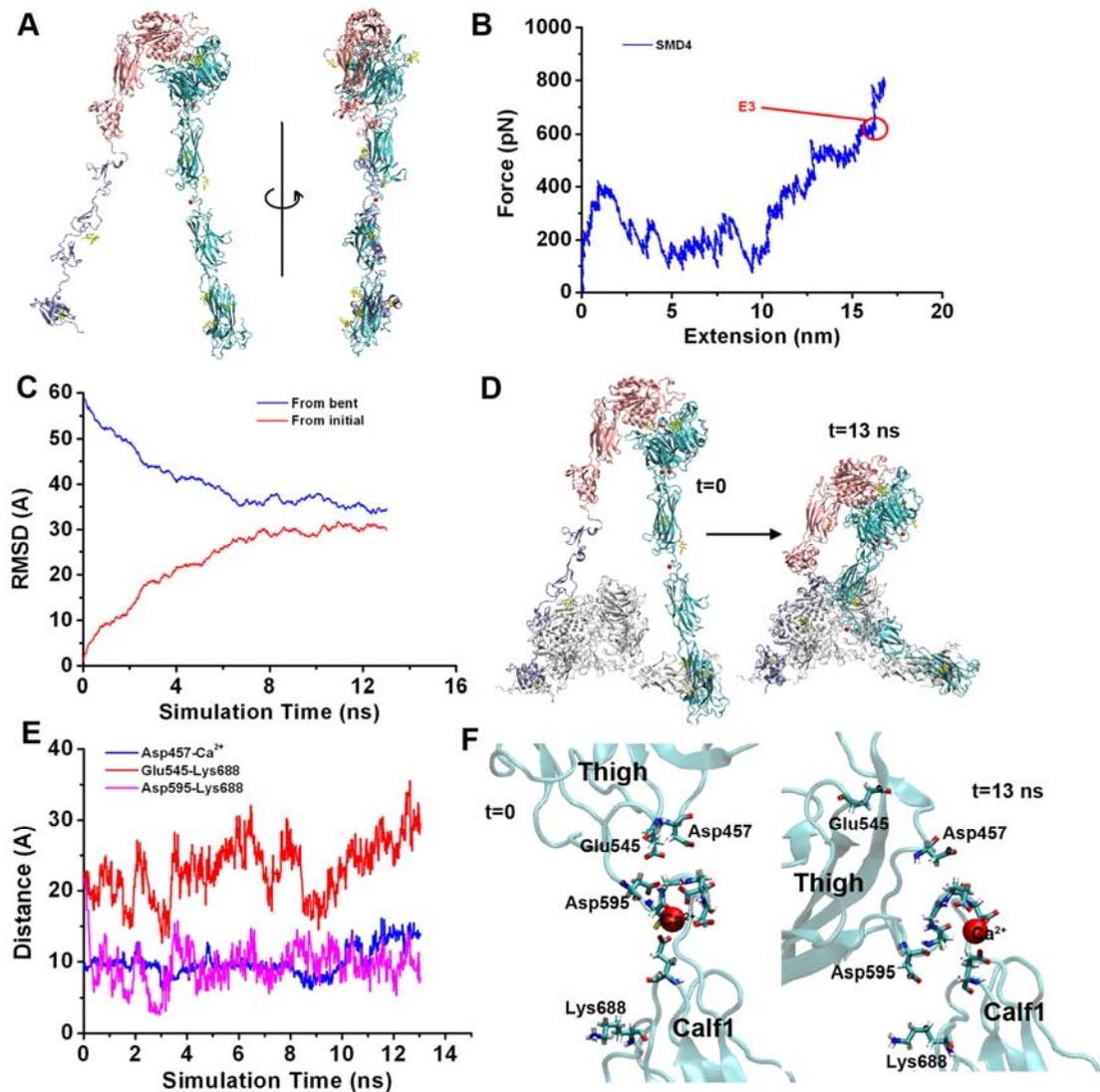
Figure 6.24 Free MD simulation of the E3 structure. A. Starting E3 structure. Right one was rotated 90 degree around the vertical axis. **B.** The starting E3 structure was chosen at indicate location when $\alpha_V\beta_3$ integrin was fully extended. **C.** RMSD of all Cα atoms from the equilibrated bent structure or the initial E3 structure. **D.** Comparison between E3 and bent structures at starting and ending time of the free MD simulation. **E.** Distances between three pairs of residues during free MD of E3: O_δ atoms of Asp457 and Ca^{2+} at α-genu, O_ε atoms of Glu545 and N_ζ atom of Lys688, and O_δ atoms of Asp595 and N_ζ atom of Lys688. **F.** Starting and ending structures of E3 at Thigh/Calf1 interface.

(Fig. 6.23C). However, P4 did not move close to the bent structure, shown by a larger

RMSD than P3 (~ 20 Å). Similarly, broken interactions at the Hybrid/EGF3 interface did

not restore (Fig. 6.23E).

Finally, a fully extended structure (E3) was extracted from the SMD4 trajectory (Fig. 6.24B). At the beginning, E3 was fully straight and most of interactions between the headpiece and the tailpiece were broken (Fig. 6.24A). During the free MD, E3 moved toward the bent conformation as indicated by decreasing RMSD from the bent structure, although RMSD was still large (~30 Å) (Fig. 6.24C). At the end of the free MD, E3 became a partially extended conformation (Fig. 6.24D). Some interactions were formed at the knee, but most of broken interactions between the headpiece and the tailpiece were not restored. In the previous free MD simulation for E1/E2 with closed legs, the fully extended structures were stable and salt bridges and calcium coordination at α-genu contributed to the stability. Consistently, Asp457-Ca^{2+} coordination and Glu545-Lys688 salt bridge were not observed during the entire free MD simulation for E3 so that the extended conformation with separated legs is not stable (Figs. 6.24 E and F).

CHAPTER 7

DISCUSSION

7.1 Glycosylation Effects on CD16-IgG Binding

In chapter 4, MD simulations and the MM-PB/GBSA method were applied to study the structural basis for glycosylation effects on CD16-IgG binding. The results showed that glycosylation of CD16 indeed changed the conformation of the CD16-IgG complex and thus their binding due to the interactions between carbohydrate chains attached on protein cores. The studies provide a structural mechanism how glycosylation regulates CD16-IgG binding affinity and offer insights into the molecular basis for the important roles of glycosylation in immune systems.

CD16 is a heavily glycosylated receptor expressed on immune cells and links the humoral to the cellular immune responses. Important roles of glycosylation of CD16 have been demonstrated by previous studies [4, 7] and the studies of our group (Ning Jiang, unpublished). The crystal structure of the CD16-IgG complex [5] provides insights into the effects of glycosylation on CD16-IgG binding. But no sugars are on the CD16 in the crystal structure, so it is still a mystery how glycosylation modulates CD16-IgG binding. In order to model a glycosylated CD16-IgG complex, an N-glycan core was added *in silico* to one of N-glycosylation sites of CD16, which locates at the binding pocket (Fig. 4.1). In the free MD simulation of the unglycosylated CD16-IgG complex, Cγ2 domains had quite large fluctuations around the Cγ2/3 junctions and showed the flexibility of IgG Fc (Fig. 4.2), which was not appreciated by looking at the static crystal structure. The flexibility of IgG Fc may be important for modulation of CD16-IgG binding. The two carbohydrate chains on IgG Fc were interacting with each other, which suggest that

sugars may not only act as spacers but also keep the integrity of the IgG Fc structure.

After adding sugars to CD16, a global conformational change was observed. CD16 was

shifted towards the Cγ2-B domain, Cγ2-B became more open, and Cγ2-A became more

closed (Fig. 4.5C). The conformational changes were caused by interactions between

sugars on CD16 and Cγ2-B of IgG (Fig. 4.6). Particularly, the Fucose residue of the sugar

on Cγ2-B was mainly involved in interactions with the sugar on CD16. Indeed, the

removal of this Fucose residue increases 3D affinity by over one order of magnitude [86].

As a result of the conformational changes, the binding pocket between CD16 and IgG Fc

was largely modified. In the unglycosylated complex, contact areas between CD16 and

IgG Fc were almost evenly distributed on Cγ2-A and Cγ2-B. After the conformational

changes, contacts at the CD16/Cγ2-A interface were less while contacts at the CD16/Cγ2-

A interface became more (Figs. 4.4 and 4.7).The number of hydrogen bonds between

CD16 and Cγ2-A decreased greatly. Most of lost hydrogen bonds involve residues of the

lower hinge of Cγ2-A (Table 4.1). This region was shown experimentally to be critical

for CD16-IgG binding by previous studies. Therefore, the observed changes of the

binding pocket suggest that the affinity of CD16-IgG binding may be decreased upon the

glycosylation of CD16.

To quantitatively evaluate the changes of the CD16-IgG, MM-PB/GBSA was

used to calculate the binding free energy of the CD16-IgG complexes. Comparison

between the computational and experimental results showed the PB method gave a better

estimate than the GB method (Fig. 4.9), which is reasonable because GB is an

approximation of PB. Although the difference of 4 kcal mol^{-1} is still much higher than the

chemical accuracy (1 kcal mol^{-1}), it is not bad considering relative cheap computation for

MM-PB/GBSA. However, PB and GB gave opposite results for comparison between the unglycosylated and glycosylated CD16 probably because the limit of sampling time and also the large error brought in by solute entropy calculation. Calculation of vibrational entropy by normal mode analysis is always very computational expensive, so much less samples were used, resulting in large error bars. These results show that it is difficult for MM-PB/GBSA to predict a small free energy difference of <10 kcal mol^{-1}. When the entropy part was excluded, both PB and GB showed less negative effective energy upon CD16 glycosylation. Decomposition of the effective energy on each residue revealed critical residues for CD16-IgG binding (Fig. 4.10). Same as seen in the structural analysis, the contribution from residues of Cγ2-A was largely decreased while that from residues of Cγ2-B was increased. Although it cannot give an accurate number of binding free energy, MM-PB/GBSA is really a powerfully tool to identify "hot-spots" in protein-protien interactions.

Another interesting thing was found in this study is that CD16 transitioned from the conformation in the complex crystal structure to the conformation in the stand-alone crystal structure. A transition state was observed, which is quite far from both conformations in the crystal structure. This observation demonstrates that the transition between two states generally does not follow a shortest distance pathway. Instead, the transition follows a pathway with the lowest energy cost, which may need a long distance movement. Furthermore, the transition was correlated well with residue Trp95 flipping, which is also seen in the crystal structure. So the conformational transition is probably controlled by the interactions between Tyr14 and Trp95. Either in crystals or in simulations, CD16 is a soluble form. When it is on membrane with a certain anchor, will

CD16 favors a certain conformation or amplifies the conformational change, and thus induces changes of affinity for IgG? Probably this could be an explanation of anchor effects on CD16-IgG binding observed in experiments [3].

7.2 Unfolding of the VWF A Domains

In chapter 5, MD simulations were used to study the unfolding mechanism of the VWF A1 and A2 domains by tensile force or under high temperature. The results showed how A1 and A2 were unfolded, providing the structural basis for VWF A domain stabilities and insights into the mechanism of the ADAMTS-13 cleavage of the A2 domain.

First, SMD was applied to simulate the forced unfolding of the VWF A1 and A2 domains. The unfolding of the A2 domain always started from the C-terminus and followed the same pathway no matter which terminus was pulled (Figs. 5.2 and 5.3). The sequential pullouts of the β-strands from the protein core generated sawtooth-like peaks in the force-extension curves, indicating that the central β-sheet was the most difficult to be unfolded by tensile force. Before unfolding, the proteolytic site for ADAMTS-13 cleavage was completely buried inside the A2 domain. After unfolding started, the sidechain of the proteolytic site began to expose. However, the backbone of the proteolytic site was not exposed until its neighboring β5 strand was pulled out. By comparison, the A1 domain was pulled at the C-terminus with or without its disulfide bond, which links its N- and C-termini. With the intact disulfide bond, the force could not propagate into the core of the A1 domain so that the core was not affected at all (Fig. 5.4). When the disulfide bond was broken *in silico* to mimic reducing conditions, the A1 domain was unfolded along a pathway similar to that of the A2 domain (Fig. 5.5).
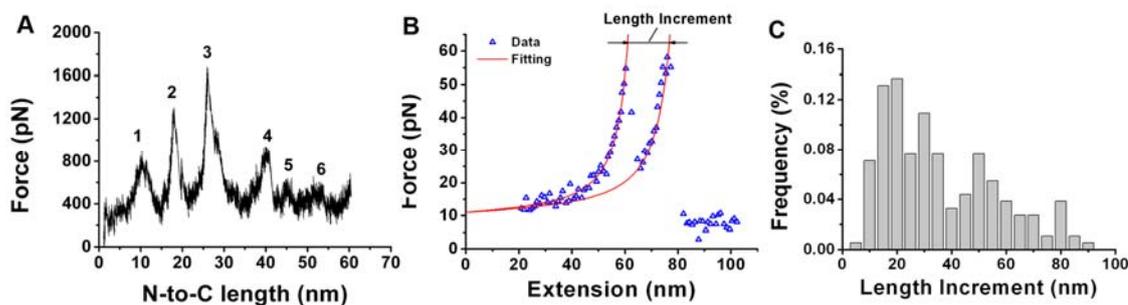
Figure 7.1 Comparison between simulations and experiments for forced unfolding of the VWF A2 domain. A. Plot of force vs. extension during the simulation of the A2 unfolding by force (replot of Fig. 5.2B). **B.** Plot of force vs. extension for the GPIb-A1A2A3 binding in one AFM experiment. Scatters are experimental data and curves are fitting by the worm-like chain model. **C.** Distribution of length increment due to structural failure observed in AFM experiments. Data in **B** and **C** are provided by Tao Wu and Jiangguo Lin.

The sawtooth-like force peak pattern (Fig. 7.1A) in the simulations of the VWF A domains unfolding is actually observed in force-ramping AFM experiments for the GPIb-A1A2A3 binding (Tao Wu et al., unpublished), where GPIb was adsorbed on AFM tips while an A1A2A3 triplet was coated on petri dishes. Multiple force peaks occurred. In the two-peak case as shown in Fig. 7.1B, the first force drop is due to certain structural failure and the second drop corresponds to the dissociation of the GPIb-A1A2A3 bond. By fitting the experimental data with the worm-like chain model, the difference between the contour lengths of the two force peaks indicates the length increment due to the structural failure. The length increment spans from 10 nm to 80 nm (Fig. 7.1C). The structural failure could be the breakage of inter-domain interactions among the three A domains or the unfolding of the A2 domain. If the structural failure is the A2 unfolding, the majority could not be the pullout of a single β strand because it generates only 10 nm length increment. Instead, it could be partial or full unfolding of A2. Considering the C-terminal portion of A2 is unfolded (first three peaks in Fig. 7.1A), the length increment is 20-30 nm, which corresponds to the first peak of the distribution in Fig. 7.1C. On the

other hand, when A2 is fully unfolded, the length increment is 50-60 nm, which corresponds to the second peak of the distribution.

It was further demonstrated that the specific topology of the A domains determines the sequential order of the pullouts of the β strands (Fig. 5.6). The N-terminal β1 strand is in the middle of the β sheet with many hydrogen bonds, while the C-terminal β6 strand is on the edge of the β sheet with only a few hydrogen bonds. Therefore, the C-terminus is much easier to be pulled out than the N-terminus. Moreover, the major resistance of the A1 and A2 domains to unfolding by tensile force came from the hydrogen network of the central β-sheet (Fig. 5.7). Two different unfolding pathways of β strands were observed: sliding and unzipping pathways. Unfolding along the former pathway would encounter a much higher energy barrier than the latter pathway (Fig. 5.8). Although the A1 and A2 domains unfolded quite similarly in terms of the pullouts of the β strands, they differed in the unfolding of their helices (Fig. 5.9).

The A2 domain hosts the proteolytic site for ADAMTS-13 while the A1 domain provides binding sites for GPIb [13, 14] and collagen [15]. Because the proteolytic site in the A2 domain is completely buried, the A2 domain should be easy to unfold to expose the proteolytic site *in vivo* for VWF to be cleaved. In contrast, the binding site for GPIb is on the surface of the N-terminal portion of the A1 domain (β3 side), so the A1 domain should be stable and not easy to unfold. Indeed, the GPIb-A1 complex structure shows little changes of the A1 domain upon binding [14]. The MD simulations demonstrate the importance of the disulfide bond in enhancing the A1 domain stability (Fig. 5.4B). The A2 domain does not have a disulfide bond, which makes it relatively easy to unfold.

The SMD simulations showed that the proteolytic site of the A2 domain was exposed step-by-step from fully buried to fully exposed (Figs. 5.2C and 5.3C). At what stage can ADAMTS-13 access and cleave the proteolytic site? The crystal structures of the catalytic domains of several ADAMTS family members, including ADAMTS-1, 4, and 5, were solved recently [87-89]. The active site locates at a shallow cleft on the surface of the catalytic domain, which may fit a single β strand. If that is the case, one side or all of the backbone of the β4 strand, where the proteolytic site is, needs to be exposed so that the β4 strand can get into the cleft of ADAMTS-13. Therefore, the A2 domain needs to be partially or fully unfolded for cleavage to occur under physiological condition. Could any intermediate states exist during the unfolding of the A2 domain? The pullouts of β5 and β4 generated the highest force peaks (Figs. 5.2B and 5.3B), showing largest energy barriers on the unfolding pathway. One intermediate between these two energy barriers could exist, in which one side of the backbone of the β4 strand is exposed. Further computational and experimental studies are required to reveal this possible intermediate state. Nevertheless, unfolding of the A2 domain is required for proteolysis by ADAMTS-13, making A2 a potential mechanosensor: only when a tensile force is above a certain threshold would the A2 domain be unfolded and cleaved by ADAMTS-13.

In the simulations of thermal unfolding of the A1 and A2 domains, the temperature was increased to 500 K and then the pressure was adjusted to 1 atm (Figs. 5.11A and 5.12A). After releasing restraints on the proteins, stepwise unfolding was observed in terms of RMSD from the starting structures (Figs. 5.11C and 5.12C). The C-terminal portion of the A2 domain was most susceptible to thermal unfolding. During the

46 ns simulation, the α7, α6 helices, and β6 strand were unfolded into loops and flipped away from their original position (Fig. 5.11E), so the sidechain of the proteolytic site was partially exposed (Fig. 5.11D). Meanwhile, part of the α5-loop was converted to a helix. In contrast, the N-terminal portion of the A1 domain was easier to be unfolded by thermal forces. During the 30 ns simulation, the β3 strand was disconnected from the β-sheet and converted to a helix. Then the β3 and α3 helices flipped upward away from their original position (Fig. 5.12D). At the same time, the α4 helix was unfolded to a loop. On the C-terminal portion of the A1 domain, the β6 strand was also converted to a helix while the β7 helix was broken into two parts. The thermal unfolding of the A1 and A2 domains was very different from the unfolding by tensile force, reflecting the different ways that tensile force and thermal force act on the proteins.

In the thermal unfolding simulations, the proteolytic site of A2 was only partially exposed. As discussed above, ADAMTS-13 may not be able to access the proteolytic site at this stage. It could be imagined that the whole domain can fall apart if simulations are run longer. Then the proteolytic site is mostly exposed for ADAMTS-13 cleavage.

7.3 Unbending of $\alpha_V\beta_3$ Integrin

In chapter 6, the unbending of $\alpha_V\beta_3$ integrin from a bent to an extended conformation was simulated by SMD. Critical residues that regulate the unbending were identified. Furthermore, the stabilities of different conformations along the unbending pathway were tested. The study suggests the structural basis for the transition between different conformations and provides insights into the mechanism of regulation of integrin functions through global conformational changes.

Global conformational changes of integrins are thought to be directly related to regulation of integrin funcions as suggested in the switchblade model [45, 47, 50]. Starting from the crystal structure of the $\alpha_V\beta_3$ ectodomain, which was the only available structure of the full integrin ectodomain at the time of the study, homology models of the two unresolved EGF1/2 domains were built based on the β_2 subunit structure to obtain a complete ectodomain structure for simulations. It turns out that the homology models are not bad. Compared to the recent published crystal structure of $\alpha_V\beta_3$ ectodomain [46], the location and the overall structure and conformation of the models are similar.

In SMD simulations, the βA domain was pulled by a force while the βTD domain was restrained. Very nicely, the simulated integrin became extended gradually (Fig. 6.7). A major force peak was observed, corresponded to the breakage of interactions at Hybrid/EGF4 and Hybrid/βTD interfaces (Fig. 6.8). Interestingly, the interactions between Hybrid and βTD are mainly hydrophilic, including a salt bridge Asp393/Arg633. The importance of residue Arg633 was recognized by Matsumoto et al. [90], although the salt bridge was not mentioned. In contrast, the interactions between Hybrid and EGF4 are mostly hydrophobic. The importance of these interactions was further emphasized by free MD simulations of a partially extended conformation. During the free MD, the hydrophilic residues at Hybrid and βTD first contacted with each other, and then the hydrophobic interactions between Hybrid and EGF4 stabilized the bent conformation (Fig. 6.11).

In the constant-velocity SMD simulations, the length increase after the first force drop due to the breakage of interactions at the Hybrid/EGF4/βTD interfaces is ~5 nm (Fig. 7.1A). In the AFM experiments for the $\alpha_5\beta_1$ integrin-Fibronectin binding (Fang
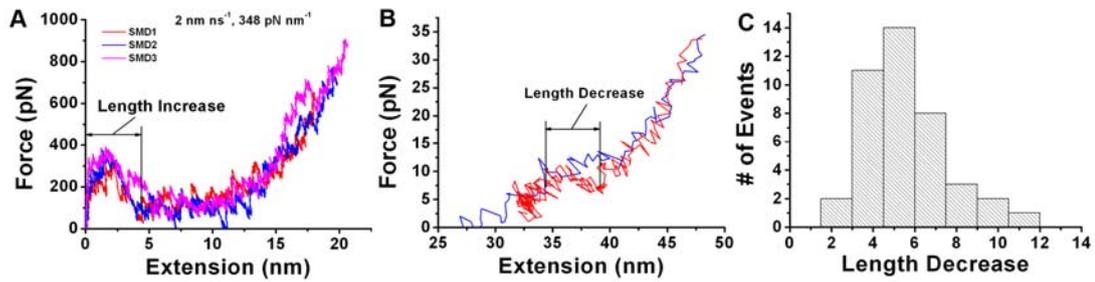
Figure 7.2 Comparison between simulations and experiments for unbending of integrin with constant-velocity pulling. A. Plot of force vs. extension during constant-velocity SMD simulations (replot of Fig. 6.7B). **B.** Plot of force vs. extension for the $\alpha_5\beta_1$ integrin-Fibronectin binding during one AFM experiment. Blue color indicates the force loading phase while red color shows the force releasing phase. **C.** Distribution of length decrease observed in AFM experiments. Data in B and C are provided by Kong Fang.

Kong, unpublished), where the cantilever first moves up at a constant speed (force loading phase) and then goes down at another constant speed (force releasing phase), a sudden length change is also observed (Fig. 7.1B). The distribution of the length decrease has a maximum at ~5 nm, similar to that in simulations. The legless $\alpha_5\beta_1$ integrin does not have the sudden length change (data not shown), therefore the change is most likely because of the breakage (or restoration) of interactions at the Hybrid/EGF4/βTD interfaces during unbending (or rebending). The full extension of $\alpha_V\beta_3$ integrin results in a length increase of ~15 nm, which is also shown in a constant-force (200 pN) SMD simulation (Fig. 7.3A). Indeed, when the pulling force is kept at a constant level in biomembrane force probe (BFP) experiments for the $\alpha_L\beta_2$ integrin-ICAM1 binding (Wei Jack Chen, unpublished), a sudden length increase was observed (Fig. 7.3B). The distribution of length increment has a maximum at ~17 nm (Fig. 7.3C). Since $\alpha_L\beta_2$ has an additional αA domain, it is reasonable that $\alpha_L\beta_2$ has a bigger extension than $\alpha_V\beta_3$. The consistence of simulation observation with experiments justifies the simulation results and also provides structural explanation for experimental data.
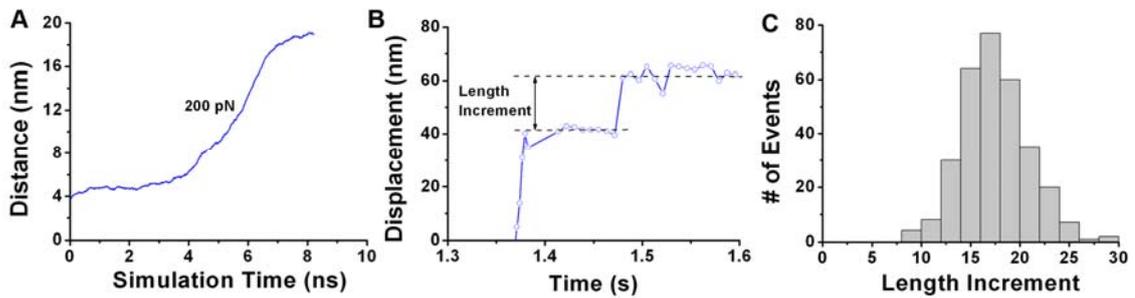
Figure 7.3 Comparison between simulations and experiments for unbending of integrin under constant force. A. Plot of distance between the βA and βTD domains vs. simulation time during a constant-force (200 pN) SMD simulation of $\alpha_V\beta_3$ integrin. **B.** Plot of displacement vs. time for the $\alpha_L\beta_2$ integrin-ICAM1 binding during one BFP experiment. **C.** Distribution of length increment observed in BFP experiments. Data in B and C are provided by Wei Jack Chen.

After the extension of simulated $\alpha_V\beta_3$ integrin, the Hybrid domain did not swing out and the leg did not separate. So extension only is not enough for Hybrid swing-out or leg separation. Probably, they are not coupled. There were no ligands in the simulated integrin. One would think maybe ligands could make Hybrid swing-out or leg separation. However, after extension, the interactions between the legs are mostly kept. That means Hybrid swing-out or leg separation is not easier after extension. Could force help it? In the SMD simulations, the force was along the extension direction (i.e., normal to the imaginary cell membrane). The force did not induce Hybrid swing-out. Instead, it made Hybrid even swing-in after the integrin was fully stretched (Fig. 6.9). This is the same as what Zhu et al. found [46]. Only if there is another force parallel to the cell membrane, the legs can be separated and Hybrid can swing out.

After fully extension of simulated integrin, free MD simulations were run to test its stability. In two simulations, the extended conformation was stable during 20 ns (Figs. 6.14 and 6.15). Two new interactions were identified at α-genu. One is that residue Asp457 of the Thigh domain coordinated with the Ca^{2+} ion at α-genu. The other is that Glu545 of Thigh formed a salt bridge with Lys688 of Calf1. These two interactions

bridged Thigh and Calf1 together and thus stabilized the extended conformation. This provides an explanation for how metal ions regulate integrin conformation. Maybe different metal ions have different abilities to form new coordination with Asp457. Mn^{2+} has higher ability so that it favors the extended conformation. In contrast, Ca^{2+} has lower ability so that it favors the bent conformation.

In the switchblade model, the leg separation can induce extension during inside-out signaling. In order to test it, SMD simulations were run to separate legs. A big force peak was observed, showing the breakage of interactions between the legs (Fig. 6.16). This force peak is larger than that during the unbending, showing the interactions between the legs are stronger than those between the headpiece and the tailpiece. Indeed, the broken interactions involve extensive contacts at Calf2/βTD, Calf2/β-ankle, Calf2/EGF4, Calf1/EGF3, and Calf1/EGF2 interfaces (Fig. 6.17). After leg separation, the interactions between the headpiece and the tailpiece of β-subunit were mostly kept. Even more new interactions formed between Hybrid and EGF3 or between PSI and EGF2. As a result, the unbending with separated legs has two additional peaks (Fig. 6.19), indicating more resistance. This demonstrates that leg separation does not facilitate extension.

The free MD simulations with separated legs showed that a fully extended conformation was not stable because the interactions that stabilized the extended conformation with closed legs were not observed in the case of separated legs (Fig. 6.24). Also, a partially extended conformation did not return to the bent conformation as what occurred with closed legs (Figs. 6.22 and 6.23). All these suggest that the interactions

between the legs stabilize both the extended and the bent conformations. With separated legs, the integrin probably favors a partially extended conformation.

Experiments suggest that talin head alone bound to the cytoplasmic tail of integrins can activate integrins. From the simulations here, it is hard to imagine talin head binding would give enough energy to separate legs and extend integrins. Most likely, some smaller conformational changes occur as suggested by the deadbolt model [51] so that βA switches to an intermediate affinity state for ligand binding. After engagement of ligands and/or with force, the integrin can be unbent and the legs can be separated so that the integrin is brought into a high affinity state.

CHAPTER 8

RECOMMENDATIONS FOR FUTURE WORK

CD16-IgG Binding

The study in this thesis provides one structural mechanism for regulation of the CD16-IgG binding by glycosylation and gives ideas about the general role of glycosylation on Ig-FcR interactions. Actually, the important role of glycosylation on antibodies has been recognized since 1980s [91]. The general interests and the success of the current project inspire further studies.

Residue Fucose of the carbohydrate chain on IgG Fc was shown to be important by MD simulations here or published experimental data [86]. It is worth to run MD simulations with the removal of the Fucose to see how the conformation and the binding will be changed.

Over 30 glycoforms have been identified for attached glycans on IgG. Individual glycoforms of IgG may contribute to modulating inflammatory responses. Kaneko et al. [92] showed that even the change of a sugar residue far away from the binding pocket can change the IgG-FcR binding and following effector functions. So it is of interests to look at the effects of those sugar residues on the IgG-FcR binding by molecular simulations and understand the structural mechanism.

CD16 conformational changes were observed during MD simulations. The conformational changes may affect the CD16-IgG binding. With different anchors, CD16 may favor different conformations or amplify the observed conformational changes, resulting in changes of the binding affinity. To test this hypothesis, MD simulations for

CD16 with anchors need to be performed to see how the conformations and dynamics of CD16 and the CD16-IgG binding will be affected.

VWF A Domains

In this study, the VWF A1 and A2 domains were unfolded in simulations either by a force or under high temperature. The observed unfolding pathways provide structural basis for understanding their stabilities and offer insights into the way ADAMTS-13 interacts with and cleaves the A2 domains. Based on current study, further research could be pursued as follows.

Does a stable intermediate state exist during the unfolding of A2? If yes, this intermediate state could imply how ADAMTS-13 accesses the cleavage site and cuts it. Constant-force SMD with relatively small force should be used to detect the intermediate because the intermediates are expected to last longer under smaller force.

An A2/ADAMTS-13 complex structure is needed for understanding the mechanism of VWF cleavage by ADAMTS-13. Currently, no ADAMTS-13 crystal structure is available. But the crystal structures of several homologous ADAMTS family members, including ADAMTS-1, 4, and 5, are available [87-89]. So the homology model of ADAMTS-13 can be made. Combined with the unfolding structures of A2, a homology model of A2/ADAMTS-13 complex can be obtained. The homology model could reveal important interactions between A2 and ADAMTS-13 and shed light on the mechanism of VWF cleavage by ADAMTS-13.

Most mutations associated with type 2A VWD locate at the A2 domain [31, 32]. These mutations may affect unfolding of the A2 domain, its interactions with other VWF domains, and/or its interactions with ADAMTS-13. Sutherland et al. [25] showed that

some of the type 2A VWD mutations caused some structural changes of the A2 domain with equilibration at 310 K. Do these mutations affect the resistance of the A2 domain to unfolding or change the unfolding pathway? Future unfolding simulations with these mutations should be performed.

Integrin Global Conformational Changes

In this thesis, the transition of integrin conformations from the bent to the extended was successfully simulated. The study provides the structural basis for understanding the mechanism of integrin functions.

At the time of preparation of this thesis, a new crystal structure of $\alpha_{IIb}\beta_3$ just came out [46]. Probably more crystal structures of other integrins will be available in the near future. Because the unresolved domains of $\alpha_V\beta_3$ are visible in the new crystal structure of $\alpha_{IIb}\beta_3$, it is worth to run simulations with the new structure to justify what was obtained here, although it may not change the major conclusions since it only affects the interactions at β-genu.

The variations of sequences across integrin family are big. The critical residues identified for $\alpha_V\beta_3$ are mostly not conserved among other integrins. It suggests that different integrins may have different stabilities of either the bent or the extended conformation, which may determine how hard integrins can be activated. Therefore, it is quite interesting to simulate conformational changes of different integrins. The comparison will give clues to understand the mechanism of integrin activation.

Just before submission of the thesis, the structure of the transmembrane segments of $\alpha_{IIb}\beta_3$ integrin by nuclear magnetic resonance (NMR) was released [93]. So the structures of most parts of integrin are available now. With these structures, a complete

structure of the whole integrin can be obtained. Therefore, it is time to simulate the whole

integrin anchored in a lipid bilayer, mimicking the membrane environment. With this

system, the pathway of signal transduction between the cytoplasmic tails and the ligand

binding head can be studied. It sounds like a daunting task because the system will be

even larger than that was simulated here. The real problem is that the time scale can be

reached is only 10 ns order, in which the interesting signal propagation is hardly seen. A

coarse-grain model may be more suitable for simulating the system. The coarse-grain

model can simulate the propagation of structural changes through multiple domains,

while all-atom simulations are still needed to identify important interactions and study

subtle changes in the binding pocket. Combination of these two types of simulations is a

promising method to reveal the structural mechanism of integrin functions.

REFERENCES

1. Sondermann, P. and V. Oosthuizen. The structure of Fc receptor/Ig complexes: considerations on stoichiometry and potential inhibitors. *Immunol Lett* **82**(1-2): 51-56, 2002.

2. Cohen-Solal, J.F., L. Cassard, W.H. Fridman, and C. Sautes-Fridman. Fc gamma receptors. *Immunol Lett* **92**(3): 199-205, 2004.

3. Chesla, S.E., P. Li, S. Nagarajan, P. Selvaraj, and C. Zhu. The membrane anchor influences ligand binding two-dimensional kinetic rates and three-dimensional affinity of FcgammaRIII (CD16). *J Biol Chem* **275**(14): 10235-10246, 2000.

4. Edberg, J.C. and R.P. Kimberly. Cell type-specific glycoforms of Fc gamma RIIIa (CD16): differential ligand binding. *J Immunol* **159**(8): 3849-3857, 1997.

5. Sondermann, P., R. Huber, V. Oosthuizen, and U. Jacob. The 3.2-A crystal structure of the human IgG1 Fc fragment-Fc gammaRIII complex. *Nature* **406**(6793): 267-273, 2000.

6. Radaev, S., S. Motyka, W.H. Fridman, C. Sautes-Fridman, and P.D. Sun. The structure of a human type III Fcgamma receptor in complex with Fc. *J Biol Chem* **276**(19): 16469-16477, 2001.

7. Drescher, B., T. Witte, and R.E. Schmidt. Glycosylation of FcgammaRIII in N163 as mechanism of regulating receptor affinity. *Immunology* **110**(3): 335-340, 2003.

8. Sadler, J.E. von Willebrand factor: two sides of a coin. *J Thromb Haemost* **3**(8): 1702-1709, 2005.

9. Franchini, M. and G. Lippi. Von Willebrand factor and thrombosis. *Ann Hematol* **85**(7): 415-423, 2006.

10. Chen, J. and J.A. Lopez. Interactions of platelets with subendothelium and endothelium. *Microcirculation* **12**(3): 235-246, 2005.

11. Takahashi, Y., M. Kalafatis, J.P. Girma, K. Sewerin, L.O. Andersson, and D. Meyer. Localization of a factor VIII binding domain on a 34 kilodalton fragment of the N-terminal portion of von Willebrand factor. *Blood* **70**(5): 1679-1682, 1987.

12. Plow, E.F., T.A. Haas, L. Zhang, J. Loftus, and J.W. Smith. Ligand binding to integrins. *J Biol Chem* **275**(29): 21785-21788, 2000.

13. Matsushita, T. and J.E. Sadler. Identification of amino acid residues essential for von Willebrand factor binding to platelet glycoprotein Ib. Charged-to-alanine

scanning mutagenesis of the A1 domain of human von Willebrand factor. *J Biol Chem* **270**(22): 13406-13414, 1995.

14.   Huizinga, E.G., S. Tsuji, R.A. Romijn, M.E. Schiphorst, P.G. de Groot, J.J. Sixma, and P. Gros. Structures of glycoprotein Ibalpha and its complex with von Willebrand factor A1 domain. *Science* **297**(5584): 1176-1179, 2002.

15.   Mazzucato, M., P. Spessotto, A. Masotti, L. De Appollonia, M.R. Cozzi, A. Yoshioka, R. Perris, A. Colombatti, and L. De Marco. Identification of domains responsible for von Willebrand factor type VI collagen interaction mediating platelet adhesion under high flow. *J Biol Chem* **274**(5): 3033-3041, 1999.

16.   Sixma, J.J., M.E. Schiphorst, C.L. Verweij, and H. Pannekoek. Effect of deletion of the A1 domain of von Willebrand factor on its binding to heparin, collagen and platelets in the presence of ristocetin. *Eur J Biochem* **196**(2): 369-375, 1991.

17.   Dent, J.A., S.D. Berkowitz, J. Ware, C.K. Kasper, and Z.M. Ruggeri. Identification of a cleavage site directing the immunochemical detection of molecular abnormalities in type IIA von Willebrand factor. *Proc Natl Acad Sci U S A* **87**(16): 6306-6310, 1990.

18.   Levy, G.G., W.C. Nichols, E.C. Lian, T. Foroud, J.N. McClintick, B.M. McGee, A.Y. Yang, D.R. Siemieniak, K.R. Stark, R. Gruppo, R. Sarode, S.B. Shurin, V. Chandrasekaran, S.P. Stabler, H. Sabio, E.E. Bouhassira, J.D. Upshaw, Jr., D. Ginsburg, and H.M. Tsai. Mutations in a member of the ADAMTS gene family cause thrombotic thrombocytopenic purpura. *Nature* **413**(6855): 488-494, 2001.

19.   Romijn, R.A., B. Bouma, W. Wuyster, P. Gros, J. Kroon, J.J. Sixma, and E.G. Huizinga. Identification of the collagen-binding site of the von Willebrand factor A3-domain. *J Biol Chem* **276**(13): 9985-9991, 2001.

20.   Bienkowska, J., M. Cruz, A. Atiemo, R. Handin, and R. Liddington. The von willebrand factor A3 domain does not contain a metal ion-dependent adhesion site motif. *J Biol Chem* **272**(40): 25162-25167, 1997.

21.   Huizinga, E.G., R. Martijn van der Plas, J. Kroon, J.J. Sixma, and P. Gros. Crystal structure of the A3 domain of human von Willebrand factor: implications for collagen binding. *Structure* **5**(9): 1147-1156, 1997.

22.   Emsley, J., M. Cruz, R. Handin, and R. Liddington. Crystal structure of the von Willebrand Factor A1 domain and implications for the binding of platelet glycoprotein Ib. *J Biol Chem* **273**(17): 10396-10401, 1998.

23.   Lee, J.O., P. Rieu, M.A. Arnaout, and R. Liddington. Crystal structure of the A domain from the alpha subunit of integrin CR3 (CD11b/CD18). *Cell* **80**(4): 631-638, 1995.

24.   Qu, A. and D.J. Leahy. Crystal structure of the I-domain from the CD11a/CD18 (LFA-1, alpha L beta 2) integrin. *Proc Natl Acad Sci U S A* **92**(22): 10277-10281, 1995.

25.   Sutherland, J.J., L.A. O'Brien, D. Lillicrap, and D.F. Weaver. Molecular modeling of the von Willebrand factor A2 Domain and the effects of associated type 2A von Willebrand disease mutations. *J Mol Model* **10**(4): 259-270, 2004.

26.   Dong, J.F., J.L. Moake, L. Nolasco, A. Bernardo, W. Arceneaux, C.N. Shrimpton, A.J. Schade, L.V. McIntire, K. Fujikawa, and J.A. Lopez. ADAMTS-13 rapidly cleaves newly secreted ultralarge von Willebrand factor multimers on the endothelial surface under flowing conditions. *Blood* **100**(12): 4033-4039, 2002.

27.   Arya, M., B. Anvari, G.M. Romo, M.A. Cruz, J.F. Dong, L.V. McIntire, J.L. Moake, and J.A. Lopez. Ultralarge multimers of von Willebrand factor form spontaneous high-strength bonds with the platelet glycoprotein Ib-IX complex: studies using optical tweezers. *Blood* **99**(11): 3971-3977, 2002.

28.   Padilla, A., J.L. Moake, A. Bernardo, C. Ball, Y. Wang, M. Arya, L. Nolasco, N. Turner, M.C. Berndt, B. Anvari, J.A. Lopez, and J.F. Dong. P-selectin anchors newly released ultralarge von Willebrand factor multimers to the endothelial cell surface. *Blood* **103**(6): 2150-2156, 2004.

29.   Auton, M., M.A. Cruz, and J. Moake. Conformational stability and domain unfolding of the Von Willebrand factor A domains. *J Mol Biol* **366**(3): 986-1000, 2007.

30.   Tsai, H.M. Deficiency of ADAMTS13 and thrombotic thrombocytopenic purpura. *Blood* **100**(10): 3839-3840; author reply 3840-3832, 2002.

31.   Castaman, G., A.B. Federici, F. Rodeghiero, and P.M. Mannucci. Von Willebrand's disease in the year 2003: towards the complete identification of gene defects for correct diagnosis and treatment. *Haematologica* **88**(1): 94-108, 2003.

32.   Michiels, J.J., A. Gadisseur, U. Budde, Z. Berneman, M. van der Planken, W. Schroyens, A. van de Velde, and H. van Vliet. Characterization, classification, and treatment of von Willebrand diseases: a critical appraisal of the literature and personal experiences. *Semin Thromb Hemost* **31**(5): 577-601, 2005.

33.   Tsai, H.M., Sussman, II, and R.L. Nagel. Shear stress enhances the proteolysis of von Willebrand factor in normal plasma. *Blood* **83**(8): 2171-2179, 1994.

34.   Yago, T., J. Lou, T. Wu, J. Yang, J.J. Miner, L. Coburn, J.A. Lopez, M.A. Cruz, J.F. Dong, L.V. McIntire, R.P. McEver, and C. Zhu. Platelet glycoprotein Ibalpha forms catch bonds with human WT vWF but not with type 2B von Willebrand disease vWF. *J Clin Invest* **118**(9): 3195-3207, 2008.

35. Furlan, M., R. Robles, and B. Lamie. Partial purification and characterization of a protease from human plasma cleaving von Willebrand factor to fragments produced by in vivo proteolysis. *Blood* **87**(10): 4223-4234, 1996.

36. Tsai, H.M. Physiologic cleavage of von Willebrand factor by a plasma protease is dependent on its conformation and requires calcium ion. *Blood* **87**(10): 4235-4244, 1996.

37. Shimaoka, M., J. Takagi, and T.A. Springer. Conformational regulation of integrin structure and function. *Annu Rev Biophys Biomol Struct* **31**: 485-516, 2002.

38. Hynes, R.O. Integrins: bidirectional, allosteric signaling machines. *Cell* **110**(6): 673-687, 2002.

39. Takagi, J. and T.A. Springer. Integrin activation and structural rearrangement. *Immunol Rev* **186**: 141-163, 2002.

40. Arnaout, M.A., B. Mahalingam, and J.P. Xiong. Integrin structure, allostery, and bidirectional signaling. *Annu Rev Cell Dev Biol* **21**: 381-410, 2005.

41. Luo, B.H., C.V. Carman, and T.A. Springer. Structural basis of integrin regulation and signaling. *Annu Rev Immunol* **25**: 619-647, 2007.

42. Xiong, J.P., T. Stehle, B. Diefenbach, R. Zhang, R. Dunker, D.L. Scott, A. Joachimiak, S.L. Goodman, and M.A. Arnaout. Crystal structure of the extracellular segment of integrin alpha Vbeta3. *Science* **294**(5541): 339-345, 2001.

43. Xiong, J.P., T. Stehle, R. Zhang, A. Joachimiak, M. Frech, S.L. Goodman, and M.A. Arnaout. Crystal structure of the extracellular segment of integrin alpha Vbeta3 in complex with an Arg-Gly-Asp ligand. *Science* **296**(5565): 151-155, 2002.

44. Xiong, J.P., T. Stehle, S.L. Goodman, and M.A. Arnaout. A novel adaptation of the integrin PSI domain revealed from its crystal structure. *J Biol Chem* **279**(39): 40252-40254, 2004.

45. Takagi, J., B.M. Petre, T. Walz, and T.A. Springer. Global conformational rearrangements in integrin extracellular domains in outside-in and inside-out signaling. *Cell* **110**(5): 599-511, 2002.

46. Zhu, J., B.H. Luo, T. Xiao, C. Zhang, N. Nishida, and T.A. Springer. Structure of a complete integrin ectodomain in a physiologic resting state and activation and deactivation by applied forces. *Mol Cell* **32**(6): 849-861, 2008.

47. Xiao, T., J. Takagi, B.S. Coller, J.H. Wang, and T.A. Springer. Structural basis for allostery in integrins and binding to fibrinogen-mimetic therapeutics. *Nature* **432**(7013): 59-67, 2004.

48. Takagi, J., K. Strokovich, T.A. Springer, and T. Walz. Structure of integrin alpha5beta1 in complex with fibronectin. *EMBO J* **22**(18): 4607-4615, 2003.

49. Nishida, N., C. Xie, M. Shimaoka, Y. Cheng, T. Walz, and T.A. Springer. Activation of leukocyte beta2 integrins by conversion from bent to extended conformations. *Immunity* **25**(4): 583-594, 2006.

50. Beglova, N., S.C. Blacklow, J. Takagi, and T.A. Springer. Cysteine-rich module structure reveals a fulcrum for integrin rearrangement upon activation. *Nat Struct Biol* **9**(4): 282-287, 2002.

51. Xiong, J.P., T. Stehle, S.L. Goodman, and M.A. Arnaout. New insights into the structural basis of integrin activation. *Blood* **102**(4): 1155-1159, 2003.

52. Adair, B.D., J.P. Xiong, C. Maddock, S.L. Goodman, M.A. Arnaout, and M. Yeager. Three-dimensional EM structure of the ectodomain of integrin {alpha}V{beta}3 in a complex with fibronectin. *J Cell Biol* **168**(7): 1109-1118, 2005.

53. Case, D.A., T.A. Darden, T.E. Cheatham, III, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, K.M. Merz, B. Wang, D.A. Pearlman, M. Crowley, S. Brozell, V. Tsui, H. Gohlke, J. Mongan, V. Hornak, G. Cui, P. Beroza, C. Schafmeister, J.W. Caldwell, W.S. Ross, and P.A. Kollman, *AMBER 8*. 2004, University of California: San Francisco.

54. Cornell, W.D., P. Cieplak, C.I. Bayly, I.R. Gould, K.M. Merz, D.M. Ferguson, D.C. Spellmeyer, T. Fox, J.W. Caldwell, and P.A. Kollman. A 2nd Generation Force-Field for the Simulation of Proteins, Nucleic-Acids, and Organic-Molecules. *Journal of the American Chemical Society* **117**(19): 5179-5197, 1995.

55. Duan, Y., C. Wu, S. Chowdhury, M.C. Lee, G. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee, J. Caldwell, J. Wang, and P. Kollman. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J Comput Chem* **24**(16): 1999-2012, 2003.

56. Woods, R.J., R.A. Dwek, C.J. Edge, and B. Fraserreid. Molecular Mechanical and Molecular Dynamical Simulations of Glycoproteins and Oligosaccharides .1. Glycam-93 Parameter Development. *Journal of Physical Chemistry* **99**(11): 3832-3846, 1995.

57. Phillips, J.C., R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R.D. Skeel, L. Kale, and K. Schulten. Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry* **26**(16): 1781-1802, 2005.

58.   Berman, H.M., J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, and P.E. Bourne. The Protein Data Bank. *Nucleic Acids Research* **28**(1): 235-242, 2000.

59.   Darden, T., D. York, and L. Pedersen. Particle Mesh Ewald - an N.Log(N) Method for Ewald Sums in Large Systems. *Journal of Chemical Physics* **98**(12): 10089-10092, 1993.

60.   Ryckaert, J.P., G. Ciccotti, and H.J.C. Berendsen. Numerical-Integration of Cartesian Equations of Motion of a System with Constraints - Molecular-Dynamics of N-Alkanes. *Journal of Computational Physics* **23**(3): 327-341, 1977.

61.   Humphrey, W., A. Dalke, and K. Schulten. VMD: Visual molecular dynamics. *Journal of Molecular Graphics* **14**(1): 33-&, 1996.

62.   Frishman, D. and P. Argos. Knowledge-based protein secondary structure assignment. *Proteins* **23**(4): 566-579, 1995.

63.   Lu, H., B. Isralewitz, A. Krammer, V. Vogel, and K. Schulten. Unfolding of titin immunoglobulin domains by steered molecular dynamics simulation. *Biophys J* **75**(2): 662-671, 1998.

64.   Needleman, S.B. and C.D. Wunsch. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J Mol Biol* **48**(3): 443-453, 1970.

65.   Marti-Renom, M.A., A.C. Stuart, A. Fiser, R. Sanchez, F. Melo, and A. Sali. Comparative protein structure modeling of genes and genomes. *Annu Rev Biophys Biomol Struct* **29**: 291-325, 2000.

66.   Shen, M.Y. and A. Sali. Statistical potential for assessment and prediction of protein structures. *Protein Sci* **15**(11): 2507-2524, 2006.

67.   Gohlke, H. and D.A. Case. Converging free energy estimates: MM-PB(GB)SA studies on the protein-protein complex Ras-Raf. *Journal of Computational Chemistry* **25**(2): 238-250, 2004.

68.   Gohlke, H., C. Kiel, and D.A. Case. Insights into protein-protein binding by binding free energy calculation and free energy decomposition for the Ras-Raf and Ras-RaIGDS complexes. *Journal of Molecular Biology* **330**(4): 891-913, 2003.

69.   Nicholls, A. and B. Honig. A Rapid Finite-Difference Algorithm, Utilizing Successive over-Relaxation to Solve the Poisson-Boltzmann Equation. *Journal of Computational Chemistry* **12**(4): 435-445, 1991.

70. Onufriev, A., D. Bashford, and D.A. Case. Modification of the generalized Born model suitable for macromolecules. *Journal of Physical Chemistry B* **104**(15): 3712-3720, 2000.

71. Weiser, J., P.S. Shenkin, and W.C. Still. Approximate atomic surfaces from linear combinations of pairwise overlaps (LCPO). *Journal of Computational Chemistry* **20**(2): 217-230, 1999.

72. Bohne-Lang, A. and C.W. von der Lieth. GlyProt: in silico glycosylation of proteins. *Nucleic Acids Res* **33**(Web Server issue): W214-219, 2005.

73. Maenaka, K., P.A. van der Merwe, D.I. Stuart, E.Y. Jones, and P. Sondermann. The human low affinity Fcgamma receptors IIa, IIb, and III bind IgG with fast kinetics and distinct thermodynamic properties. *J Biol Chem* **276**(48): 44898-44904, 2001.

74. Carl, P., C.H. Kwok, G. Manderson, D.W. Speicher, and D.E. Discher. Forced unfolding modulated by disulfide bonds in the Ig domains of a cell adhesion molecule. *Proc Natl Acad Sci U S A* **98**(4): 1565-1570, 2001.

75. Wiita, A.P., S.R. Ainavarapu, H.H. Huang, and J.M. Fernandez. Force-dependent chemical kinetics of disulfide bond reduction observed with single-molecule techniques. *Proc Natl Acad Sci U S A* **103**(19): 7222-7227, 2006.

76. Wiita, A.P., R. Perez-Jimenez, K.A. Walther, F. Grater, B.J. Berne, A. Holmgren, J.M. Sanchez-Ruiz, and J.M. Fernandez. Probing the chemistry of thioredoxin catalysis with force. *Nature* **450**(7166): 124-127, 2007.

77. Bond, C.S. TopDraw: a sketchpad for protein structure topology cartoons. *Bioinformatics* **19**(2): 311-312, 2003.

78. Brahms, S. and J. Brahms. Determination of protein secondary structure in solution by vacuum ultraviolet circular dichroism. *J Mol Biol* **138**(2): 149-178, 1980.

79. Day, R., B.J. Bennion, S. Ham, and V. Daggett. Increasing temperature accelerates protein unfolding without changing the pathway of unfolding. *J Mol Biol* **322**(1): 189-203, 2002.

80. Gianni, S., N.R. Guydosh, F. Khan, T.D. Caldas, U. Mayor, G.W. White, M.L. DeMarco, V. Daggett, and A.R. Fersht. Unifying features in protein-folding mechanisms. *Proc Natl Acad Sci U S A* **100**(23): 13286-13291, 2003.

81. Beck, D.A. and V. Daggett. Methods for molecular dynamics simulations of protein folding/unfolding in solution. *Methods* **34**(1): 112-120, 2004.

82. Duan, J. and L. Nilsson. Thermal unfolding simulations of a multimeric protein--transition state and unfolding pathways. *Proteins* **59**(2): 170-182, 2005.

83. Chong, L.T., W.C. Swope, J.W. Pitera, and V.S. Pande. Kinetic computational alanine scanning: application to p53 oligomerization. *J Mol Biol* **357**(3): 1039-1049, 2006.

84. Walser, R., A.E. Mark, and W.F. van Gunsteren. On the temperature and pressure dependence of a range of properties of a type of water model commonly used in high-temperature protein unfolding simulations. *Biophys J* **78**(6): 2752-2760, 2000.

85. Shi, M., S.Y. Foo, S.M. Tan, E.P. Mitchell, S.K. Law, and J. Lescar. A structural hypothesis for the transition between bent and extended conformations of the leukocyte beta2 integrins. *J Biol Chem* **282**(41): 30198-30206, 2007.

86. Ferrara, C., F. Stuart, P. Sondermann, P. Brunker, and P. Umana. The carbohydrate at FcgammaRIIIa Asn-162. An element required for high affinity binding to non-fucosylated IgG glycoforms. *J Biol Chem* **281**(8): 5032-5036, 2006.

87. Gerhardt, S., G. Hassall, P. Hawtin, E. McCall, L. Flavell, C. Minshull, D. Hargreaves, A. Ting, R.A. Pauptit, A.E. Parker, and W.M. Abbott. Crystal structures of human ADAMTS-1 reveal a conserved catalytic domain and a disintegrin-like domain with a fold homologous to cysteine-rich domains. *J Mol Biol* **373**(4): 891-902, 2007.

88. Mosyak, L., K. Georgiadis, T. Shane, K. Svenson, T. Hebert, T. McDonagh, S. Mackie, S. Olland, L. Lin, X. Zhong, R. Kriz, E.L. Reifenberg, L.A. Collins-Racie, C. Corcoran, B. Freeman, R. Zollner, T. Marvell, M. Vera, P.E. Sum, E.R. Lavallie, M. Stahl, and W. Somers. Crystal structures of the two major aggrecan degrading enzymes, ADAMTS4 and ADAMTS5. *Protein Sci* **17**(1): 16-21, 2008.

89. Shieh, H.S., K.J. Mathis, J.M. Williams, R.L. Hills, J.F. Wiese, T.E. Benson, J.R. Kiefer, M.H. Marino, J.N. Carroll, J.W. Leone, A.M. Malfait, E.C. Arner, M.D. Tortorella, and A. Tomasselli. High resolution crystal structure of the catalytic domain of ADAMTS-5 (aggrecanase-2). *J Biol Chem* **283**(3): 1501-1507, 2008.

90. Matsumoto, A., T. Kamata, J. Takagi, K. Iwasaki, and K. Yura. Key interactions in integrin ectodomain responsible for global conformational change detected by elastic network normal-mode analysis. *Biophys J* **95**(6): 2895-2908, 2008.

91. Arnold, J.N., M.R. Wormald, R.B. Sim, P.M. Rudd, and R.A. Dwek. The impact of glycosylation on the biological function and structure of human immunoglobulins. *Annu Rev Immunol* **25**: 21-50, 2007.

92. Kaneko, Y., F. Nimmerjahn, and J.V. Ravetch. Anti-inflammatory activity of immunoglobulin G resulting from Fc sialylation. *Science* **313**(5787): 670-673, 2006.

93.   Lau, T.L., C. Kim, M.H. Ginsberg, and T.S. Ulmer. The structure of the integrin alphaIIbbeta3 transmembrane complex explains integrin transmembrane signalling. *EMBO J*, 2009.