

Computational modeling of fibronectin, integrin and their complexes: search for inhibitors of its interactions with integrins



Student:

Pompeo Marco Gaudiosi S279983

Supervisors:

Jack A. Tuszynski

Marco A. Deriu

Abstract

The interactions between extracellular matrix (ECM) fibronectin (FN) and various integrins (INs) through specific receptors are of fundamental importance in cellular response (Bachman et al., 2015), cell-cell interactions, immunological and inflammatory events and thrombus formation (Johansson et al., 1997). The integrin binding domains of FN have been widely investigated and they are represented by the Arg-Gly-Asp (RGD) domain (Johansson et al., 1997) and the nearby areas that help to extend the contact surface, represented by the Pro-His-Ser-Arg-Asn (PHSRN) domain (Leahy et al., 1996), found in the 9th and 10th type III FN modules. The ability to modulate the various aspects of cellular adhesion, motility and proliferation identified the FN-IN interactions as potential target for therapeutic treatments, especially in cancer and associated conditions (like thrombosis, angiogenesis, osteoporosis) (Sheldrake & Patterson, 2014). Some integrins types are particularly critical intermediaries in a wide spectrum of cancer abilities, especially angiogenesis (Reardon & Cheresch, 2011). In the recent years, computational modelling, analysis, and simulation offer the possibility to study and evaluate molecules in a completely simulated computational environment, avoiding expensive experiments and reducing the simulation time, particularly in drug development, where the number of compounds to test can range in millions. By offering these possibilities to test ligand affinity and analyzing compound docking, the task to isolate the most promising compounds that could make, for example, excellent inhibitors of the FN-IN interactions, is made much easier, and this study could be extended focusing on the interactions that promote and support cancer growth and maintaining. In this work crystallographic fragments of FN and a computational model of the whole protein, obtained during a previous group project with professor Tuszynski using M.O.E. (Molecular Operating Environment, <https://www.chemcomp.com/index.htm>) will be used to analyze the sites of interaction between FN and IN, focusing on the integrins involved in cancer processes and maintenance, and to perform docking and molecular dynamics simulations with different molecules to evaluate possible inhibitors of such interactions, to use as potential therapeutic drug candidates. The analysis on such mechanism could also elucidate the processes of these interactions and the conformational changes that follow, that will be compared to in vitro simulations and results.

The complete homology model of FN was obtained during a project work in the Rational Drug Design course at Politecnico of Torino, led by professor J. Tuszynski and M. Miceli. I'm grateful to all my colleagues that participated in the project work and helped obtaining the model used for this work:

Martina Birocco

Gabriele Giambattista

Eleonora Molinari

Sara Groppo

Alessandro Masoero

1. Introduction

1.1. Biological Background

1.1.1. Fibronectin

Structure and physiological importance

FN is a large glycoprotein almost ubiquitous in the human body, and it is found on cell surface, in connective tissue matrix and extracellular fluids (Sharma, 1999). First studies referred to a “cold-insoluble globulin” found in plasma at concentration of $330\mu\text{g/ml}$, with a total molecular weight (Mw) of around 450 kDa composed by two distinct chains (Mw of 215 and 220 kDa) covalently bound (Mosesson et al., 1975). In 1984, Mosher (Mosher, 1984) analyzed the many roles of FN both in soluble and insoluble forms. It is present as a polymeric fibrillar network in the ECM and as soluble protomers in body fluids. FN as protomer is organized in a V-shape dimer of two arms, each one of approximately 2000 amino acids (220 kDa) and about 2×60 nm dimension, with angle between the arms of 70° degree circa (Mosher, 1984). The dimers bind themselves with two disulphide bridges near the carboxyl terminus of the protein in antiparallel manner (Mosesson et al., 1975). Each monomer is composed by the repetition of three internal homologous repeating sequence elements, classified as type I, II or III repeats (FnI, FnII and FnIII)(Petersen et al., 1983). The precise combination of these modules is 12 FnI, 2 FnII, 15-17 FnIII, which account for 90% of FN sequence. FN is a good example of mosaic protein, in which the primary sequence is entirely consisting of these three repeating modules, except for the V segment. Several variants of FN are formed by alternative splicing of the pre-mRNA, generated by transcription of a single gene, at three positions: AFnIII between 11 and 12 FnIII, BFnIII between 7 and 8FnIII and the variable region V (or segment, as indicated before), subjected to a more complex pattern of splicing (Johansson et al., 1997). Thus the length of FN can span between 2156 and 2325 amino acids, depending on which internal splicing has taken place (Kornblihtt et al., 1985). Currently, 20 different human variants originated from a single gene have been identified. The three modules differ in length and in tertiary structure:

- FnI module consists of about 45 amino acids and is located in the amino-terminal and carboxy-terminal fibrin-binding domains and in the collagen-binding domain of FN. FnI module contains four cysteine residues, which form two disulphide bonds in a 1-3, 2-4 configuration. The only exception is made for 12FnI which has an extra disulphide bond (Potts & Campbell, 1994).
- FnII module contains approximately 60 residues. The central core of the module is composed of two double-stranded anti-parallel β -sheets oriented approximately perpendicularly to each other and two irregular loops, one separating the two β -sheets and the other between the two strands of the second β -sheet (Potts & Campbell, 1994).
- FnIII modules are made of approximately 90 amino acids, depending on the isoform. Unlike FnI and FnII modules, FnIII modules do not contain disulphide bonds, and most are encoded by two exons rather than one. The tertiary structure's dominant feature of FnIII modules is a ‘sandwich’ formed by two anti-parallel β -sheets enclosing a hydrophobic core (Potts & Campbell, 1994) Type III repeats of FN form $\alpha\beta$ -sandwich with four β -strands on one side and three on the other. The core structure of the type III repeats is conserved, with differences in function attributed to key residues, often located in the loop regions. Due to their abundance, these repeats are involved in a multitude of binding interactions. Thus, structural elements responsible for these interactions continue to be mapped. (Sharma, 1999).

FN interacts with different cell surface receptors and matrix components to modulate and assist cell adhesion, spreading, migration, and maintenance and remodeling of tissue and embryogenesis (Pickford et al., 2001). As a soluble dimer in plasma, FN is involved in blood coagulation (wound healing, hemostasis, and thrombosis) through its affinity for fibrin and platelets. As an insoluble network, in the form of fibrillar aggregate (diameter from nm to μm , highly interwoven) (Zollinger & Smith, 2017) and of dimer, FN interacts with cell surface receptors and with other matrix components such as collagens and proteoglycans, thus assisting cell adhesion, spreading and migration, maintenance and remodeling of tissue integrity and embryogenesis (Pickford et al., 2001). During embryogenesis, fibronectin appears before or at the onset of gastrulation in all vertebrates examined, and it is abundant at times and sites of cell migration. Alternative mRNA splicing is used during development as a mechanism to create different forms of fibronectin within the extracellular matrix by inclusion or exclusion of the AFnIII, BFnIII and V segments. The expression of integrins has also been demonstrated to be developmentally regulated, where some of the fibronectin receptor subunits are continuously expressed, while others are not. Injection of antibodies to fibronectin or RGD-containing peptides inhibits gastrulation in several species, indicating that the interactions between fibronectin and integrins are important during that particular stage of development. Mouse embryos lacking fibronectin die on embryonic day 8.5, and they have defects in the development of mesoderm, neural tube and blood vessels. Fibronectin-null blastocysts hatch and implant into the uterine wall, initiate gastrulation and form mesoderm in complete absence of embryo derived fibronectin. The presence of fibronectin in these early stages is believed to be important since the oocyte alone is probably contributing to enough maternal fibronectin for these processes. From embryonic day 8.0 and onwards the mutant embryos develop deformities and deteriorate during day 10-11 of gestation: show shortened anterior-posterior axes, fail to develop certain mesodermally derived structures like notochord and somites, and develop abnormal heart and blood vessels, all probably result of a deficit in the mesodermal layer. The lack of notochord and somites has later been shown to be a result of fibronectin being critical for the organization or maintenance of the notochord precursor cells and for the condensation of precursor cells into somites. Neural folds are formed in the absence of fibronectin, while the neural tube becomes kinked. Primitive red blood cells do develop in the fibronectin-null embryos, while blood vessels do not, strongly suggesting a role for fibronectin in vasculogenesis but not in hematopoiesis. Since there are several receptors for fibronectin, the total effect of the fibronectin-null mutation is likely to be made up of separate effects due to lack of binding between fibronectin and its individual receptors.

The broad range of functions possessed by FN is due to the fact that the three FN modules are organized into functional domains several binding sites, specific for ECM proteins (e.g. collagen), cell-surface receptors (e.g. integrins), circulating blood proteins (e.g. fibrin), glycosaminoglycans (e.g. heparin, chondroitin sulphate), signaling molecules and growth factors (Dalton & Lemmon, 2021). These interactions are responsible for cell mechanical and chemical cues that induce peculiar cell behaviors (e.g. differentiation, epithelial-mesenchymal transition) (Griggs et al., 2017), and their impairment promote scarring, tumorigenesis, fibrosis and developmental defects (Bae et al., 2013).

The main binding sites, obtained through proteolytic fragmentation or recombinant DNA analyses (Moore, 2001; Mosher, 1984), are here described:

- Collagen-binding domain is composed by 6FnI 1-2FnII 7-9FnI modules. It can be isolated as a 42 kDa fragment, particularly affine for heat-denatured collagen (gelatin). An analysis of the physiological state of collagen indicates that the triple helix is likely to unfold locally at body temperature, which suggests that this FN domain could be involved in interactions with native collagen in vivo (Pickford et al., 2001).

- Heparin-binding domains: there are three domains that interact with heparan sulfate proteoglycans. The high-affinity heparin domain, towards the C-terminal part (12-14 FnIII), can also bind to a widely distributed glycosaminoglycan (chondroitin sulfate). The weaker heparin-binding domain, towards the N-terminal end (1-5 FnI), contains a Staphylococcus-aureus-binding site. Hep3 is at the central part of the molecule (4-6 FnIII) and shows a very low affinity for heparin (Mezzenga & Mitsi, 2019; Pankov & Yamada, 2002). Heparin-binding activity in FN has been proposed to act with integrins to stimulate the formation of focal adhesions. In addition to the known heparin-binding capability of FN12–14, several reports have indicated that it may possess binding sites for integrins (Sharma, 1999).
- Fibrin-binding domains are mainly specific for Fibrin I and Fibrin II. The major site is in the N-terminal domain and is formed by 4-5 FnI modules. The interaction of FN with fibrin is important for cell adhesion or cell migration into fibrin clots. Moreover, the interaction of FN with fibrin may also be involved in macrophage clearance of fibrin from circulation after trauma or inflammation (Pankov & Yamada, 2002).
- Integrin-binding domains are represented primarily by the **Arg-Gly-Asp** motif (**RGD**), located in 10FnIII. This is one of the most important sites for integrin recognition, able to recognize different types of this receptor, such as: $\alpha 3\beta 1$, $\alpha 5\beta 1$, $\alpha 8\beta 1$, $\alpha V\beta 3$, $\alpha V\beta 6$, $\alpha IIb\beta 3$ (Johansson et al., 1997). Furthermore, it was discovered that also areas near RGD site help to the contact surface between integrins and ligands.
 Another site that works in synergy with RGD, important for integrins-binding is **Pro-His-Ser-Arg-Asn** domain (**PHSRN**), located in the crystal structure of 9FnIII (Leahy et al., 1996; Obara et al., 1988). RGD, in fact, for some integrins, such as $\alpha 5\beta 1$, requires PHSRN because this sequence interacts with the $\alpha 5$ subunit, promoting the binding of $\alpha 5\beta 1$ to FN. The distance between these two sites (RGD-PHSRN) is about 32 Å-55 Å, depending on the conformation to these two domains (Bachman et al., 2015). Besides the PHSRN sequence in 9FnIII, Aota et al. located an other peptide region, probably involved in the contribution to the adhesive capacity of FN (Aota -i. et al., 1991).
 The binding of FN to its integrin receptors is further complicated by the fact that integrin affinity and specificity can be modulated by events within the cell, tuning the interaction depending on the wide range of integrin interactions among cells (Hynes, 1992).

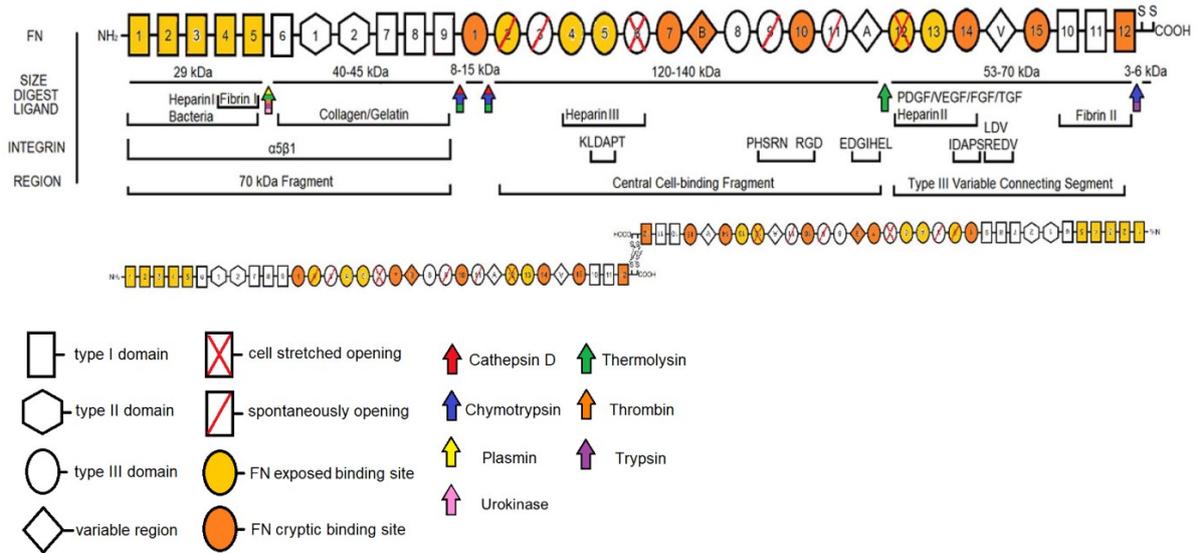


Figura 1: A schematic of the domains of FN with relevant structural features, cleavage sites, nomenclature, and integrin binding. FN is an approximately 250 kDa protein that is secreted as a dimer. Individual domains are classified as Type I (rectangles), Type II (hexagons), Type III (ovals), or a variable region (diamond). Domains that spontaneously open are shown with a single red slash, while domains that are mechanically unfolded are shown with a red X. Domains that have exposed FN–FN binding sites are shown in gold, while FN domains that have been shown to exhibit cryptic FN–FN binding sites are shown in orange. Molecular weight of regions is directly below. Enzymes known to digest FN are shown at their specific sites with arrows and color coded appropriately. Regions and/or specific sequences that have been shown to bind other ECM constituents are labeled based on size and ligand, then integrins are listed below that, and common terminology for each FN fragment is listed below that. The dimerization of FN at its C-terminus is shown at bottom (Dalton & Lemmon, 2021).

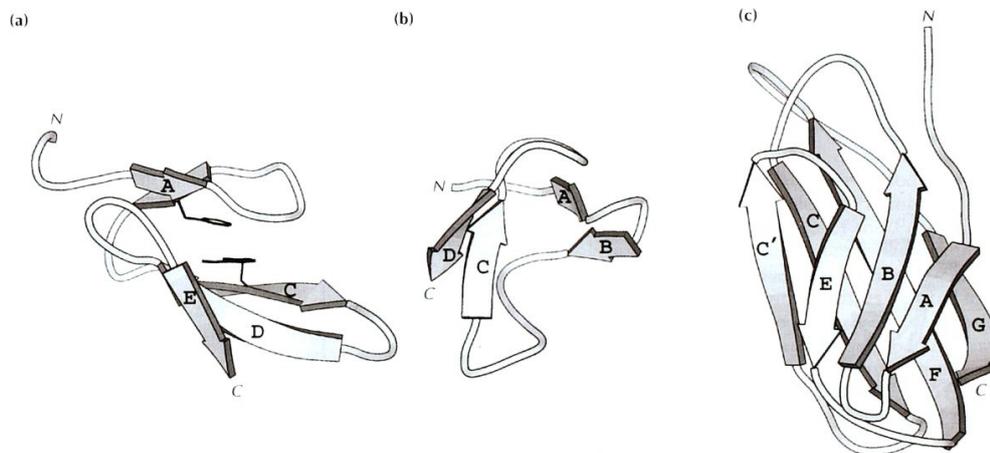


Figura 2: Ribbon diagram of typical structure of Fnl (a), FnII (b), FnIII (c) modules (Potts & Campbell, 1994), with β strands labelled.

Crystallography structures and analysis on PDB entries

Due to its many functions in the human body, building a computational model of FN is fundamental to better visualize the protein and study its physiological properties. Furthermore, in silico trials can be performed to mimic the interactions with other biological molecules or compounds. By now, various fragments of FN have been crystallized independently and can be found in the Protein Data Bank (PDB). However, many fragments of FN are still missing, and a model of the complete protein is not available, furthermore since FN is isolated usually in unphysiological condition (Johansson et al., 1997).

Various fragments of FN have been crystallized independently, but many are still missing and a complete crystallized protein model is absent. In our previous work, a complete homology model of FN was obtained by merging crystallized structures of FN fragments already obtained and homology models computed to simulate the missing fragments.

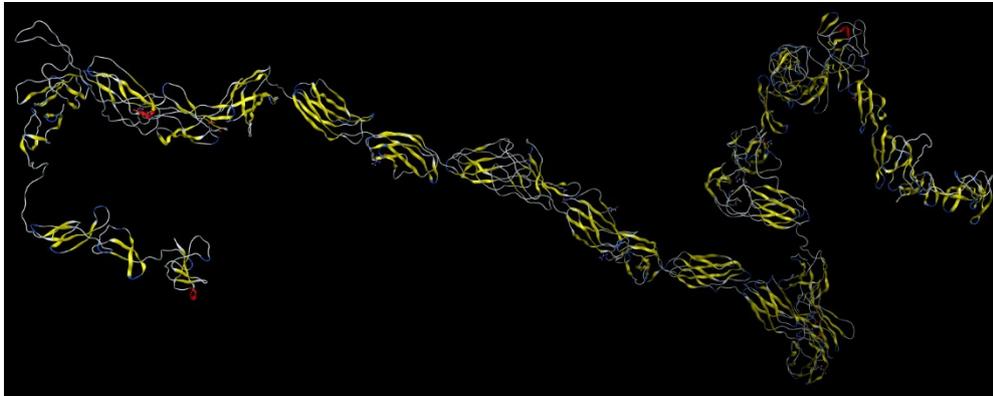


Figura 3: homology model of FN obtained on MOE using existing crystallized fragments and homology model of lacking structures.

The first 48 amino acid sequence of FN still doesn't have a crystallography structure in PDB, likely due to it being the beginning sequence of the protein and a very liable, irregular, and flexible region. Literature does not report any specific binding site in this area of the protein, more likely to be considered as non-structured fragment, as the linker between modules all along the protein. In our previous work, in order to create a usable model, the available structures of FN were aligned to the query sequence and the percentage identity (%ID) was evaluated. The PDB structure named 1Q38 showed a %ID of 24.2, the highest found, and therefore was used to model this specific fragment.

The first 5 different FnI modules are the sites of interaction for heparin, fibrin and also bacteria (Mezzenga & Mitsi, 2019). The query sequence in residues from 48 to 140 comprehends 1FnI-2FnI modules, corresponding to PDB entries 1O9A (Bingham et al., 2008) and 1QGB (Potts et al., 1999), while in residues 93 to 183 it comprehends the 2FnI-3FnI modules, found at PDB entries 2CG6, 2CG7, 2CKU, 2RKZ, 3CAL, 3ZRZ, 4PZ5.

Superposition of 2CG6, 2CG7, 2CKU, 2RKZ, 3CAL, 3ZRZ, 4PZ5 (**figure 4b**), reporting the structure of 2FnI and 3FnI modules, shows a similar situation. While 2FnI module appears substantially the same, 3FnI module structures differ from each other in terms of internal rotation. Entries 2RKZ, 3CAL, 3ZRZ, 4PZ5 refer to complex of FN and different proteins, while the unmatched superposition between

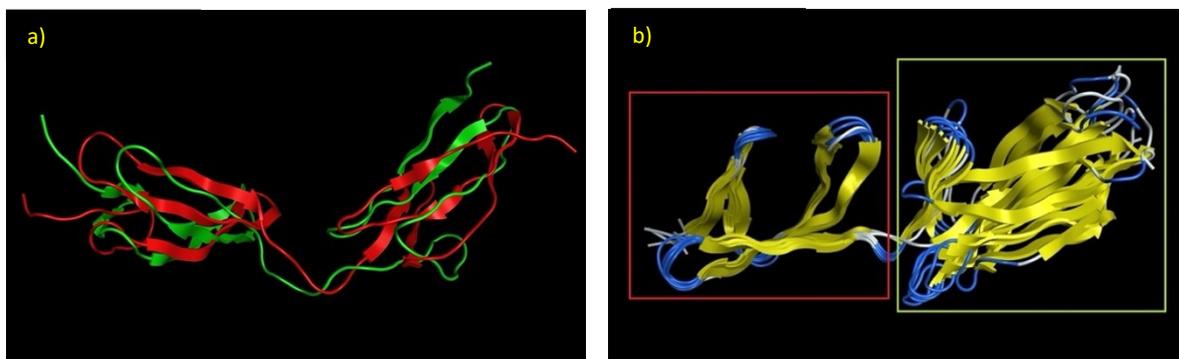


Figura 4: a) Superposition of 1O9A (red) and 1QGB (green) structures; b) 2CG6, 2CG7, 2CKU, 2RKZ, 3CAL 2F1 and 3F1 modules can be identified by the red and green boxes respectively.

2CG6 and 2CG7 structure is due to a different crystallization condition, resulting in a rotation of the

structure. As reported by the author (Rudiño-Piñera et al., 2007) structure 2CG7 is more consistent with NMR structure (2CKU) and, also, with superposition with 4FnI-5FnI modules (found at 1FBR PDB entry).

Superposition of 1O9A, 1QGB and 2CG7 structure (**figure 5a**) shows different relative position between modules, that FN can assume whether bound to other protein or not.

Another fragment from residue 183 to 275 contains the 4FnI-5FnI modules. The different crystallizations can be found on PDB at these entries: 1FBR, 2RKY, 2RL0. Superposition of these structures does not underline major differences in the folding. For reasons of resolution, the 2RKY entry was chosen as template for the model.

Since 2CG7 and 2RKY structure did not share any residues, a query for the gap was submitted to BLAST, and the best match was chosen as the one sharing the most residues with the query. PDB structure named 2C9E contained the amino acid sequence "PIAEK" which links the 3FnI and 4FnI modules as shown in the alignment in **figure 5b**. This structure was used to model the connection between these two modules.

No sequence is available in literature to model the residues from 275 to 297, that represent the linker between 5FnI and 6FnI. The homology model of this sequence was carried out by a search on MOE protein search tool of the query sequence, which resulted in a composite structure of 3GXE and 1E8B PDB entries with 47.4 %ID. This model was further used in the final model to predict relative position between 5FnI and 6FnI modules.

The sequences 3M7P, 1E88, 1E8B, 1Q06, 3MQL, 2FN2 (residues 297 to 604) identify the modules 6FnI 1-2FnII 7-9FnI. These modules are specific for the interaction with collagen or gelatin (Hynes & Yamada, 1982). After performing alignment, 3M7P was chosen, due to the very high-quality of chain.

1E88, 1E8B, 1Q06, 3MQL and 2FN2 describe properly only a portion of the gelatin binding site. Thus, having available the entire sequence of this binding site, they are all discharged.

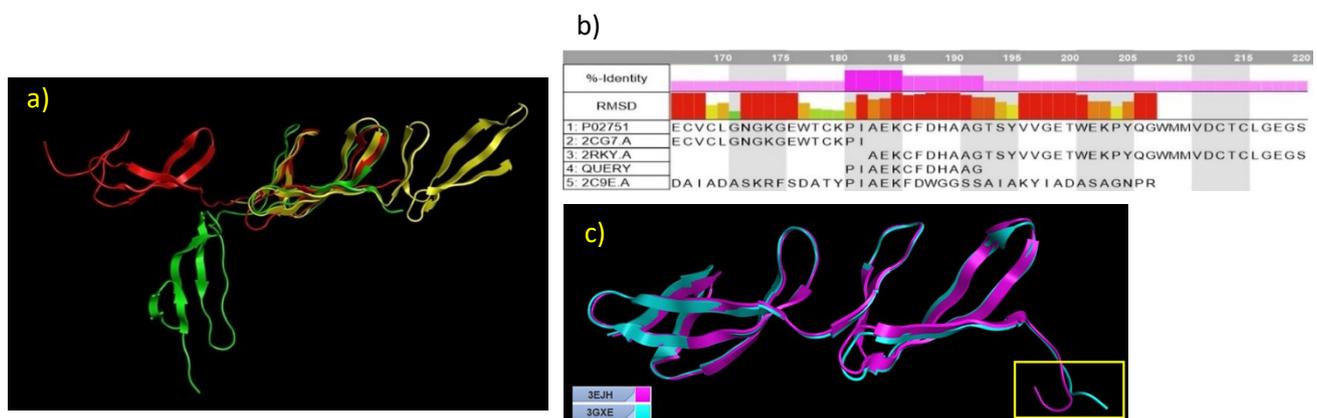


Figura 5: a) Superposition of 1O9A (red), 1QGB (green) and 2CG7(yellow); b) Alignment of 2CG7 and 2RKY along with query sequence and linker sequence 2C9E. RMSD and %ID are also shown above; c) Superposition of 1O9A (red), 1QGB (green) and 2CG7(yellow).

The sequences 3EJH, 3GXE (residues 516 to 608) identify the modules 8-9 FnI, which are specific for the interaction with collagen or gelatin (Hynes & Yamada, 1982). 3EJH.A is selected for alignment, considering the best resolution of the sequence (2.10 Å). After alignment, percentage of identity and superposition are performed, giving a 100%ID and an almost total superposition of the entries. The superposition highlights a different conformation of the last four residues of the sequence due to the

different binding: 3EJH (pink) interacts with collagen and 3GXE (light blue) interacts with low affinity collagen (**figure 5c**).

Depending on the isoform, between 15 and 17 FnIII modules can be found, each one of approximately 90 amino acids of length (Potts et al., 1999). Residues 608 to 1268 of P02751 FN sequence include repeats 1FnIII to 7FnIII, none of these fragments forming any binding site with other proteins. 1FnIII can be found at PDB entry 1OWW and 2HA1, the latter also includes 2FnIII module (figure 6).

The 1OWW entry includes residues 608-701 and displays a higher quality of chain than 2HA1; for this reason, 1OWW was used to improve the 1FnIII module already present in 2HA1 structure.

Repeats 3FnIII to 7FnIII are found at PDB entries 2NK1 (3FnIII) and 6MFA (4FnIII-7FnIII), both displaying a high resolution and exceptional quality of chain, with 6MFA being one of the largest fragments viable containing 369 fragments. Other PDB entries covering 608 to 1268 residues were 1Q38, 5J6Z, 2MNU, 5N48 and 6MSV but due to their lower quality of chain, length, or resolution, all these sequences were discarded. Alignment of P02751 FN sequence and the chosen structures was performed.

The superposition of 2HA1 and 1OWW fragments is shown (**figure 6a**): 1FnIII and 2FnIII are clearly visible as conglomerations of β -strands joined by a long loop. There are no main differences between the two chains, apart for the first few residues of 1OWW that diverge from 2HA1 loops. Also, PDB entries 2H41 and 2H45 include residues 721-809, both being different mutations of the 2FnIII repeat. These sequences have been aligned to the sequence of Fibronectin and superposed with 2HA1. The two sequences differ from 2HA1, since the loop modelling the first few residues is rotated in opposite directions. 2H41 and 2H45 were discarded (**figure 6b**).

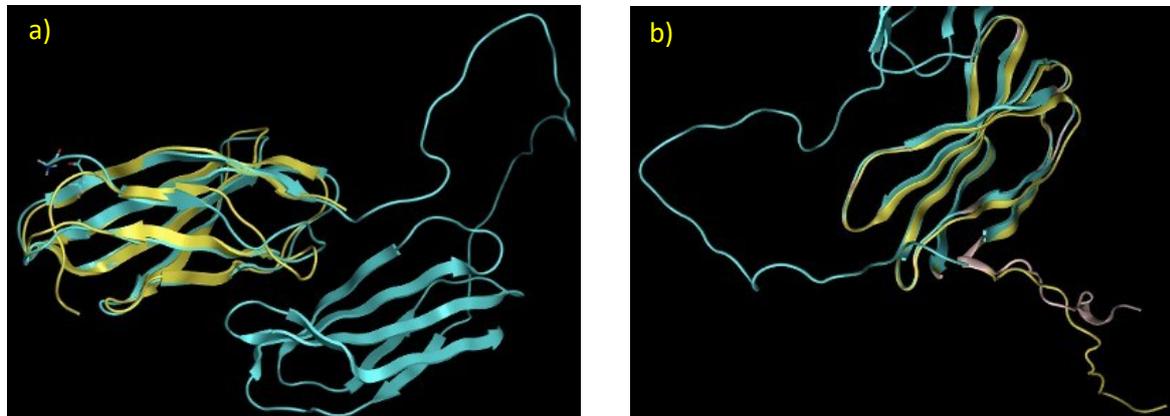


Figura 6: 1OWW with the chosen fragment 2HA1 and mutations of this fragment 2H41 and 2H45, the opposite rotation of the first few residues is shown.

The superposition of 2NK1 and 2HA1 differs from previous superpositions since the two chains share only two residues, that are modelled differently: one as β strand, the other as loop (**figure 7a**).

Fragments 2NK1 and 6MFA, were aligned and superposed, as previously; the two chains share few fragments, but in this case the two are modelled as loops. Structure 6MFA clearly displays 4th to 7th repeats of FnIII, as groups of β strands joined by loops (**figure 7b**).

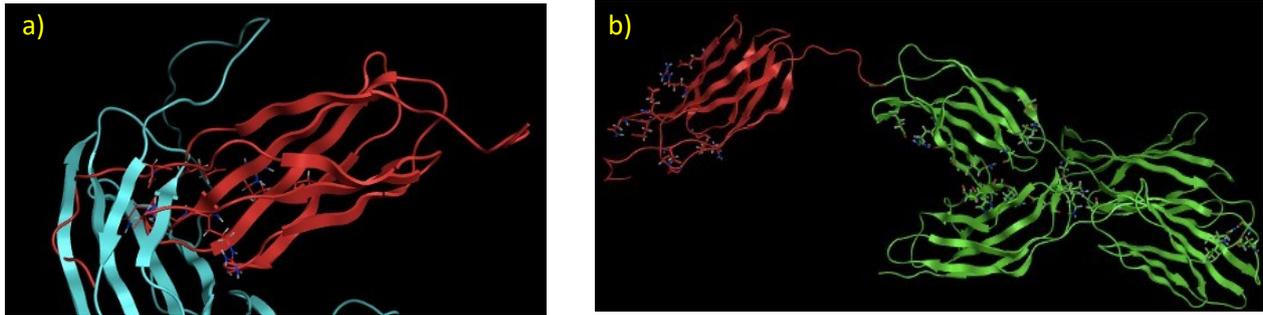


Figura 7: a) Superposition of 2NK1 and 2HA1 fragments showing the difference in the secondary structure of the residues of the connection between fragments; b) Superposition of 2NK1 and 6MFA fragments, the two share a few residues with the same secondary structure.

The 7FnIII, the Extra-Domain B (EDB), 8FnIII and 9FnIII modules are represented by the following sequences: 1FNF, 3T1W, 4GH7, 7NWL, 5N47, 2NFB, 2GEE. Despite the better resolution of 1FNF, 3T1W was selected, due to its overall better-quality of chain (Schiefner et al., 2012). 4GH7, 7NWL and 5N47 were discarded not only because of their lower resolution, but also for their quality chain, that was equal or inferior to 3T1W (Schiefner et al., 2012, 2018; Schumacher et al., 2021). Structures 2NFB and 2GEE represent residues 1266 to 1356 and 1266 to 1477; because the residues are already represented in 3T1W, that presents better resolution, these sequences were discarded.

The 9FnIII and 10FnIII modules represent the integrin-binding domains and corresponding residues are represented by 4LXO, 5DC9, 1TTF and 1TTG PDB entries. To model the integrin binding site 4LXO was selected: it presents better resolution and quality chain than 1TTF and 1TTG. Despite the high resolution and the better quality of chain, 5DC9 was discarded, because it represents the crystal structure of the AS25 monobody, that share a high %ID with the 10FnIII sequence (Wojcik et al., 2016).

Thus, (Schiefner et al., 2012)(Schiefner et al., 2012, 2018; Schumacher et al., 2021)(Wojcik et al., 2016)the selected sequences are 3T1W and 4LXO, that overlap over the 9FnIII module.

3T1W contains an x-ray structural analysis of EDB in the context of its neighboring domains 7FnIII, 8FnIII and 9FnIII. 4LXO contains the Crystal structure of 9, 10 FnIII-elegantin chimera, used to evaluate the binding of Integrin $\alpha 5$ - $\beta 1$ with 9FnIII and 10FnIII. It presents 13 discrepancies from the reference sequences UNP-P02751 (**figure 8**). From the sequence editor the discrepancies of 4LXO are visible in the aforementioned residues, while 3T1W differs only in the first residues from UNP-P02751.

The superposition of 3T1W and 4LXO results optimal, as can be seen in **figure 9**.

1: 3T1W	1435	1440	1445	1450	1455	1460	1465	1470	1475	1480	1485
2: 4LXO.A											
1: 3T1W	1490	1495	1500	1505	1510	1515	1520	1525	1530	1535	1540
2: 4LXO.A											
1: 3T1W	1545	1550	1555	1560	1565	1570	1575	1580	1585	1590	1595
2: 4LXO.A											
1: 3T1W	1600	1605	1610	1615	1620	1625	1630	1635	1640	1645	1650
2: 4LXO.A											

Figura 8: Discrepancies in the residue sequence of 4LXO compared to UNP-P02751.

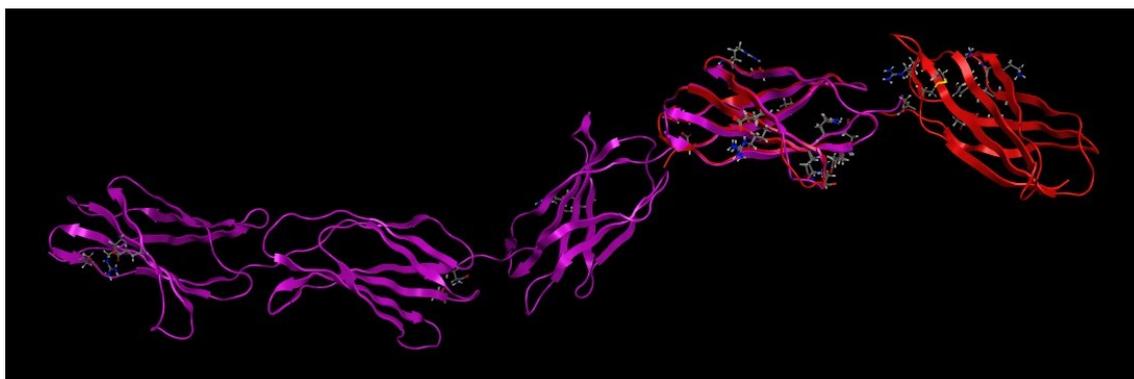


Figura 9: Alignment and superposition of 3T1W.A and 4LX0, focused on the shared residues 1448-1539.

To further optimize the representation of the 10FnIII module, the sequence 4MMX, 4MMY, 4MMZ, 5DC4 and 6NAJ are analyzed. 4MMX was chosen, corresponding to the coordinates and structure factor of integrin α V- β 3 bound to wtFnI0 (10th type III RGD-domain of wild-type fibronectin) (Van Agthoven et al., 2014a). It has only 5 discrepancies between the modelled and the reference sequence, but the quality of chain is not exceedingly high. There are other PDB entries, 4XXY and 4MMZ, that cover the same residues, but they have been discarded because they are fragments of a mutant form of FN.

Fragments 6XAX and 6XAY both describe the 10FnIII, 11FnIII, Extra-domain A and 12FnIII modules. 6XAY presents more residues than 6XAX, but is a mutant form and presents a missing part in the center of the sequence compared to P02751 FN FASTA. Despite that, there are only 6 discrepancies between the modelled and the reference sequence. To better represent the structure corresponding to these fragments, both are chosen to create an improved homology modelling of the corresponding sequence. The two entries have been aligned and superposed (**figure 10a**).

There are more fragments that represent the residues corresponding to the 10FnIII module, as 5DC0, 5J7C and 1FNA but due to these residues being already represented in other more suitable entries, they are all discarded. Entries 5DFT and 1J8K represent the 11FnIII module and, being already represented in 6XAX and 6XAY with better resolution, they are both discarded. Entries 1FNH and 3R8Q both represent 12FnIII, 13FnIII and 14FnIII modules, from residues 1812 to 2082. Following the described selection criteria, 3R8Q is chosen as structure to perform homology model, due to its higher resolution compared to 1FNH (2.40Å vs 2.80 Å).

In MOE, alignment and superpose of these two structures have been performed, showing that 1FNA and 3R8Q present a relatively wide overlap (**figure 10b**).

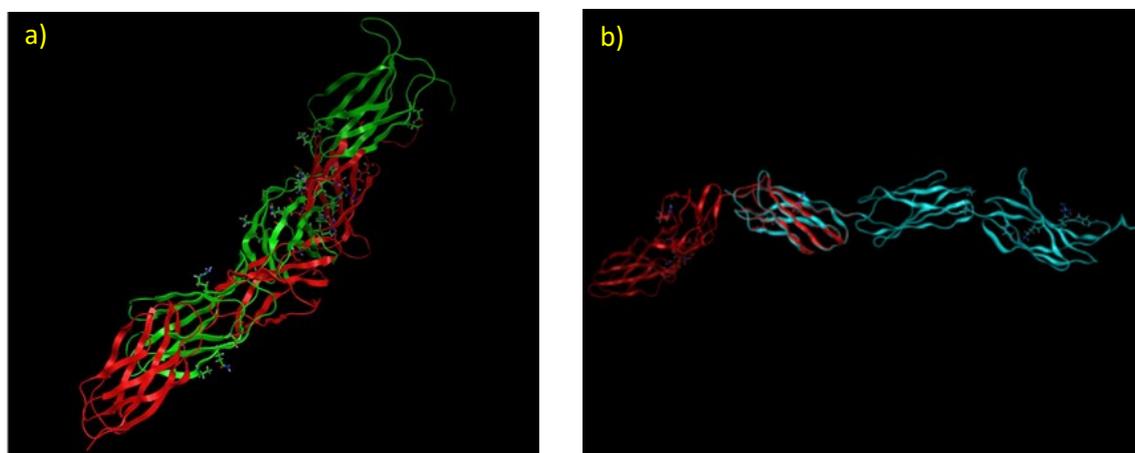


Figura 10: a) alignment and superposition of 6XAX.B (green) and 6XAY.D (red) on UNP-P02751; b) Superposition between 6XAX (light blue) and 3R8Q (red)

Literature and Data Banks do not provide information about the crystallization of the variable region between 14 and 15 FnIII modules. Thus, FN sequence was cut to recreate a query sequence for homology modelling. Residues from 2077 to 2203 were considered: this sequence consists of 116 residues of the variable region to which 5 residues at both ends were added. Several experimental attempts were performed to obtain a homology model of this fragment, selecting different FnIII modules based on the highest %ID (**table 1**).

Sequence used as template	Modules of the sequence	Number of residues	%ID between FASTA of the missing residues and template
3R8Q	12,13,14 FnIII	270	4%
1FNH	12,13,14 FnIII	270	8,7%
6XAX	11 FnIII, EDA,12 FnIII	273	5,6%
6XAY	10,11,12 FnIII	365	11,1%
4LX0	9,10 FnIII	183	6,3%
2GEE	8 FnIII	181	7,9%

Table 1: the entry with the highest percentage of identity (6XAY) is chosen to perform the homology modeling of V-region.

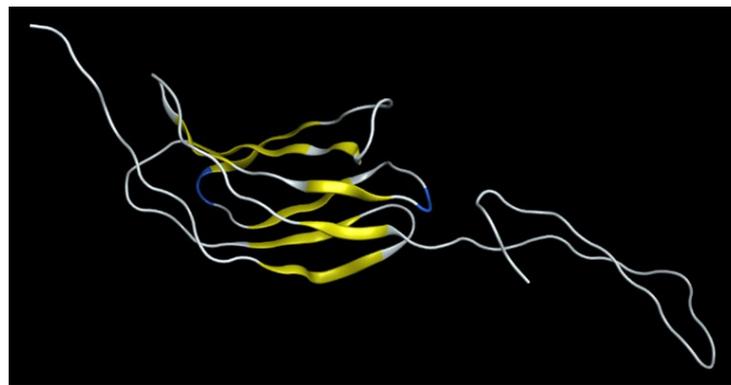


Figure 11: homology model from 2078th to 2203rd residue (variable region).

Due to the highest percentage of identity with the query sequence (11,1%), 6XAY was chosen as template for modelling the missing fragment (**figure 11**).

The module 15FnIII is represented with the 5M0A entry, the only reported one for 2199 to 2284 residues.

No x-ray or NMR crystallography of the 10FnI module was found in literature. For this reason, it has been chosen to select the structure of a homologous FnI module to perform homology modelling (**table 2**).

Sequence used as template	Modules of the sequence	Number of residues	%ID between FASTA of the missing residues and template
3M7P	6FI-1FII-2FII-7FI-8FI-9FI	307	6.9%
2EC3	11 FNI	60	8,8%
2CKU	2FI-3FI	89	1,7%

Table 2: the entry with the highest percentage of identity (2EC3) is chosen to perform the homology modeling.

PDB entry 2EC3.A (11-Fnl original module) was selected as template of residues from 2273 to 2330 due to its 8.8 %ID and its quality of chain (**figure 12a**).

The obtained structure will represent the missing 10Fnl module in this gap (**figure 12b**).

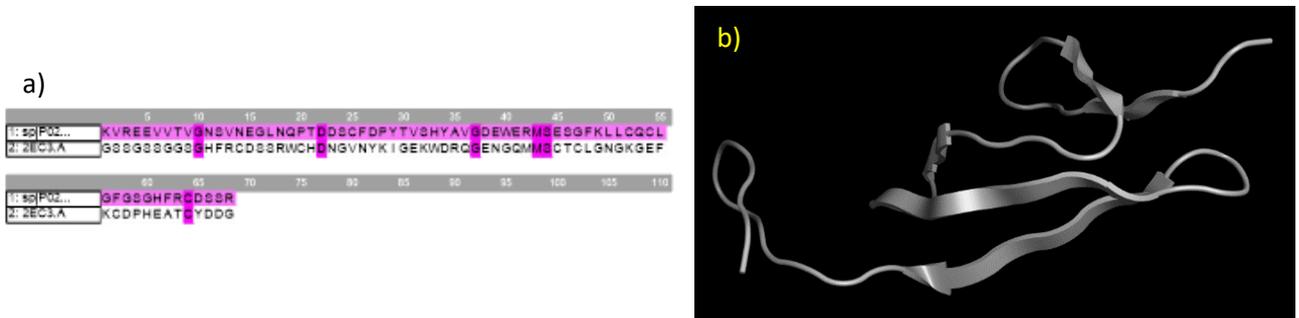


Figura 12: a) Alignment of the missing residues from 2284 to 2339 and 2EC3.A; b) visual rendering in new cartoon of the homology model for the 10-Fnl module

The 11Fnl module is represented by 2EC3.A, as it is the only entry representing residues from 2330-2390.

The terminal part of FN molecule consists of the 12Fnl module, followed by the two disulphide bonds and the COOH terminal. The two disulphide bonds represent the binding site of the two chains of the dimer of FN.

Figure 21: Alignment of the missing residues from 2284 to 2339 and 2EC3.A

Literature and Data Banks do not provide any crystallized structure of these final residues;

thus, FN sequence was again cut to form a query sequence to model this gap. Entries 3M7P.A, 1E8B.A, 1E88.A were chosen as candidates to represent residues from 2379 to 2477 (**table 3**).

Sequence used as template	Modules of the sequence	Number of residues	%ID between FASTA of the missing residues and template
3M7P	6FI-1FII-2FII-7FI-8FI-9FI	307	4,1%
1E8B	6FNI, 1-2FNII	305	9,2%
1E88	6FNI, 1-2FNII	305	9,2%

Table 3: the entry with the highest percentage of identity (1E88) is chosen to perform the homology modeling.

PDB entry 1E88.A was selected, due to its highest %ID with the gap and quality of chain (3M7P.A=4.1 ID%, 1E8B=9.2 ID%, 1E88=9.2 ID%) (**figure 13a-b**).

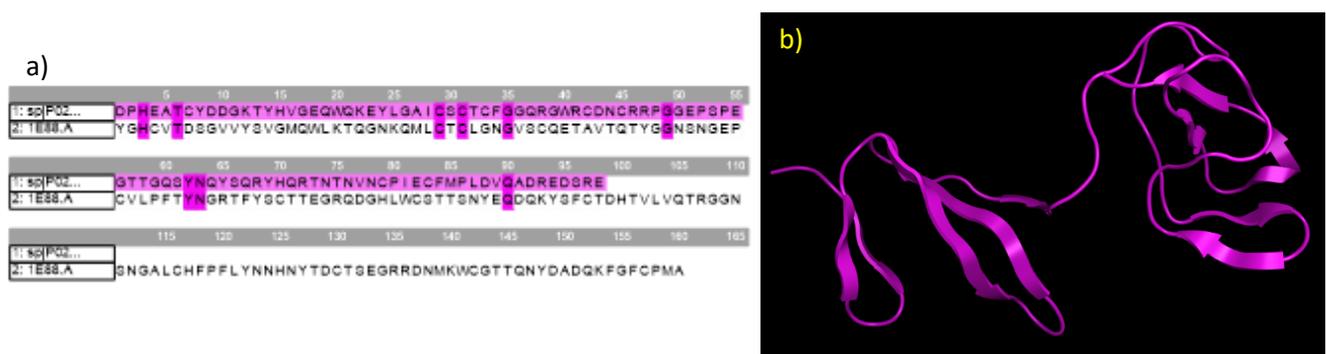


Figura 13: a) Alignment of 1E88.A with the query sequence of missing residues; b) visual rendering in new cartoon of the homology model for the 12-Fnl module.

The 10FnIII module, being the best main protagonist of FN-IN interactions, is analyzed more in depth: from analysis on the 10FnIII module (Main et al., 1992) it can be observed how the module is composed by seven β strands, that assemble in a “sandwich” of two antiparallel β sheets. One containing 2 strands (ABE) and the other one containing four strands (C’CFG).

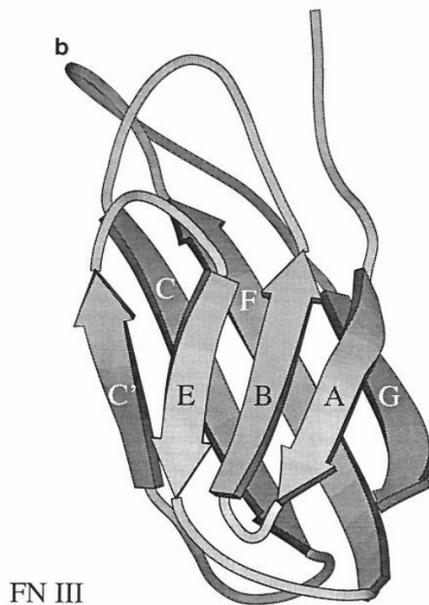


Figura 14: MOLSCRIPT diagram of type III module, with β strands annotated (Kraulis, 1991).

The triple stranded β sheet consists of residues Glu9-Thr14 (A), Ser17-Asp23 (B), and Thr56-Ser60 (E). There is a “classic” β bulge between Val11, Ala12 and Leu19 (Richardson et al., 1978), usually occurring between a pair of closely spaced hydrogen bonds, formed from HN and CO of Leu19 and Val11 in approximately α -helical conformation (ϕ -80° , ψ -45°) and Ala12 with an approximately normal β sheet conformation (ϕ -180° , ψ $+165^\circ$). The turn between strands A and B is well defined with a 2:2 turn, seemingly a distorted type I β turn. The four-stranded β sheet consists of residues Tyr31-Glu38 (C), Gln46-Pro51 (C’), Val66-Thr76 (F), and Ile88-Thr94 (G). Both the β sheets have a right-handed twist and they stack on top of each other to enclose a hydrophobic core (Main et al., 1992).

Analyzing the conserved residues in the type III module family, it can be observed how the majority of the conserved residues contribute to the hydrophobic core, with the invariant hydrophobic residues Trp22 and Tyr68 lying toward the N-terminal and C-terminal ends of the core respectively.

Other module types often contain highly conserved Gly residues, that are necessary for the formation of certain types of β turns (Wilmot & Thornton, 1990). Type III modules contain only one tight β turn that does not require Gly to form. Another case where the residues may be conserved for structural purposes are in the case of Pro25, the loop between strands E and F, and the Pro near the N-terminus, that may help in the correct formation of interfaces between sheets or modules (Main et al., 1992). The connection between strands E and F is a conserved five-membered loop in all type III modules, with the first and last residue showing a Gly preference. In a similar way, also the turn between strands A and B is of consistent length.

All the remaining loops in the module are all highly variable in length, with the insertion of the RGDS sequence in the F-G loop being particularly important.

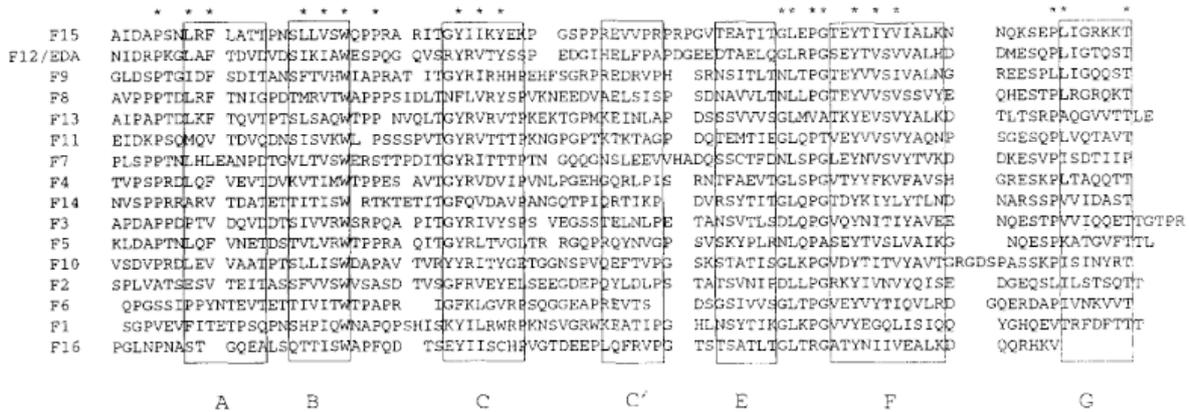


Figura 15: Sequence alignment of the type III modules of human with labelled strands FN. F12 corresponds to the ED-A sequence, which is not always present as a consequence of alternative splicing of mRNA. Asterisks indicate the most conserved residues.

In the work of Main and Harvey, the NOE distribution shows how the loops between strand C and C' and F and G present long range NOEs and thus high RMSDs, suggesting that these loops are conformationally labile; on top of that the heteronuclear NOEs for residues in these loops are significantly smaller than the rest of the molecule, suggesting again considerable conformational flexibility. Heteronuclear NOEs for residues Gly79, Asp80, Ser81, Lys83 and Ser84 are all of similar magnitude, implying that they undergo similar amplitudes of motion, possibly a conformational equilibrium with hinge motion rather than complete disordered motion. The β strands are much less flexible and seem to provide a rigid framework upon which functional, flexible loops are built, as suggested by all the studies. The heteronuclear NOE experiment described in Main and Harvey work provides direct evidence that the RGD-loops may be flexible and not in a specific conformation. The RGD motifs of potent integrin inhibitors also lie at the apex of solvent-exposed, flexible loops suggests that they may be responsible for fast recognition and fitting to the receptor. As already stated above, the RGD sequence alone does not account for the full cell adhesive properties of FN and are required additional synergistic regions for full activity, located in 8FnIII and 9FnIII. By internal deletion, deleting the region at the center of 9FnIII resulted in the greatest activity loss. By homology with 10FnIII, this region corresponds to the C and C' strands and the turn between them. By alignment of the modules, it seems that the CC' loop of 9FnIII is of the same length of 10FnIII, thus it probably possesses the same dynamic properties and flexibility. This loop may lie at the domain-domain interface between 9FnIII and 10FnIII and is a viable candidate for interaction with either the RGD loop or the integrin, depending on the exact nature of the module-module orientation. It was thought that the CC' loops was responsible for the synergy observed, but as already discussed this synergy site lies in the C'E loop, in the PHSRN sequence residues (Leahy et al., 1996).

From the work of de Vos et al. it results that the two adjacent type III modules participate in the interaction between FN and its integrin receptor in a way similar done in the observation of structurally related cytokine receptor modules (de Vos, 1993).

Assembling and evaluating the FN homology model

The entire FN molecule was assembled starting from the known structures and homology models obtained from the missing sequences. Every overlapping sequence was merged, two fragments at a time, by MOE homology modelling tool using FN sequence (P02751) as blank sequence; the first structure was used as template and second one as template override. The generalized Born/volume

integral (GB/VI) score was used for models' scoring and the one with the lowest value was chosen as final model of the merged fragment. Finally, once the total structure was obtained, the model was evaluated by geometrical analyses on MOE including Ramachandran plot, backbone bond length and angles, dihedral angles, rotamers, atom clashes and contact energy. Also, Z-score and local quality were measured with Protein Structure Analysis ([ProSA](#)), Whatcheck and ERRAT, found on [SAVES](#) server. All the structures available are listed in the table below (**table 4**); green highlighted entries are selected to assemble the fibronectin model.

Identifier Fn	Method	Resolution (Å)	Position	module (Fn I-II-III)	DOI
1O9A	NMR	-	48-140		http://doi.org/10.2210/pdb1O9A/pdb
1QGB	NMR	-	48-140	1-2FnI	http://doi.org/10.2210/pdb1QGB/pdb
2CG6	X-ray	1.55	93-182		http://doi.org/10.2210/pdb2CG6/pdb
2CG7	X-ray	1.20 Å	93-182	2-3FnI	http://doi.org/10.2210/pdb2CG7/pdb
2CKU	NMR	-	93-182		http://doi.org/10.2210/pdb2CKU/pdb
2RKZ	X-ray	2.00 Å	93-182		http://doi.org/10.2210/pdb2RKZ/pdb
3CAL	X-ray	1.70 Å	93-182		http://doi.org/10.2210/pdb3CAL/pdb
3ZRZ	X-ray	1.70 Å	93-182		http://doi.org/10.2210/pdb3ZRZ/pdb
4PZ5	X-ray	1.96 Å	93-182		http://doi.org/10.2210/pdb4PZ5/pdb
1FBR	NMR	-	183-275		http://doi.org/10.2210/pdb1FBR/pdb
2RKY	X-ray	1.80 Å	183-275	4-5 FnI	http://doi.org/10.2210/pdb2RKY/pdb
2RLO	X-ray	2.00 Å	184-272		http://doi.org/10.2210/pdb2RLO/pdb
3M7P	X-ray	2.50 Å	297-604	6FnI-1FnII-2FnII-7-9FnI	http://doi.org/10.2210/pdb3M7P/pdb
1E88	NMR	-	305-464		http://doi.org/10.2210/pdb1E88/pdb
1E8B	NMR	-	305-464		http://dx.doi.org/10.2210/pdb1e8b/pdb
1QO6	NMR	-	305-405		http://dx.doi.org/10.2210/pdb1qo6/pdb
3MQL	X-ray	3.00 Å	308-515		http://doi.org/10.2210/pdb3MQL/pdb
2FN2	NMR	-	406-464		http://doi.org/10.2210/pdb2FN2/pdb
3EJH	X-ray	2.10 Å	516-608	8-9 FnI	http://doi.org/10.2210/pdb3EJH/pdb
3GXE	X-ray	2.60 Å	516-608		http://doi.org/10.2210/pdb3GXE/pdb
1OWW	NMR	-	608-701	1FnIII	http://doi.org/10.2210/pdb1OWW/pdb
2HA1	NMR	-	609-809	1FnIII-2FnIII	http://doi.org/10.2210/pdb2HA1/pdb
1Q38	NMR	-	631-705		http://doi.org/10.2210/pdb1Q38/pdb

2H41	NMR	-	721-809		http://doi.org/10.2210/pdb2H41/pdb
2H45	NMR	-	721-809		http://doi.org/10.2210/pdb2H45/pdb
5J6Z	NMR	-	805-834		http://doi.org/10.2210/pdb5J6Z/pdb
2N1K	NMR	-	808-905	<i>3FnIII</i>	http://doi.org/10.2210/pdb2N1K/pdb
6MFA	X-ray	1.75 Å	903-1268	<i>4FnIII-5FnIII-6FnIII-7FnIII</i>	http://doi.org/10.2210/pdb6MFA/pdb
2MNU	NMR	-	907-995		http://doi.org/10.2210/pdb2MNU/pdb
5N48	X-ray	1.60 Å	907-995		http://dx.doi.org/10.2210/pdb5n48/pdb
6MSV	X-ray	2.40 Å	1085-1173		http://doi.org/10.2210/pdb6MSV/pdb
1FNF	X-ray	2.00 Å	1173-1265		http://doi.org/10.2210/pdb1FNF/pdb
3T1W	X-ray	2.40 Å	1173-1539	<i>7FnIII-EDB,8-9FnIII</i>	http://doi.org/10.2210/pdb3T1W/pdb
4GH7	X-ray	2.60 Å	1173-1447		http://doi.org/10.2210/pdb4GH7/pdb
7NWL	EM	3.10 Å	1173-1631		http://doi.org/10.2210/pdb7NWL/pdb
5N47	X-ray	3.00 Å	1173-1456		http://doi.org/10.2210/pdb5N47/pdb
2FNB	NMR	-	1266-1356		http://doi.org/10.2210/pdb2FNB/pdb
2GEE	X-ray	2.01 Å	1266-1447		http://doi.org/10.2210/pdb2GEE/pdb
4LXO	X-ray	1.42 Å	1448-1631	<i>9-10 FnIII</i>	http://doi.org/10.2210/pdb4LXO/pdb
5DC9	X-ray	1.56 Å	1537-1631		http://doi.org/10.2210/pdb5DC9/pdb
1TTF	NMR	-	1538-1631		http://doi.org/10.2210/pdb1TTF/pdb
1TTG	NMR	-	1538-1631		http://doi.org/10.2210/pdb1TTG/pdb
2CK2	X-ray	2.00 Å	1538-1633		http://doi.org/10.2210/pdb2CK2/pdb
6XAY	X-ray	2.48 Å	1538-1903	<i>10-12 FnIII</i>	http://doi.org/10.2210/pdb6XAY/pdb
2OCF	X-ray	2.95 Å	1539-1631		http://doi.org/10.2210/pdb2OCF/pdb
4JE4	X-ray	2.31 Å	1539-1631		http://doi.org/10.2210/pdb4JE4/pdb
4JEG	X-ray	2.30 Å	1539-1631		http://doi.org/10.2210/pdb4JEG/pdb
4MMX	X-ray	3.32 Å	1539-1631	<i>10 FnIII</i>	http://doi.org/10.2210/pdb4MMX/pdb
4MMY	X-ray	3.18 Å	1539-1631		http://doi.org/10.2210/pdb4MMY/pdb
4MMZ	X-ray	3.10 Å	1539-1629		http://doi.org/10.2210/pdb4MMZ/pdb
5DC4	X-ray	1.48 Å	1539-1631		http://doi.org/10.2210/pdb5DC4/pdb

6NAJ	X-ray	3.10 Å	1539-1629		http://doi.org/10.2210/pdb6NAJ/pdb
5DC0	X-ray	2.23 Å	1540-1631		http://doi.org/10.2210/pdb5DC0/pdb
5J7C	X-ray	2.54 Å	1540-1631		http://doi.org/10.2210/pdb5J7C/pdb
1FNA	X-ray	1.80 Å	1543-1633	10-FnIII	http://doi.org/10.2210/pdb1FNA/pdb
6XAX	X-ray	2.40 Å	1630-1903	11-FnIII-EDA-12-FnIII	http://doi.org/10.2210/pdb6XAX/pdb
5DFT	X-ray	2.50 Å	1638-1726		http://doi.org/10.2210/pdb5DFT/pdb
1J8K	NMR	-	1722-1815		http://doi.org/10.2210/pdb1J8K/pdb
1FNH	X-ray	2.80	1812-2082		http://doi.org/10.2210/pdb1FNH/pdb
3R8Q	X-ray	2.40 Å	1812-2082	12-14FnIII	http://doi.org/10.2210/pdb3R8Q/pdb
6HNF	NMR	-	1995-2082		http://doi.org/10.2210/pdb6HNF/pdb
5MOA	NMR	-	2199-2284	15 FnIII	http://doi.org/10.2210/pdb5MOA/pdb
2EC3	NMR	-	2330-2390	11 FnI	http://doi.org/10.2210/pdb2EC3/pdb

To compute the entire fibronectin molecular structure, the chosen PDB entries and the homology models obtained to represent each FN module were merged with the help of the homology modelling tool, found in MOE. The template structure and override structure were chosen performing two trials with both combinations, selecting the result with best GB/VI score.

The homology model of the first 48 residues, obtained using the 1Q38.A structure, was linked to 1O9A.A, using the first as template override. In this way the homology model of the fibronectin was obtained from residue 1 to 140 (HM 1-140). This model was linked to 2CG7.A using the same procedure described before, choosing the homology model as template override. Thus the HM of the protein from residue 1 to 182 was obtained (HM 1-182).

In the end, the obtained linked chains were connected, subsequently to 2C9E, 2RKY.A and the 3GXE-1E88 hybrid chain, always using the resulting homology model as template override. The resulting homology model represents residues from 1 to 285 (HM 1-285).

Entries 3M7P and 3EJH were chosen to build the homology model of residues 279 to 608. Firstly, the superposition of 3M7P and 3EJH were obtained, showing the not perfect overlap of the sequences, even if they share 89 residues, this could be due to their different resolutions (2.50 Å and 2.10 Å respectively). Then 3M7P was used as a template and 3EJH is used as template override, obtaining the homology model of residues 279 to 608 (HM 279-608) (**figure 16a**).

Regarding the homology model of residues 608-809 (HM 608-809), the 1OWW entry displays both a high resolution and a higher quality of chain than 2HA1, for this reason 1OWW was used to improve the 1FnIII module already present in 2HA1 structure.

To model residues from 609 to 905 (HM 609-905), 2HA1 was used again as template, due to its lower resolution, and 2NK1 as template override, while to model residues from 808 to 1268 (HM 808-1268), 2NK1 was used as template and 6MFA as template override, because of its higher resolution.

Using the aforementioned procedure, all the previous homology models were linked together, the former homology model used as template and the latter as override.

The homology model of residues from 1 to 1268 was thereby obtained (HM 1-1268).

To perform homology modelling of the residues from 1173 to 1631 (HM 1173-1268) PDB entry 4LXO.A was selected as template, and 3T1W.A was used as template override, due to a resulting higher GB/VI score. This model, that represents the 7FnIII, 8FnIII, 9FnIII, 10FnIII modules and EDB module, was then linked to the previous homology model (HM 1-1268), obtaining a homology model from residue 1 to 1631 (HM 1-1631).

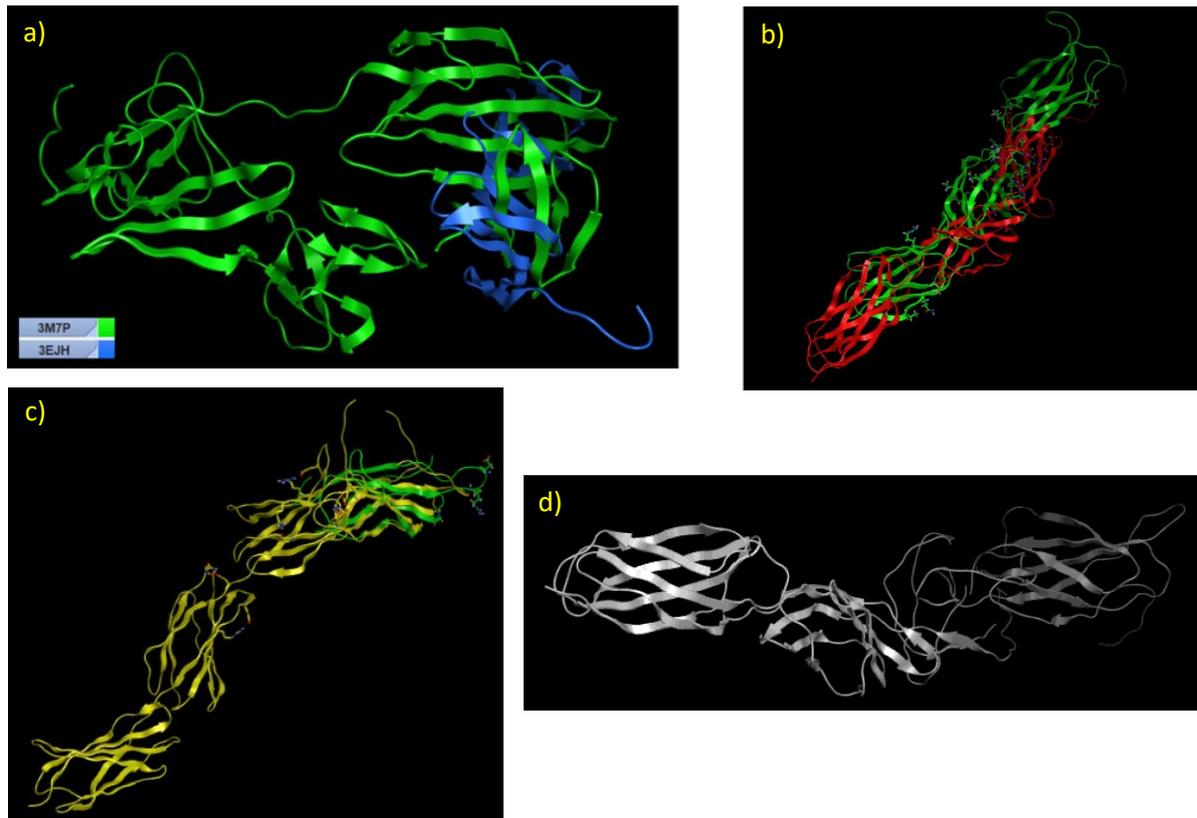


Figura 17: a) Superposition of 3M7P (green) and 3EJH (blue) structures; b) alignment and superposition of 6XAX.B (red) and 6XAY.D (green); c) alignment and superposition of 6XAX.B-6XAY.D homology model (yellow) and 4MMX.C (green) on UNP-P02751; d) homology model obtained after the alignment and superposition of 6XAX.B-6XAY.D homology model and 4MMX.C.

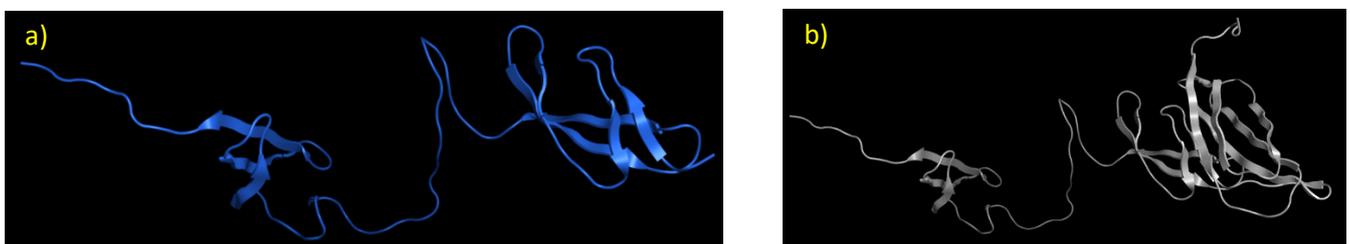


Figura 16: a) visual rendering in new cartoon of the homology model for 10-11FnI modules; b) visual rendering in new cartoon of the homology model for 15-FnIII and 10-11FnI modules.

In order to obtain a more refined structure, homology modelling was performed selecting 6XAY.D, 6XAX.B and 4MMX.C, to represent 10FnIII, 11FnIII, and 12 FnIII. Alignment and superposition of P02751 FN sequences and the chosen entries were performed (**figure 16b-c**). Entries 6XAY and 6XAX and

4MMX were merged, obtaining the homology model of residues 1539 to 1903 (HM 1539-1903) (**figure 16d**). The homology model of fibronectin from residue from 1 to 1903 (HM 1-1903) was obtained, linking HM 1-1631 with HM 1539-1903, using the latter as template override. The 10FnI module homology model was attached to the 11FnI module, represented by the 2EC3.A entry, using the latter as override. In this way we obtained an HM from residue 2273 to 2390 (HM 2273-2390) (**figure 17a**). This HM is subsequently linked to 15FnIII module structure (5MOA), using the former as template, finally obtaining the homology model from residue 2199 to 2390 (HM 2199-2390) (**figure 17b**).

Then, the variable region homology model was connected to the homology model previously obtained (2199 to 2390 residues), choosing the last fragment as template, due to the highest absolute value of GB/VI score of the obtained model from residue 2078 to 2390 (HM 2078-2390).

PDB entry 3R8Q.A, representing the 12-14FnIII modules, was linked to the resulting homology model above mentioned (HM 2078-2390), using the former as override, giving the highest GB/VI score, obtaining thereby a homology model from residue 1812 to 2390. This last model was merged with the 12FnI module HM, obtaining a homology model from residue 1812 to 2477 (HM 1812-2477) (**figure 18**).

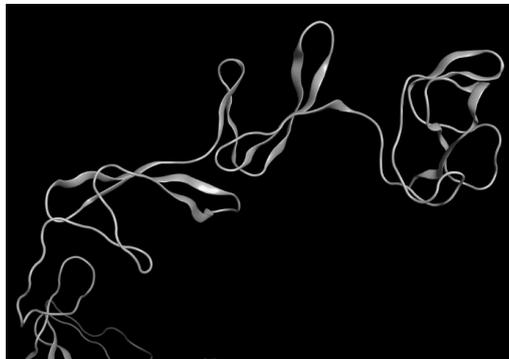


Figura 18: Homology model of residues 1812 to 2477, focusing on the 10FnI, 11FnI and 12FnIII modules.

Finally, HM from 1 to 1903 was linked to HM 1812-2477. The best GB/VI score is obtained using HM 1812-2477 as template and HM 1-1903 as override. Thus, we obtained the complete homology model of Fibronectin (**figure 19**).

FN homology model evaluation

For what concerns the main integrin binding site, the 9-10FnIII modules contained in the 4LXO PDB entry were separately analyzed. Both the plot (**figure 21a-b**) and the analysis (**table 6**) of the data showed the goodness of the model with 99.5% of residues in the core and allowed region and only 0.5% (only 1 proline residue) in outlier region.

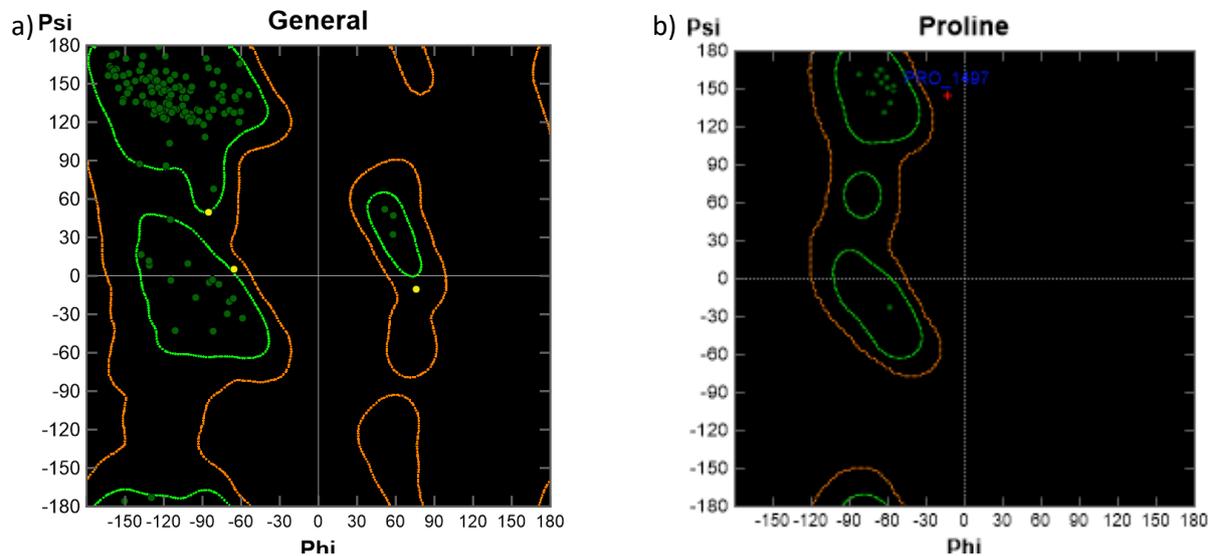


Figura 21: a) Ramachandran plot of integrin binding domain (9-10FnIII) b) Ramachandran plot of integrin binding domain, highlighting the outlier proline residue with proline filter contour

	CORE	ALLOWED
9-10 FnIII	97.8%	1.6%

Table 6: Percentage of atoms in the most favorable region (core), atoms in the additional allowed region (allowed) and atoms in disallowed region (outlier) of the model of integrin binding domain.

The bond lengths and bond angles plots show the quality of both backbone bond lengths and bond angles on a Z-Score scale. It was set by default a Z-Score cutoff of 4.0, represented by a dotted horizontal red line in **figure 22a**.

The results of Bond Length plot (**figure 22a**) show the positive absence of lengths above the Z-score cutoff, thus indicating that no anomalous bond lengths are present along the backbone. Compared to the expected values of bond lengths of a reference database (Protein Data Bank), almost the totality of bond lengths values results within the expected range, although N-C α and C=O bonds present lower Z-score value than expected (N-C α and C=O bonds reference ranges: 1.459 ± 0.012 Å and 1.233 ± 0.012 Å respectively). As a result of this, N-C α and C=O bonds lengths are not represented faithfully in the model.

Looking at the Bond Angle plot (**figure 22b**), it is easy to notice some values above the Z-score cutoff: there are present some anomalous bond angles along the backbone, as pC-N-C α angles are above the Z-Score cutoff, (angles' values superior to the reference range, with expected values: $121.60^\circ \pm 1.554^\circ$), analogous to pC α -pC-N and C α -C=O angles, that present values above the Z-score cutoff. Thus these angles will not be properly represented in the model.

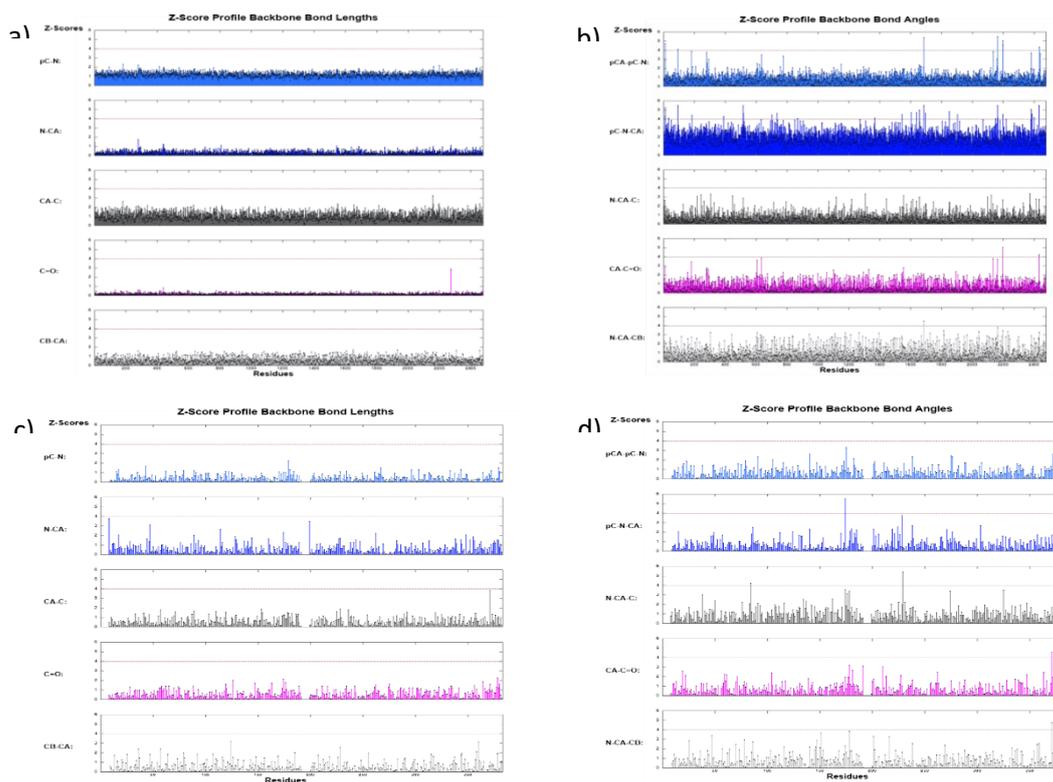


Figura 22: a) Z-score values for each type of backbone bond lengths; b) Z-score values for each type of backbone bond angles; c) Z-score profiling backbone bond lengths of the integrin-binding site (9FnIII and 10FnIII modules); d) Z-score profiling backbone bond angles of the integrin-binding site (9FnIII and 10FnIII modules).

For what concerns the integrin binding site, bond length plot does not present any outlier and bond length values are consistent with the reference values. Whereas bond angle values are more spread because of the presence of many outliers in the bond angle plot, as shown in N-C α -C, pC-N-C α , C α -C-O and N-C α -C β angles' plots(**figure 22c-d**). Only the pC α -pC-N angles' values are consistent with the expected values.

The evaluation of dihedral angles' plot shows the profile of omega, phi, psi, C-beta dihedrals, and planarity along the backbone: there are some anomalous values for the representations of phi and psi angles, already highlighted in the Ramachandran plot, but in relation to the expected values present in the backbone dihedral definitions' table, most of the values are included. For omega and C-beta only a few outliers are presented, proving the goodness of AF_HMFN models, also demonstrated by the absence of anomalies in the planarity diagram (**Figure 23a**).

Analyzing the integrin binding domain, it shows that there are only some anomalous values for omega angles, while no outliers are present in the planarity, phi/psi angles and C-beta plot (**figure 23b**).

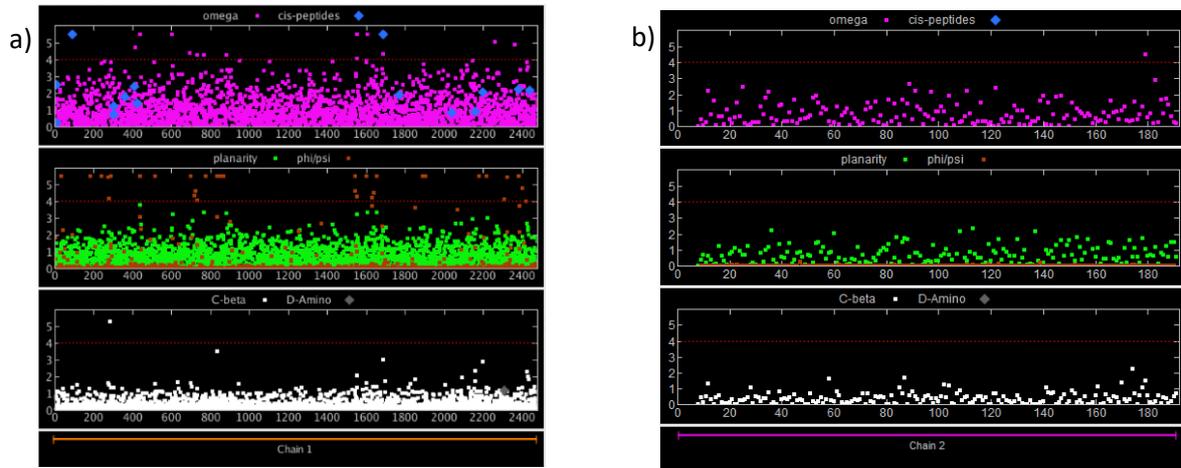


Figura 23: a) dihedral angles evaluation after energy minimization; b) dihedral angles focused on the integrin domain.

Through evaluation of the rotamers, the scoring energy along the backbone is computed in order to predict sidechains conformations. The lower is the E-value (Expected energy value), the better the result of rotamers prediction is. By analyzing the graphic (**figure 24a**), it is shown that major part of energy-based scoring is included in the E-threshold [-1; +5], imposed by the system.

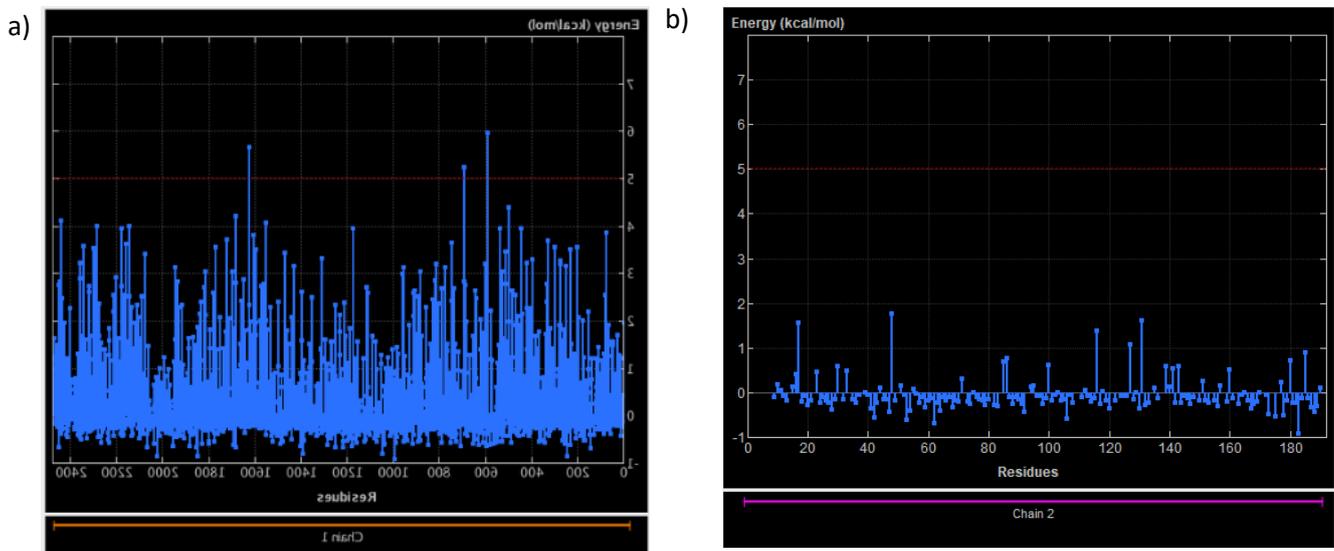


Figura 24: a) fibronectin backbone rotamers energy plot; b) Integrin-binding domain rotamers energy values.

The evaluation of rotamers in the integrin-binding domain (4LXO) proved the goodness of this sequence because of the low energy values (**figure 24b**).

Plotting the atom clashes of AF_HMFN, no atom collides in any way with other non-bounded atoms. Thus, the plot will represent no values, as it can be shown in figure 43. This is clearly the result of the minimization executed on the homology model BF_HMFN. Energy minimization provides a more stable and energy minimized configuration, by removing the atom clashes registered in the final residues (**figure 25a-b**), that are the ones missing by the protein.

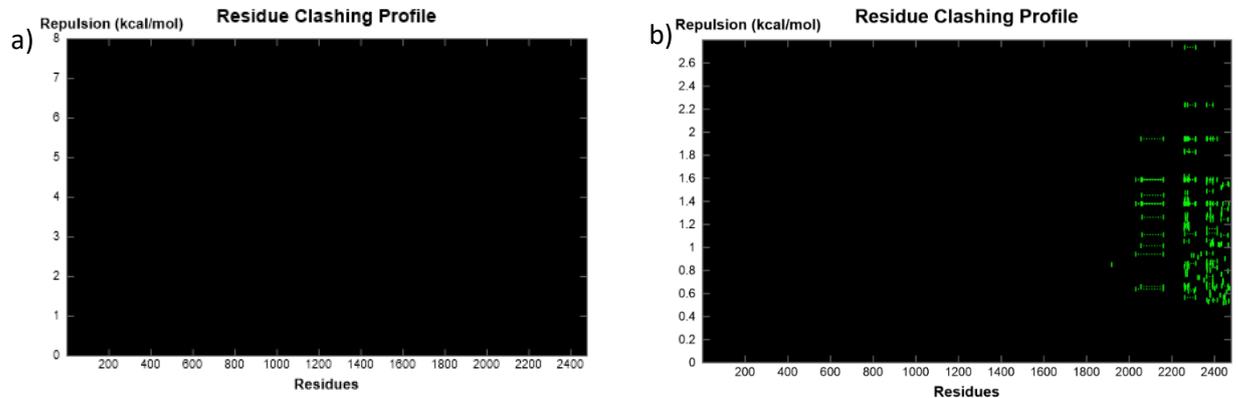


Figura 25: plotted atom clashes after (a) and before (b) energy minimization. The plot of the atom clashes computes and represents the sum of the repulsion energies for each residue's atoms. The clashing of two non-bounded atoms is defined by a positive repulsion energy term (in kcal/mol). No value is plotted after energy minimization, even after setting the E-Threshold to 0.1.

Focusing on the integrin-binding domain, the 4LX0 taken out of the fibronectin structure presents some atom clashes, that subside when analyzing the entire final FN homology model. (**Figure 26**). This effect could be due to the fact that the 4LX0 terminus, where the atom clashes are present, is bound to the next fragment.

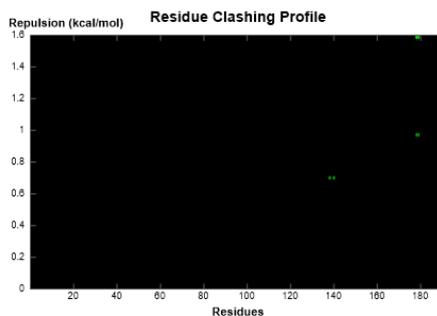


Figura 26: Atom clashes' plot through the integrin-binding domain.

The effective atomic contact energies are plotted and analyzed (**figure 27a-b**). The outer residues present positive contact energies, while the residues more entangled within the protein present a more negative contact energy. This result is consistent with the fact that the more hydrophobic residues present large negative values and are buried in the protein environment, while the more hydrophilic residues present positive values and are expected to be in exposed regions of the protein (Yan et al., 2008).

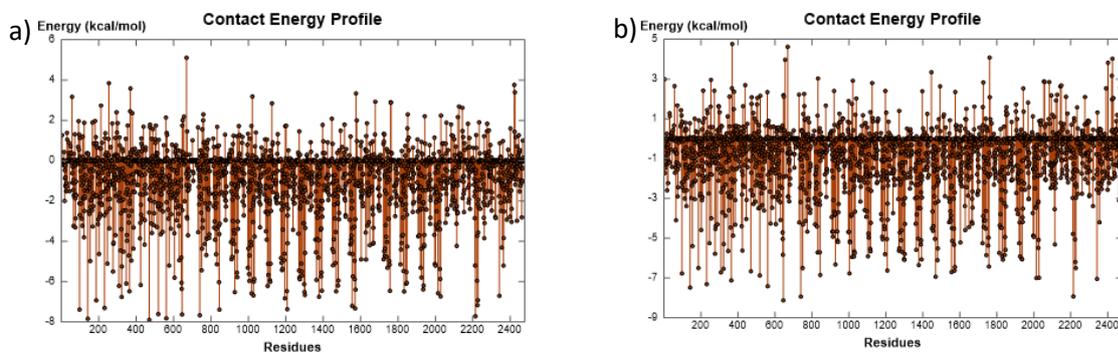


Figure 27: plotted contact energies of homology model before (a) and after (b) energy minimization . The effective atomic contact energies are thus obtained, calculated for heavy atoms of standard amino acids within a contact range of 6 Å. These energies are summed for each residue in the system.

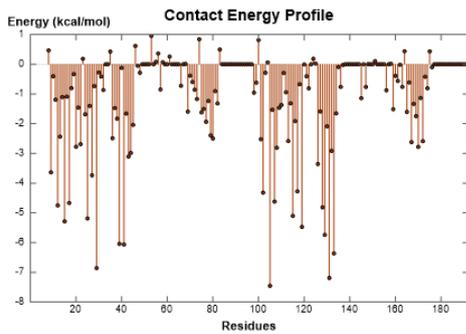


Figure 28: Integrin-binding site contact energies.

Overall and Local Model Quality

The AF HMFN Z-score was calculated with ProSA (Protein Structure Analysis), and results were evaluated with the help of WhatCheck tool. An overall model quality Z-score of -13.7 was obtained, meaning that the overall structure obtained is worse than average structures present in the Data Base (**figure 29a**). The chi-1/chi-2 angle correlation score was good (-0.736), because the angles of all residues are within expected ranges for well-refined structures. Also, Ramachandran Z-score was of -3.480, expressing that the backbone conformations of all residues have an acceptable correspondence to the known allowed areas in the Ramachandran plot, in agreement with the previous results obtained with Moe. Unfortunately, the backbone conformation shows that the fold in the structure is very unusual, resulting in a very low Backbone conformation Z-score of -21.043. This last result could be due to the absence of an exploitable template for fragments 2083-2198, 2285-2329 and 2391-2477, in literature as well as in Data Banks (PDB and Blast). For this reason, it is possible that the folding of the backbone in these fragments is not appropriate.

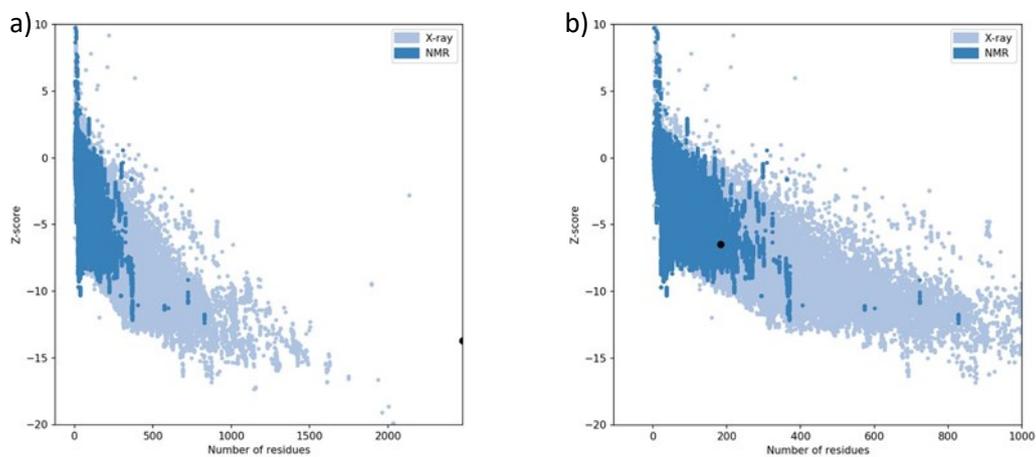


Figure 29: Z-score residues distribution for the complete FN Homology model (a) and the integrin binding site (b) (fragment 4LXO).

On the other hand, the integrin binding site analyzed obtained a z-score of -6.48, which means that the homology model represents this site on average good (**figure 29b**). In fact, both Ramachandran and chi-1/chi-2 angle correlation z-score resulted being around zero, respectively of -0.562 and 0.172, because backbone conformations of all residues are within expected ranges to the known allowed areas in the Ramachandran plot, as well as to the populated areas in the chi-1/chi-2 angle database. Again, the backbone conformation z-score obtained was lower than usual: -23.131. This last result may not be reliable, since the structure chosen to represent the integrin binding site (4LXO) has already been crystallized, so this fragment was taken directly from the PDB, showing both high resolution and a good overall quality of 83.7681% [ERRAT].

Figure 30a is obtained from local model evaluation, showing local model quality by plotting energies as a function of amino acid sequence position. Positive values correspond to problematic or erroneous parts of the input structure, whereas negative parts define a well-refined structure. As expected, high energies are plotted mostly at the beginning and at the end of the sequence. In fact, 1-48 residues, as well as 2083-2198, 2285-2329 and 2391-2477 residues were modeled using the proper aminoacidic sequence, but an estimated structure, due to lack of this information on Data Banks or literature.

Thus, the high energy confirms the possibility that problematic or erroneous parts may occur in the missing fragments of FN modeled in this study.

For what concerns only the integrin binding site (4LXO) the plot of energies as a function of amino acid sequence position shows negative values, confirming again the well-refined structure of this domain (**figure 30b**).

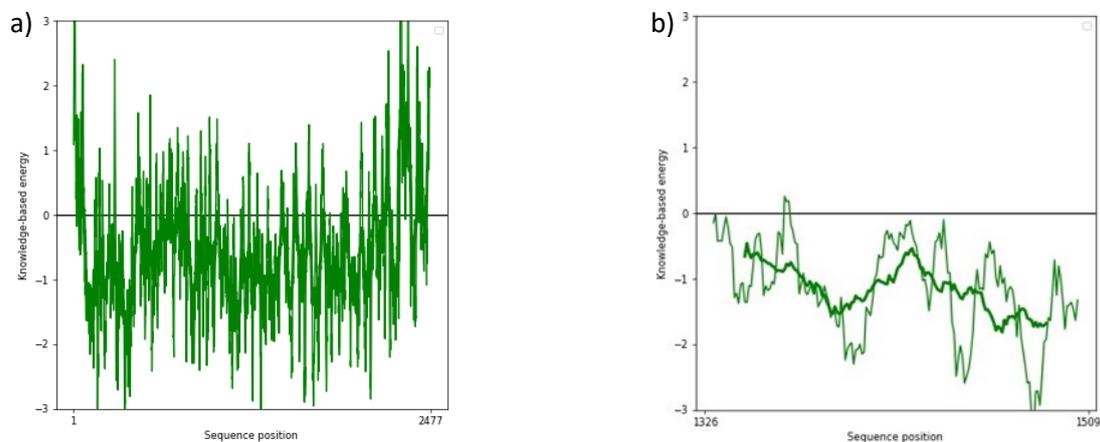


Figure 30: a) local model energy of AF_HMFN: positive values show problematic or erroneous parts of the input structure, negative parts define a well-refined structure; b) local model energy of integrin binding site.

The energy of AF_HMFN is visualized on ProSA-web, on the 3D structure of the input protein using the molecule viewer Jmol. Residues are colored from blue to red, from lower to higher residue energy. As a result, red zones occur both at the beginning and at the end of the sequence, confirming again the modelling of these fragments is not remarkably precise, due to the absence of usable templates (**figure 31a**).

Focusing on the integrin binding site, the energy visualization of the 3D structure shows the prevalence of low energy residues, plotted in blue. Thus, the good quality of the structure of the domain is again validated (**figure 31b**).

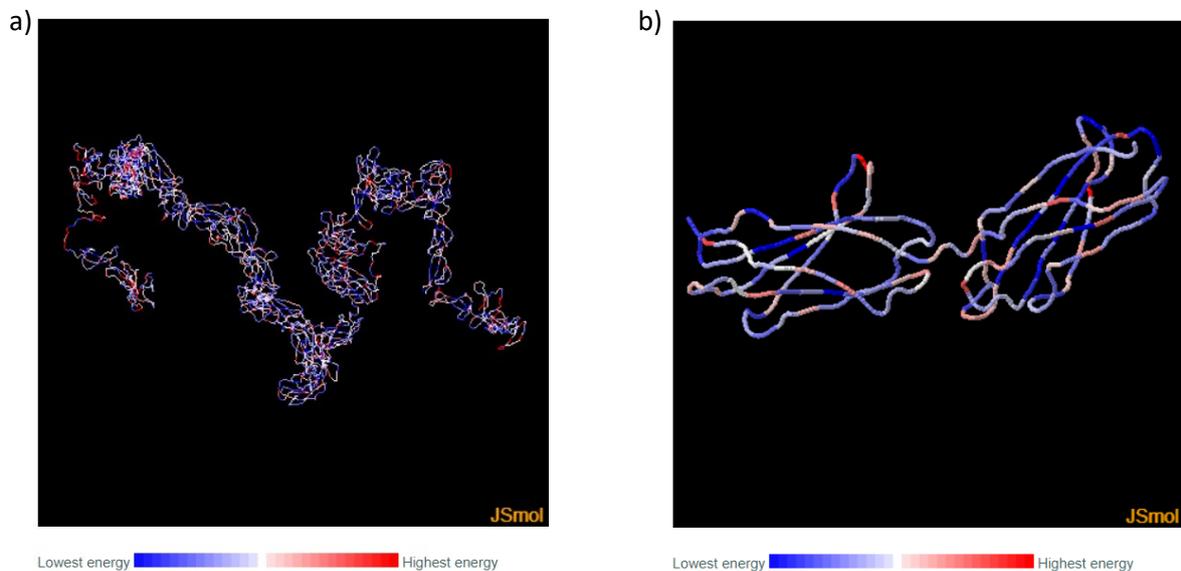


Figura 31: a) Energy visualization on the 3D structure of AF HMFN: in blue are shown residues with lowest energy, whereas in red the ones with the highest energy; b) energy visualization on the 3D structure of integrin binding site: in blue are shown residues with lowest energy, whereas in red the ones with the highest energy.

In conclusion the most problematic residues have been: 1-47, 2083-2198, 2285-2329 and 2391-2477. Thus, the reliability of the model in these areas and the relative position of adjacent modules cannot be very high, as confirmed by the results of geometry analysis and local model quality.

Although these unfavorable results, the quality of the integrin binding sites is good enough, thanks to the fact that a PDB structure of the fragments was already present in literature. In conclusion, the good overall results obtained from Ramachandran plot, backbone bond length and atom clashes analysis make this model a good starting point for further investigation and refinement of the structure of FN.

1.1.2. Integrins

Structure and physiological importance

Integrins are a superfamily of cell adhesion receptors that recognize mainly ECM ligands, cell surface ligands and some soluble ligands. Integrins are so named because they integrate the extracellular and intracellular environments by binding to ligands outside the cell, the cytoskeleton and signaling molecule inside cells (B. Luo et al., 2007). Integrins function as traction receptors that can both transmit and detect changes in mechanical force acting on the ECM. They present in the form of heterodimers 280 Å long, made up of type I transmembrane α (150 to 180 kD) and β (90 kDa circa) subunits noncovalently associated, of which 18 α subunits and 8 β are currently known. α and β integrin structure, along with the inserted α -I domain (or α -A), has been highly conserved during the evolution of vertebrates (Takada et al., 2007). The combinations of specific subunit types generate 24 heterodimers. Both subunits are composed of a large extracellular domain, a single-pass transmembrane (TM) domain and a small cytoplasmic tail. The diversity in subunit composition induces the different ligand recognition.

Usually, different integrin heterodimers are presented in different cell or tissue, with some types being almost ubiquitous (Mould et al., 2014):

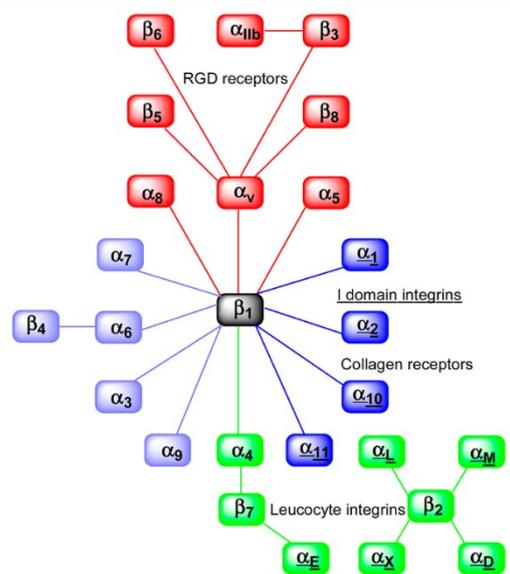
- **$\alpha 5\beta 1$** : fibronectin-specialized integrin, is expressed by many cell types and it represents the major fibronectin receptor in most of them. The RGD loop in 10FnIII is the critical recognition site, but the PHSRN in 9FnIII is required to acquire high affinity binding. It is the first integrin identified involved in the fibronectin network formation. It localized to the sites where fibronectin fibrils are in contact with the cell, and antibodies to $\alpha 5$ or $\beta 1$ can inhibit the polymerization in fibroblast cultures. Overexpression of $\alpha 5$ in CHO cells increases the deposition of fibronectin in the matrix. Mouse embryos deficient in the $\alpha 5$ gene still contained extracellular fibronectin networks, indicating that also other integrins could promote the polymerization process.
- **$\alpha 11\beta 3$** : apart from platelets it has only been found in megakaryocytes. Its task is to bind fibrinogen during thrombus formation. Both the RGD site and the synergy site DRVPHSRNSIT contribute to the binding of FN, in a similar manner to $\alpha 5\beta 1$. The interaction with the synergy site alone is 30-fold weaker compared to RGD peptides. These peptides can inhibit the binding of FN or fibrinogen to isolated receptors or to platelets.
- **$\alpha 4\beta 1$** : expressed in white blood cells and some types of adherent cells. It can mediate cell-cell as well cell-ECM contacts through VCAM-1 and FN respectively. Both process are important for leucocyte extravasation and immunological and inflammatory events. Even in unactivated state is capable of binding to VCAM-1 permitting to roll activated endothelial cells. It interacts mainly with the 14FnIII-V region in FN. CS1 is the predominant site in the FN that present it. CS1 and CS5 cross-inhibit each other for binding $\alpha 4\beta 1$, but they probably don't compete for the same binding site because recombinant FN fragments for the receptor present higher affinity when both sites are included. H1 doesn't inhibit the binding of CS1 to $\alpha 4\beta 1$. $\alpha 4\beta 1$ can also recognize the RGD domain, but it is obtained only in presence of integrin activating antibodies and its physiological relevance is to be determined...
- **$\alpha 3\beta 1$** : expressed in epithelial cells in the skin, endothelial cells in the digestive tract, kidney mesangial cells and tumor cells in vivo, and by most cell culture in vitro. Laminin 5 appears to be an important ligand for the receptor and is been observed also weak binding to FN and other ligand under certain conditions. It is unable to mediate cell adhesion except for laminin 5, so it clearly isn't an FN receptor in intact cells. It's been suggested to function as secondary receptor with post adhesion functions.
- **$\alpha 8\beta 1$** : expressed mainly in epithelial cells, smooth muscle cells. Myofibroblasts, embryonic neural cells. It binds to RGD sites in FN and vitronectin, and to tenascin-C via both RGD-dependent and -independent interactions. On these ligands **$\alpha 8\beta 1$** is capable of mediating cell adhesion and neurite outgrowth.
- **$\alpha v\beta 3$** : abundantly expressed by many cell cultures in vitro and by endothelial cells, osteoclasts and tumoral cells. It's important for angiogenesis and is therefore a potential target for inhibition of tumor growth. It binds FN, vitronectin and several other cell adhesion proteins via the RGD domain. It does not require the synergy site of FN for stable interaction.
- **$\alpha v\beta 6$** : expressed by epithelial cells during development and wound healing, and in many epithelial tumors. The 11 unique C-terminal amino acids of the cytoplasmic part of $\beta 6$ contain a proliferation promoting activity. This cytoplasmic tail also contains three regions required for the localization to focal contacts (membrane proximal region and two NPXY motifs). The corresponding region in $\beta 1$ and $\beta 3$ share the same function. Its main ligand is FN and in a weaker manner tenascin-C. It binds to FN using the RGD site and it does not require the synergy site in a similar manner to $\alpha v\beta 3$.

- **$\alpha 4\beta 7$** : expressed on subsets of lymphocytes, in particular a subset of T memory cells. In a similar manner to $\alpha 4\beta 1$, it binds to the 14FnIII-V region in FN, VCAM and the LDV site in the $\alpha 4$ subunit. It also recognized MadCAM-1, important interaction for homing of specific lymphocyte populations to mucosal site
- **$\alpha E\beta 7$** : expressed in T cells, dendritic cells, and mast cells in mucosal tissues.
- **$\alpha ?\beta 8$** : The $\beta 8$ subunit is expressed mainly in the brain, spinal cord, kidney and embryonic muscle. $\beta 8$ sequence is rather different from other integrins and in particular their cytoplasmic domain has no similarity. The question mark means that the α subunits associated with $\beta 8$ haven't been identified yet for FN: $\alpha V\beta 8$ can form in $\beta 8$ transfected 293 cells, but it does not bind FN. $\beta 8$ can probably combine with several different α subunits.
- **$\alpha 6\beta 4$** : expressed in keratinocytes.
- **$\beta 2$** integrins are widely found in leukocytes.

Regarding ligand specificity, mammalian integrins can be grouped distinguishing the different protein binding:

- laminin-binding integrins
- collagen-binding integrins
- leukocyte integrins
- RGD-recognizing integrins

Figure 32: human integrin superfamily and their combinations to form heterodimeric integrins. Integrin subunits that bind to each other to form heterodimer are connected by solid lines. Each integrin has distinct ligand-binding specificity and tissue and cell distribution (Takada et al., 2007).



Contacts between the two subunits involve mainly their NH_2 halves, which together form a globular head, while the remaining portions form two rod-shaped tails that span into the plasma membrane (Catterall et al., 2001).

The two dimers are totally distinct, and no homology is detectable between them; sequence identity among α subunits is about 30% and among β subunits 45%, indicating that both subunits originate from gene duplication. In some integrins is also present the I (insertion or interaction) or A domain, which play an important role in ligand binding and intercellular adhesion. The I-domain integrin α subunits are closely related to each other. Also, the α subunits that recognize the RDG binding site ($\alpha V, \alpha 8, \alpha 5$, and $\alpha 11b$) and the family of laminin-binding α subunits ($\alpha 3, \alpha 6$, and $\alpha 7$) are closely related to each other (**figure 32**). α and β subunits structure, along with the inserted α -I domain, has been highly conserved during the evolution of vertebrates. The different β -subunits in the cytoplasmic tail present high homology between them, while α -subunits are highly divergent except for a GFFKR motif next to the transmembrane region, important for the association with the β -tail (Takada et al., 2007).

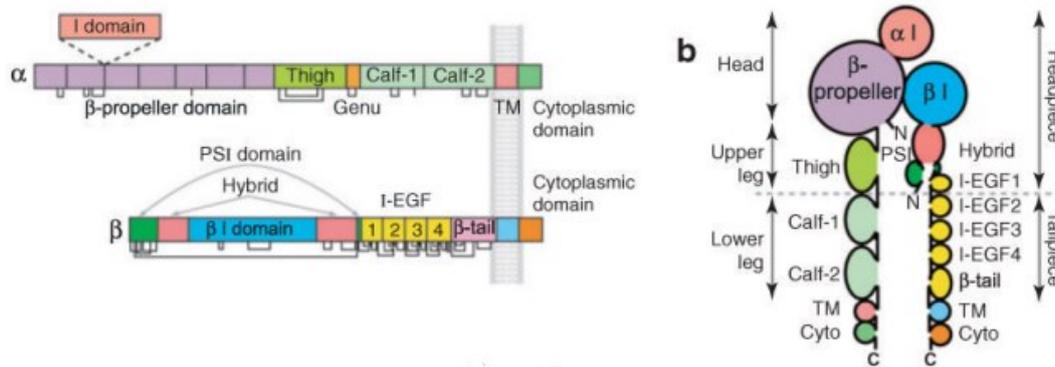


Figure 33: a) Organization of domains within the primary structures. Some α subunits contain an I domain inserted in the position denoted by the dashed lines. Cysteines and disulfides are shown as lines below the stick figures; b) schematic of the course of α and β subunit polypeptide chains through domains from the N to C termini (B. Luo et al., 2007).

The overall structure of integrins is that of a large “head” region supported by two long “legs”. The head region includes the β -propeller domain if the α subunit and the β -I (or β -A) and hybrid domain of the beta subunit (figure 33).

At least three major conformational states of integrins have been identified (figure 34) (Mould et al., 2014):

- closed form with bent legs and closed headpiece (low affinity).
- intermediate form with extended legs and closed headpiece (medium affinity).
- open form with extended legs and open headpiece (high affinity).

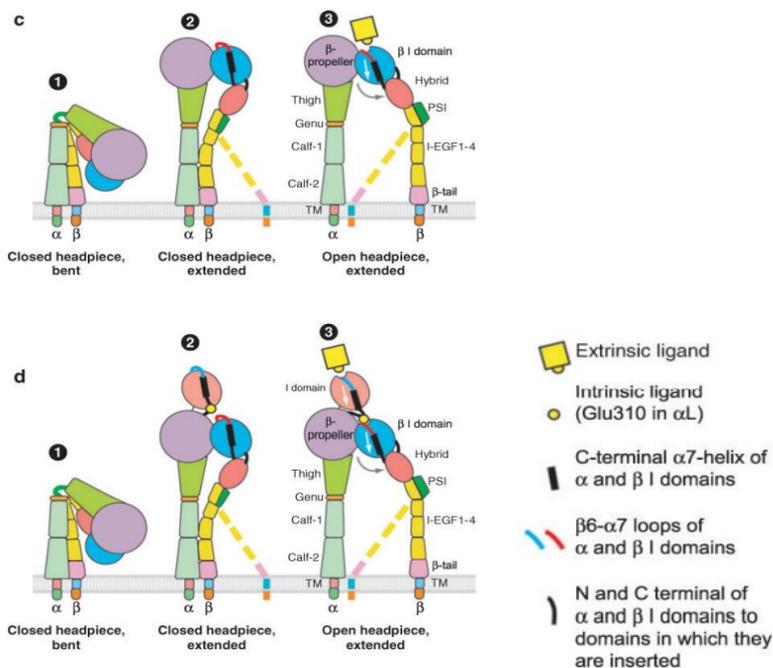


Figure 34: shifting between integrins' conformational states, with and without I domain, following binding of an extrinsic ligand (B. H. Luo et al., 2007; B. H. Luo & Springer, 2006).

Integrins are capable of binding multiple ligands and individual integrins have unique ligand specificities: from analysis on the protein bound to integrins that is no evidence that there would be separate binding sites for different ligands, and the majority recognized the RGD sequence as binding motif. Usually, a cell adhesion protein can bind to more than one type of integrin, and fibronectin represents an extreme case, being capable of binding to multiple integrin types: this is very probably

due to different intracellular signals being generated by the ligand depending on which integrin it associates. This is further proven by gene knockout analysis (Johansson et al., 1997). Many signaling pathways can regulate integrin activation, with recent studies focusing on the talin as indispensable mediator of this process (Ginsberg et al., 2005).

Focusing on the integrins types that can bind to fibronectin, we have:

- $\alpha 4\beta 1$
- $\alpha 5\beta 1$
- $\alpha 8\beta 1$
- $\alpha V\beta 1$
- $\alpha D\beta 2$
- $\alpha IIb\beta 3$
- $\alpha V\beta 3$
- $\alpha V\beta 6$
- $\alpha 4\beta 7$

A large subset of IN is comprised by those recognizing the RGD sequence in their ligands:

- $\alpha 3\beta 1$
- $\alpha 5\beta 1$
- $\alpha 8\beta 1$
- $\alpha V\beta 1$
- $\alpha IIb\beta 3$
- $\alpha V\beta 3$
- $\alpha V\beta 5$
- $\alpha V\beta 6$
- $\alpha V\beta 8$

As can be noted by comparison with the previous group, many of these recognize the RGD site in the prototypical IN ligand FN (**figure 35**: underlined the FN-binding INs in common with the RGD-recognizing INs).

Integrins	Ligands
$\alpha 1\beta 1$	Laminin, collagen
$\alpha 2\beta 1$	Laminin, collagen, thrombospondin, E-cadherin, tenascin
✘ $\alpha 3\beta 1$	Laminin, thrombospondin, uPAR
✘ $\alpha 4\beta 1$	Thrombospondin, MAdCAM-1, VCAM-1, fibronectin, osteopontin, ADAM, ICAM-4
✘ $\alpha 5\beta 1$	Fibronectin, osteopontin, fibrillin, thrombospondin, ADAM, COMP, LI
$\alpha 6\beta 1$	Laminin, thrombospondin, ADAM, Cyr61
$\alpha 7\beta 1$	Laminin
✘ $\alpha 8\beta 1$	Tenascin, fibronectin, osteopontin, vitronectin, LAP-TGF- β , nephronectin,
$\alpha 9\beta 1$	Tenascin, VCAM-1, osteopontin, uPAR, plasmin, angiostatin, ADAM [25], VEGF-C, VEGF-D [26]
$\alpha 10\beta 1$	Laminin, collagen
$\alpha 11\beta 1$	Collagen
✘ $\alpha v\beta 1$	LAP-TGF- β , fibronectin, osteopontin, LI
$\alpha L\beta 2$	ICAM, ICAM-4
$\alpha M\beta 2$	ICAM, iC3b, factor X, fibrinogen, ICAM-4, heparin
$\alpha X\beta 2$	ICAM, iC3b, fibrinogen, ICAM-4, heparin, collagen [27]
✘ $\alpha D\beta 2$	ICAM, VCAM-1, fibrinogen, fibronectin, vitronectin, Cyr61, plasminogen
✘ $\alpha 11b\beta 3$	Fibrinogen, thrombospondin, , fibronectin, vitronectin, vWF, Cyr61, ICAM-4, LI, CD40 ligand [28]
✘ $\alpha v\beta 3$	Fibrinogen, vitronectin, vWF, thrombospondin, fibrillin, tenascin, PECAM-1, fibronectin, osteopontin, BSP, MFG-E8, ADAM-15, COMP, Cyr61, ICAM-4, MMP, FGF-2 [29], uPA [30], uPAR [31], LI, angiostatin [32], plasmin [33], cardiotoxin [34], LAP-TGF- β , Del-I
$\alpha 6\beta 4$	Laminin
✘ $\alpha v\beta 5$	Osteopontin, BSP, vitronectin, CCN3 [35], LAP-TGF- β
✘ $\alpha v\beta 6$	LAP-TGF- β , fibronectin, osteopontin, ADAM
✘ $\alpha 4\beta 7$	MAdCAM-1, VCAM-1, fibronectin, osteopontin
$\alpha E\beta 7$	E-cadherin
✘ $\alpha v\beta 8$	LAP-TGF- β

Figure 345: Different integrin heterodimers and their ligands. Highlighted in red are the FN-binding integrins. Marked with a X are the IN that recognize RGD as binding motif in their ligand (Takada et al., 2007).

The RGD sequence in FN was originally identified as an IN-binding motif, and indeed this and related sequences act as IN-binding motifs in vivo in ECM molecules. Individual integrins are also specific for particular protein ligands. Immunologically, the intercellular adhesion molecules (ICAMs) are important integrin ligands (Takada et al., 2007).

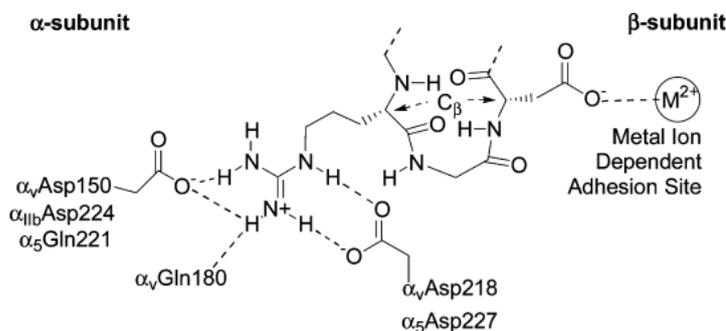


Figura 6: Key features of RGD-binding site (Sheldrake & Patterson, 2014).

Integrin Signaling and Ligand binding

Integrin signaling is bidirectional: “inside-out” signaling (or priming) regulate integrin affinity for adhesive ligands, in particular high-affinity ligand bindings requires the “activation” of the integrin by conformational changes induced by inside-out signals; in turn “outside-in” signals induced by integrin

ligation regulate cellular responses to adhesion, motility and gene expression (Ginsberg et al., 2005). Many cytoskeletal and signaling proteins bind to β cytoplasmic tails and have been found to interact with specific α tails. Most integrin β tails contain one or two NPxY/F motifs (x=any amino acid), part of canonical recognition sequence for phosphotyrosine-binding (PTB) domains. Phosphorylation of this tyrosine may be involved in the regulation of IN interaction with proteins on the cytoplasmic face of plasma membrane. The integrin tails recruit several proteins (e.g. talin that binds to actin filaments connecting to the cytoskeleton) (Takada et al., 2007).

Integrin β subunit cytoplasmic domains are required for IN activation, while in most cases the α cytoplasmic domain plays a regulatory role. The shift between active and inactive state is dynamically controlled by the cell through energy dependent interactions involving the IN cytoplasmic tail: this induces a change in conformation of the extracellular domain, that induces changes between the active and inactive states and regulates changes ligand binding affinity. A third conformational state is induced after ligand occupation. Experimentally the three state can be stabilized by binding different monoclonal antibodies to the two subunits at epitopes distinct from the ligand-induced binding site (LIBS), exposed following ligand binding (Ginsberg et al., 2005; Johansson et al., 1997).

In adherent cells the integrins are mainly in the active state, but they can become inactivated during specific situations, like cell migration. In circulating cells the integrins are on the cell surface in inactive state, and can be activated with exposition to factors that trigger intracellular reactions (Johansson et al., 1997).

Receptor ligand binding can involve either induced-fit (ligand binds to a low affinity form of receptor which then converts to high affinity form) or selected-fit (in which binding to high affinity form is preferred to low-affinity) interactions. Regarding the RGD-binding integrins there is a debate concerning the relative importance of each type (Mould et al., 2014).

In the nine α I-containing integrins, α I is necessary and sufficient for divalent cation-dependent binding to physiology ligands. From the structure of isolated α I domains in the liganded and unliganded conformations was revealed that a metal ion is coordinated at the ligand binding interface of α I through a conserved five amino acid motif, the metal ion-dependent adhesion site (MIDAS), and the metal coordination is completed by a glutamate from the ligand or a water molecule. In α I-lacking integrins, ligand recognition requires an α I-like domain, the β I domain, present in all β subunits. The crystal structure of the extracellular segment of the α I-lacking integrin α V β 3 was previously determined in presence of Ca^{2+} (B. H. Luo et al., 2007).

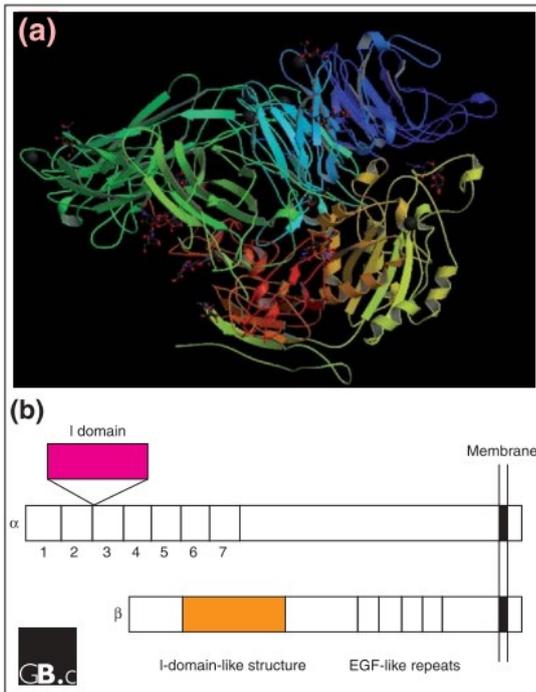


Figure 37: a) Crystal structure representing a net form of integrin $\alpha V\beta 3$ with no bound RGD peptide (PDB code 1JV2). B) The I (inserted or interactive) domain is present in seven human α subunits between β -propeller repeats 2 and 3, and is involved in ligand binding. An I-like domain is present in all human integrin β subunits along with four EGF-like repeats. Both the I and I-like domains have a Rossmann fold (Takada et al. 2007).

From analysis on the crystal structure of human integrins $\alpha V\beta 3$ and $\alpha IIb\beta 3$ show that the extracellular portion of an integrin heterodimer consists of multiple domains (**Figure 37a**). The headpiece of $\alpha V\beta 3$, which contains the ligand-binding site, consists of the β -propeller domain and the plexin-semaphorin-integrin (PSI) domain of the αV subunit, and the β I-like domain and the hybrid domain of the β subunit. The β -propeller domain contains seven repeats of about 60 amino acids each that fold into a seven-bladed β -propeller structure similar to the β subunit of a hetero- trimeric G protein. I domains contain a metal-ion-dependent adhesive site (MIDAS) and I-like domains contain a structurally similar metal-binding motif. The RGD-binding site is located at the interface between the β -propeller domain and the β I-like domain and amino-acid residues from the two domains interact directly with the RGD peptide of a ligand (**Figure 37b**) (Takada et al., 2007). comparing the crystal structures of RGD-bound and unbound forms it has been found that the disulfide-linked loop structure in the β I-like domain undergoes conformational changes and the α helix 7 and α helix 1 in the β I-like domain moves downward on ligand bindings. Also, the hybrid domain swings outward from the β -propeller upon integrin activation. In the I-domain integrins, the I domain can be in either open (=active) or closed (=inactive) conformations, inducing conformational changes that affect ligand binding in the head of the molecule. Loops on the upper surface of the β -propeller and the top face of the β I domain form the ligand-binding pocket.

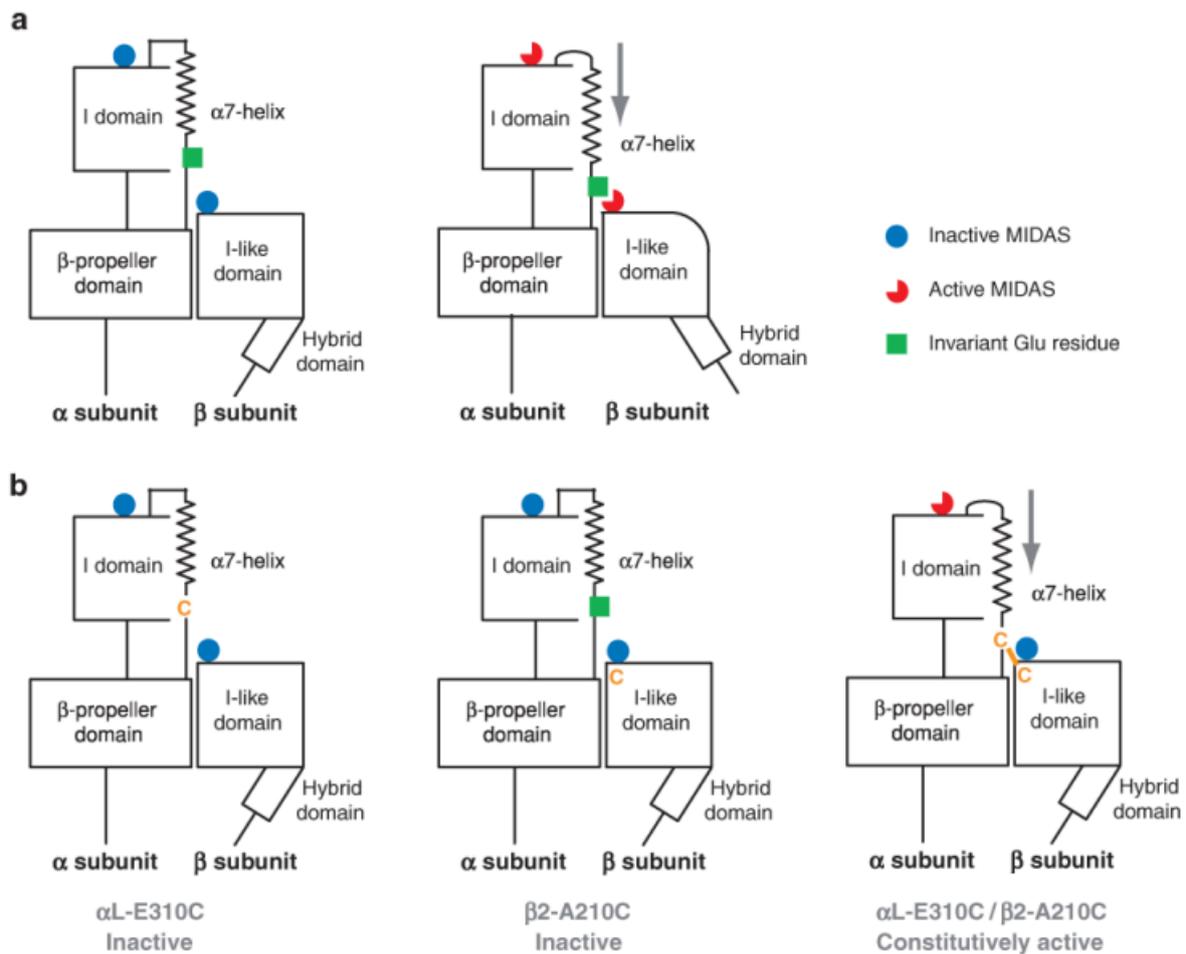


Figure 38: (a) Communication between alpha I and beta I domains. It has been proposed that alphaL-Glu-310 acts as intrinsic ligand that binds to the beta2-I domain MIDAS and thus, axially displaces the alphaL I domain alpha7-helix in the C-terminal direction, reshapes the beta6-alpha7 loop and activates the alphaL-I domain MIDAS. (b) Individual mutation of alphaL-Glu-310 or beta2-Ala-210 cysteine abolishes I domain activation, whereas the double mutation of alphaL-E310C with beta2-A210C forms a disulfide bond that constitutively activates ligand binding (B. H. Luo et al., 2007)

Strictly correlated with the conformational changes in integrins is their requirements for extracellular divalent cations for ligand binding ability. There is evidence for at least three functionally important coordination sites for divalent cations in the extracellular domain of integrins, each with different ion preferences (Johansson et al., 1997):

- **MIDAS:** metal-ion dependent adhesion site
It involves coordination of a divalent cation to residues Ser 121, 123 and Glu220 at the top of the β -propeller. In RGD-binding a carboxylate group of Asp contacts the divalent cation in MIDAS.
- **AdMIDAS:** Adjacent to MIDAS
Depending on the conformational state it can coordinate several residues with a divalent cation. The most important are Asp126, Asp127, Ser 134, Asp137, Asp138, Ala342. During RGD ligand AdMIDAS coordination by $\beta6$ - $\alpha7$ loop backbone carbonyl is replaced by an Asp side chain that coordinates through a water molecule to the divalent cation in MIDAS.
- **SyMBS:** Synergistic metal-binding site (or LIMBS)
Occupied by Asp158 and Glu220 in the unliganded state, it becomes available and hosts the coordination of a divalent cation with the carboxylate group of Glu220, the side chains of Asp158,

Asn215 and Asp217, and the carbonyl oxygen of Asp217 and Pro219. The ion doesn't directly contact the ligand, but ion coordination depends on it.

These sites are structurally interlinked and are located on the top face of the β -I domain (Mould et al., 2014).

Several good candidates for ion coordination sites have been identified in fibronectin binding integrins: different α subunits contain three or four homologous sequences (DxDxDGxxD), similar to the Ca^{2+} binding EF-hand motif, and a second type of cation binding motif (DxSxS), located in β subunits. A 14 amino acid peptide from the β 3 subunit (residue 118-131), conserved among all β subunits, was found to bind Mn^{2+} , Mg^{2+} , or Ca^{2+} with 1:1 stoichiometry. This site could also bind RGD peptides. It has been hypothesized that one of the integrin-bound ions may be directly involved in bridging to an aspartic acid residue of the ligand, but binding of RGD-peptides to a mutant form lacking ion binding seem to confute this hypothesis. The interaction between integrin and RGD-based ligands seem to be independent from ions involvement, but ions and ligands mutually influence their interactions with the receptor through allosteric regulation of the receptor conformation (Johansson et al., 1997).

These bivalent ions (Mg^{2+} , Mn^{2+} , Ca^{2+}) each have distinct effects on integrin affinity and conformational state through binding to these sites. Mn^{2+} or Mg^{2+} (or more rarely Ca^{2+}) occupancy of the MIDAS is critical for ligand recognition through coordinating to a carboxyl oxygen, such as in the side chain of aspartate in ligands containing the RGD sequence. Ca^{2+} has multiple effects on ligand binding and can be stimulatory or inhibitory, depending on the Ca^{2+} concentration and whether Mn^{2+} or Mg^{2+} is also present. The SyMBS mediates the positive effects of low concentrations of Ca^{2+} on Mg^{2+} -supported ligand binding, whereas the AdMIDAS mediates the negative effects of higher Ca^{2+} concentrations on Mn^{2+} -supported ligand binding. Mn^{2+} binding promotes the formation of the open form of the headpiece, whereas Ca^{2+} favours the closed headpiece (Mould et al., 2014).

- Mg^{2+} is most likely the dominating ion which supports ligand binding to integrins under physiological conditions. It is unknown if Mg^{2+} binds to more than one site in integrin.
- Ca^{2+} presents a concentration-dependent behavior; at millimolar concentrations inhibits the ligand binding ability of most integrins inducing the inactive state, while at low concentrations (micromolar) stimulates ligand binding. This is evidence that Ca^{2+} binds to at least two different sites in integrins.
- Mn^{2+} stimulates ligand binding to integrins even more strongly supporting ligand binding probably in the same way as Mg^{2+} but is uncertain if this is of physiological relevance. Mn^{2+} can induce/stabilize a ligand binding conformation of otherwise inactive receptors. Besides, it can activate integrins independently by the physiological intracellular events (e.g Mg^{2+}). By equilibrium gel filtration three Mn^{2+} ions were found to bind each α IIb β 3 integrin.
- It is unclear if any of the metal divalent ion sites can accept alternative divalent ions.

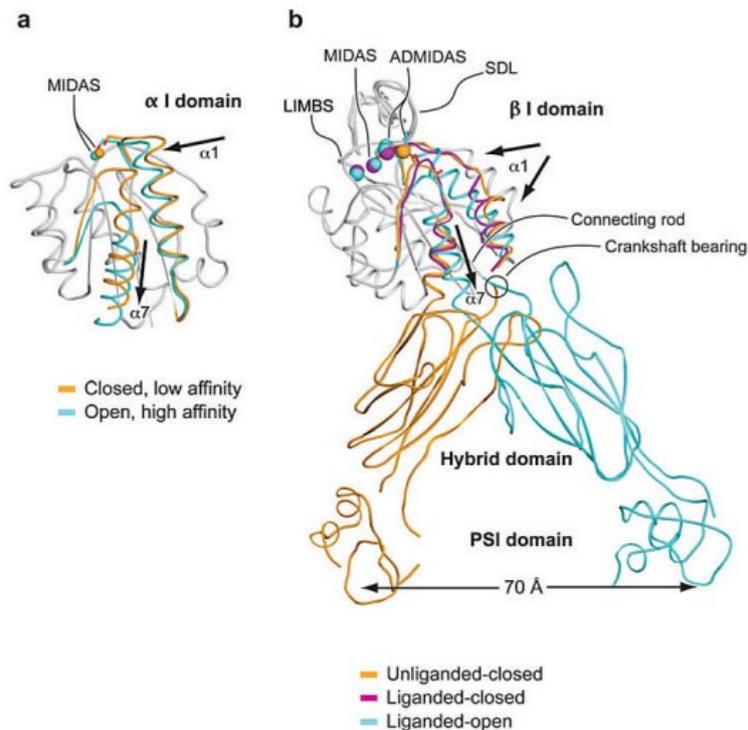


Figure 39: Conformational changes and transmission allostery by α and β I domains. (a) The α I domain: non-moving segments of the backbone are shown as a gray worm- The moving segments of the backbone and the MIDAS metal ions are closed (gold) and open (cyan). The direction of movement is indicated with arrows (IJLM and 1IDO PDB codes). (b) The β I domain and its linkage to the hybrid and plexin/semaphoring/integrin (PSI) domain. Non-moving segments of the β I backbone are shown as gray worm. Moving segments and metal ions are color coded as in legend. Directions of α 1- and α 7-helices are shown with arrows (1U8C, 1L5G and 1TXV PDB codes) (B. H. Luo et al., 2007; B. H. Luo & Springer, 2006).

Integrin α I and β I domains coordinate acidic residues in their ligands, such as the Asp side chain of RGD, through a Mg^{2+} ion held in a metal ion-dependent adhesion site (MIDAS). β I domains also contain flanking Ca^{2+} ions coordinated by residues in the adjacent to MIDAS (ADMIDAS) and synergistic metal ion binding site (SyMBS) (Xia & Springer, 2014). The ligand-binding headpiece of integrins has two states:

- A high-affinity, open conformation with the β -subunit hybrid domain swung out at its interface with the β I domain.
- A low-affinity, closed conformation with the hybrid domain swung in.

The altered conformations at the β I-hybrid domain interface are transmitted to an ~ 3 Å rearrangement in the MIDAS- and AdMIDAS- coordinating β I domain β 1- α 1 loop, which alters affinity by 1000 fold (Figure 38-39).

SyMBS and AdMIDAS have both important roles in regulating ligand binding affinity. The AdMIDAS is a negative regulatory site responsible for integrin inhibition by high concentration of Ca^{2+} and for activation by Mn^{2+} , which competes with Ca^{2+} for binding to ADMIDAS, with a mechanism still unknown (Xia & Springer, 2014).

Three metal ions are at the heart of the ligand binding site in integrin β -subunits. Furthermore, motions at many domain-domain junctions in integrins regulate ligand binding affinity.

A recent structural study revisited the role of Ca^{2+} at the AdMIDAS (Nagae et al., 2012): the $\alpha 5\beta 1$ headpiece was characterized at 2.9 Å in resolution bound to the SG/19 Fab, that stabilizes the closed conformation of $\beta 1$ integrins by binding to the β I-hybrid domain interface. Soaking with RGD peptide

induced movement of the β I domain β 1- α 1 loop and α 1 helix, resulting in a conformational state intermediate between closed and open. A decrease in electron density for the ADMIDAS Ca^{2+} ion was also observed.

Role in cancer processes of some IN phenotypes

Regulation of integrins by inside-out signals is widely used in physiology and abused or impaired in diseased states: in 2013 it was proved that high α 5 β 1 integrin expression is associated with a worst clinical prognosis of cervical cancer (Wang et al., 2013), and specifically the α V β 3 (CD51/CD61) integrin is a known receptor for tumor angiogenesis and metastasis, inflammation and bone resorption (J. Xiong et al., 2001). Ligand-mimetic RGD-based compounds act as competitive antagonist and have been developed for treatment of cancer (Mould et al., 2014; Reardon & Cheresch, 2011; Sheldrake & Patterson, 2014). However, many current RGD-based anti-integrin drugs have low specificity and act primarily as agonists for all integrins recognizing the RGD motif (Van Agthoven et al., 2014b). In contrast, some monoclonal antibodies present high specificity and are probably easier to use and produce; for example, the high specificity of LM609 for α V β 3 and the absence of selection for activated conformational states makes it a promising antagonistic candidate for future work directed toward α V β 3-positive cancerous tumors and as a competitive binder against RGD-containing viruses targeting integrin α V β 3 as a receptor for infection. Finally, the specificity of LM609 for α V β 3 raises the possibility of developing antibody- drug conjugates directing therapeutic compounds to α V β 3 integrin-expressing cell-types (Borst et al., 2017). In contrast with RGD-based inhibitors, also Ca^{2+} (but not Mg^{2+}) was able to greatly increase the dissociation rate of integrin–FN complexes. The effect of Ca^{2+} was overcome by co-addition of Mn^{2+} , but not Mg^{2+} . A stimulatory anti- β 1 monoclonal antibody (mAb) abrogated the effect of Ca^{2+} on α 5 β 1–FN complexes; conversely, a function-blocking mAb mimicked the effect of Ca^{2+} (Mould et al., 2014).

The platelet and megakaryocyte-specific integrin, α IIb β 3 (also commonly known as GPIIb-IIIa) has received considerable attention as a drug target due to its requisite role in platelet aggregation, a significant mechanism in the mediation of arterial thrombosis. GPIIb-IIIa is the key receptor which mediates platelet aggregation by adhesive cross-linking of the divalent plasma proteins fibrinogen and von Willebrand factor. Adhesive protein ligands are recognized and bind to GPIIb-IIIa through the specific amino acid tripeptide sequence RGD contained within surface loops of each of these adhesive ligands or a similar adhesive ligand sequence (KQAGDV) found on the carboxyl terminus of the γ -chain of fibrinogen (Ferrari et al., 2004; Scarborough & Gretler, 2000).

1.1.3. Fibronectin-Integrins interactions in depth

Fibronectin network formation

α 5 β 1 is the first integrin identified involved in the fibronectin network formation. It localized to the sites where fibronectin fibrils are in contact with the cell, and antibodies to α 5 or β 1 can inhibit the polymerization in fibroblast cultures. Overexpression of α 5 in CHO cells increases the deposition of fibronectin in the matrix. Mouse embryos deficient in the α 5 gene still contained extracellular fibronectin networks, indicating that also other integrins could promote the polymerization process. Also, α V β 3 and α IIb β 3 have shown this potential, but the first with less efficiency and the latter is not proven yet to act in vivo processes. Expression of α V β 1 or α 4 β 1 in CHO cells did not promote fibronectin polymerization, proving that mere binding of fibronectin to the cell surface by any integrin is not sufficient. A connection of β subunit to actin filaments is known to be required for fibronectin

fibril formation on the cell surface. Other factors influencing the ability of integrin to promote fibril formation are the recognition site preference in FN (RGD, LDV, etc.) and the receptor-ligand binding affinity.

The polymerization process can be separated in two phases: initiation (nucleation) and extension. In one study (R. Winklbauer & C. Stoltz, 1995) it has been reported that fibrils grow only at one end, indicating that they are polarized, with the direction of the growth reflecting the migration of the cell. It is unknown if new protomers are added to the fibril at the cell surface or somewhere else (**Figure 39**).

Several regions of fibronectin are involved in different phases of polymerization, through binding to cell surface components or to a neighboring fibronectin molecule. Binding of RGD region to integrins may be the initial event which triggers the subsequent reactions, maybe by inducing an altered fibronectin conformation. The inability of recombinant fibronectin lacking the RGD motif to initiate fibril formation supports this hypothesis. However, it is unclear if soluble fibronectin can bind to $\alpha 5\beta 1$. Fibronectin fragments rather than the intact protein are commonly chosen for affinity isolation of $\alpha 5\beta 1$. It can be hypothesized that folding at the suggested hinge between 9FnIII and 10FnIII in soluble fibronectin would prevent $\alpha 5\beta 1$ from reaching both the synergy site and the RGD loop. Thus, modulation of fibronectin structure by some other interaction may precede binding to $\alpha 5\beta 1$.

The isolated N-terminal 1-5FnI has been shown to bind cell layers at sites of fibronectin fibril formation, and to inhibit further incorporation of intact fibronectin molecules. Deletion or disruption of this functional unit in recombinant fibronectin prevents incorporation of the protein into fibrils. The N-terminal domain has been suggested to bind to a non-integrin "matrix assembly receptor", of nature still unknown. Instead, it's probable that 1-5FnI binds either to conformationally altered fibronectin or to a site on $\alpha 5\beta 1$ which becomes exposed after binding RGD ligands. In either case, this domain would be involved in fibril extension rather than initiation. The postulated conformational changes could result directly from the interactions or from a pulling force of the actin filament system. The importance of the cytoskeleton has been demonstrated by use of cytochalasin B, and fibronectin fibrils running between two cells appear to be under tension as indicated by immunofluorescent staining.

A cryptic site in 1FnIII can bind to 1-5FnI, and a similar interaction was described between unfolded 10FnIII and 1FnIII. At present it is unknown which of these potential interactions mediate inter- and/or intramolecular bridging between fibronectin domains. The difficulties in studying these events is a major reason why the exact mechanism of the polymerization process still is unsolved (Johansson et al., 1997)

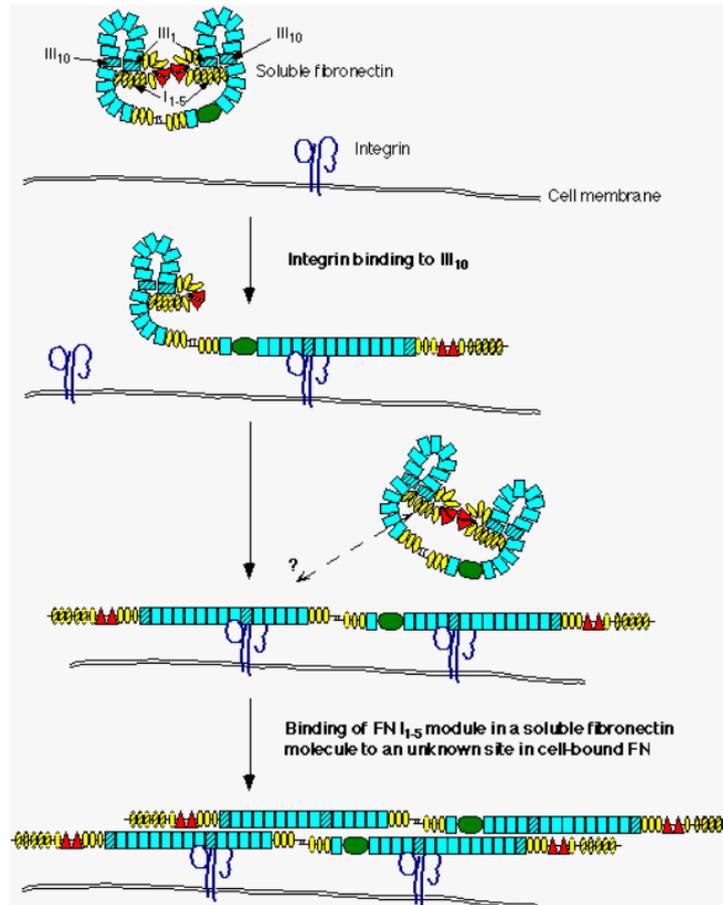


Figure 39: hypothetical fibronectin network formation model. Soluble fibronectin first binds to integrin on the cell membrane through the RGD site in 10FnIII, eventually helped by PHSRN in 9FnIII. The interaction induces a conformational change in the fibronectin molecule which becomes elongated. The cell bound fibronectin exposes binding sites (unidentified) for I1-5 in another soluble fibronectin molecule. This interaction brings the second molecule to “open up” and another FN molecule would bind through the 1-5FnI modules to the attached fibronectin, in a chain reaction that could bring the polymerization process ahead. Pulling force from the actin filament system may be required for the postulated conformational changes. Domains in fibronectin which may be involved in interactions with other fibronectin domains are marked with black stripes (Johansson et al., 1997).

Site of binding and mechanisms

Two regions in each FN subunit possess cell binding activity: 9-10FnIII and 14FnIII-V (V=variable region). The RGD sequence, the most important recognition site for about half of the integrins, is located in 10FnIII. Fibronectin can be a ligand for a dozen members of the integrin receptor family. The type III repeat 10 of FN (10FnIII) has been shown to bind integrins via an RGD sequence. In the crystal structure of 7-10FnIII, 9FnIII presents a synergy site (PHSRN) for additional interactions with integrins on the same face of FN as the RGD segment. Fibronectin binds to integrin receptors on cells, partly via an RGD sequence in 10FnIII. In 10FnIII, these residues are in a loop between strands F and G, which exhibits considerable conformational flexibility relative to the more highly conserved beta-sheet regions. Inspection of sequence alignments clearly shows that the functional RGD residues have been incorporated into a stable protein scaffold provided by the FnIII module. Studies of fibronectin binding to integrins identified an 11-residue peptide from 9FnIII that contains a novel integrin-binding site. The peptide corresponds to parts of the C' and E strands and their connecting loop in the FnIII consensus fold. Examination of the structure of the FnIII pair from neuroglian suggests that the C'-E

loop in 9FnIII would be solvent-exposed and available for interaction. Such secondary sites are of interest, as they may be involved in the specificity of ligand/integrin recognition. The alternatively spliced IIICS domain and the adjacent heparin-binding domain provide a second major cell-binding region of fibronectin, which recognizes the integrins $\alpha 4\beta 6$, and $\alpha 4\beta 7$. The IIICS region encodes two integrin-binding sites specific to cell type. A key minimal sequence, Leu-Asp-Val, has been identified in the IIICS domain, but the tripeptide exhibits only a fraction of the activity of the full-length spliced sequence, suggesting that synergistic sites are involved in this fibronectin/integrin (**Figure 40**) interaction (Potts & Campbell, 1994). The affinity of the integrins for the short RDG-containing peptides varies depending on the integrin type, but for all receptors the affinity for larger protein fragments or the intact proteins is order of magnitude higher than the previous. This is due to the amino acid sequence PHSRN, located in 9FnIII, that contributes to the contact surface between ligands and integrins.

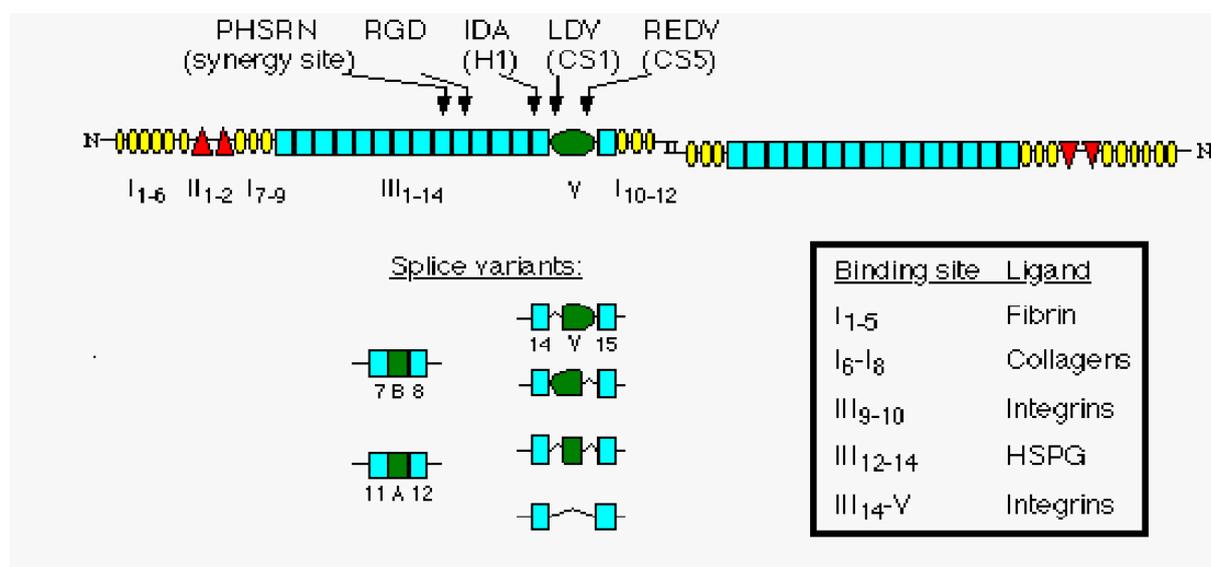


Figure 40: Integrin binding sites in FN as well as in other proteins usually contain an aspartic acid as critical amino acid (Johansson et al., 1997).

The 14FnIII-V region, or CS1 site, is recognized by $\alpha 4\beta 1$ and $\alpha 4\beta 7$ integrins in three binding sites (figure). This site has 20-fold higher affinity for integrins than the H1 and CS5 site (in the 13FnIII-14FnIII and V-10FnI region respectively). However, CS1 and CS5 can be spliced out so they could be missing in some fibronectin phenotype, while H1 is always present. The IDV and IDA sequences are the key cell binding structures in CS1 and H1 respectively, while in CS5 is been identified as REDV in human and RGDV in mice and bovine proteins. This indicated that the CS5 site presents similarity to both RGD and LDV motif (**Figure 41**) (Johansson et al., 1997).

Receptor	Main binding site in FN ²	Additional binding sites in FN	Other protein ligands
$\alpha 3\beta 1$	RGD		LN 5
$\alpha 4\beta 1$	LDV in CS1	H1, CS5, (RGD)	VCAM-1
$\alpha 5\beta 1$	RGD	PHSRN	
$\alpha 8\beta 1$	RGD		VN, TN
$\alpha V\beta 1$	RGD		
$\alpha V\beta 3$	RGD		VN, OP, vWF, FG, TSP
$\alpha IIb\beta 3^1$	RGD	PHSRN	FG, vWF, VN
$\alpha V\beta 6$	RGD		TN
$\alpha 4\beta 7$	LDV in CS1		MAdCAM-1, VCAM-1
$\alpha ?\beta 8$	CS1		

Figure 41: integrin types interacting with FN and their sites of binding (Johansson et al., 1997).

State of the art in FN-IN interactions and inhibitors knowledge RGD ligands, RGD mimetics and monoclonal antibody inhibitors

Several studies exist on the interaction between integrins and their ligands, in an attempt to understand the mechanism underlying these interactions and a way to therapeutically exploit them. $\alpha V\beta 3$ is an integrin known for being involved in angiogenesis and metastatic processes in tumoral cells. In a study the interactions between $\alpha V\beta 3$ and vitronectin, fibronectin or RGD-containing peptides were observed (Orlando & Cheresh, 1991). It can be observed that $\alpha V\beta 3$ binds to vitronectin or fibronectin in a non-dissociable manner that is independent by the actin cytoskeleton, while binding to RGD-containing peptides is specific but completely dissociable with a K_d of 9.4×10^{-7} M. Besides, chemical modification of the heterodimer with limited glutaraldehyde treatment resulted in vitronectin binding in an RGD-dependent and dissociable manner. This indicates that receptor conformational changes or specific amino acid residues proximal to the ligand binding sites are involved in the stabilization event. Further studies established that macromolecular ligands, but not RGD peptides bind to $\alpha V\beta 3$ in a stabilized manner (K_d circa 0). After macromolecular ligand binding, excess soluble ligand or RGD peptides are unable to dissociate the complex, contesting the hypothesis to use RGD-mimetics ligand to disrupt already existing IN-FN complexes. A conformational change in the integrin is necessary for binding to become stabilized (Orlando & Cheresh, 1991).

In the study carried on by Mould et al (Mould et al., 2014) it was examined the binding in real time of four different RGD-binding integrins to a fragment of fibronectin using surface plasmon resonance. It was demonstrated that binding of each integrin to fibronectin is non-reversible by RGD-based inhibitors. Formation of non-reversible state is very fast, and not related to the slow formation of a state with very low dissociation rate. Probably the formation of the non-reversible state is dependent on conformational changes within the β -I domain; perturbing these changes with allosteric inhibitors lead to rapid dissolution of integrin-ligand complexes. This is an indicator that allosteric integrin antagonists may have a better therapeutic effect than competitive (RGD-based) integrin antagonists. The recombinant integrin observed were $\alpha 5\beta 1$, $\alpha V\beta 1$, $\alpha V\beta 3$ and $\alpha V\beta 6$, analyzing their interaction with a 50k fragment of FN. Each integrin interacted with 50k FN, but with different dissociation rates. In each case a biphasic dissociation phase was observed, in which the dissociation rate in the early stage of dissociation was higher than in the later part. This could be either due to the integrin samples containing a mixture state with a form having high association rate and an another one with low dissociation rate, or that a form of integrin with high dissociation rate is gradually converted in the form with lower dissociation rates. Although it wasn't possible to distinguish from selected-fit and

induced-fit models based, it was found that at longer time periods after the dissociation phase, the dissociation rate became lower, so it does suggest a important role of a time-dependent stabilization of a high-affinity state. The effect of different reagents on pre-formed IN-50kFN complexes and a truncated form of $\alpha 5\beta 1$ -Fc (TR $\alpha 5\beta 1$ -Fc), containing just the head region of $\alpha 5$ and $\beta 1$ subunits, was investigated (Mould et al., 2014):

- RGD-based peptides

Observing the effect of the two RGD-based inhibitor peptides, cilengitide and cRGD (cyclic peptide containing respectively the RGD(N-Me)V and the GRGDSP sequence from fibronectin), it resulted that no RGD-based antagonist increased the dissociation rate of IN-FN complexes, even with 10-fold higher inhibitor concentrations. Since the inhibitors were added 60 seconds after the end of dissociation phase, this means that is not necessary the presence of the time-stabilized state to these complexes to be resistant to the effects of the antagonists. Since also the IN-FN complexes with high initial dissociation rates ($\alpha V\beta 1$ and $\alpha V\beta 6$) were unaffected by these inhibitors, the non-reversibility may also be a feature of low-affinity states. Using TR $\alpha 5\beta 1$ -Fc it can be observed that its interaction with FN50k is weaker, with a lower K_d than $\alpha 5\beta 1$ -Fc with 50k. Here too it is showed that the component at low K_d increase with time, indicating a time-dependent stabilization of the high-affinity state (supporting the two-phase model). RGD-based inhibitors are also incapable of dissociating the complexes.

- EDTA (Ethylenediaminetetraacetic acid)

EDTA is known for disrupting the integrin-ligand interactions by removal of the metal ion at MIDAS. EDTA could rapidly disrupt $\alpha V\beta 1$ - and $\alpha V\beta 6$ -FN complexes, and more slowly $\alpha 5\beta 1$ -FN complexes. In contrast, $\alpha V\beta 3$ -FN complexes were resistant to EDTA (with only a 1.5-fold increase in K_d). As control, no binding to FN50k was observed if EDTA was added to each integrin before association. TR $\alpha 5\beta 1$ -FN50k complexes are dissociated by EDTA, but as already stated, not by RGD-based inhibitors: these suggests that the head region alone is sufficient to make ligand binding not reversible by RGD-based peptides; beside, since the peptides were added during the initial phase of dissociation, the high-affinity form is not required for the non-reversibility.

- Bivalent cations on reversibility of $\alpha V\beta 3$ -FN by RGD-base peptides

Since $\alpha V\beta 3$ -FN is resistant also to dissociation to EDTA, the effect of bivalent cations was evaluated on the reversibility of $\alpha V\beta 3$ -FN complexes by RGD-based peptides. This resistance indicated that MIDAS cations are very stable bound to complex and hence unlikely to be lost or replaced by other bivalent ions during the dissociation phase.

Testing whether $\alpha V\beta 3$ -FN complexes were reversible by RGD-based peptides if Mg^{2+} and Ca^{2+} rather than Mn^{2+} , it was found that both Mg^{2+} and Ca^{2+} enabled fibronectin binding, but the dissociation rate K_d was much higher than Mg^{2+} . While complexes formed in solution containing Mg^{2+} could not be dissociated by RGD-based peptides, complexes formed in Ca^{2+} were weakly dissociated by RGD-based peptides (K_d in presence of CRGD or cilengitide 3-times fold higher than in absence of peptides); EDTA had very small effect on K_d in both cases of Mg^{2+} and Ca^{2+} .

- Dissociation of IN-FN complexes by Ca^{2+}

Ca^{2+} is capable of influencing whether RGD-based inhibitors could dissociate $\alpha\text{V}\beta 3$ -FN complexes, it was tested if Ca^{2+} itself is capable of influencing the dissociation of IN-FN complexes formed in presence of Mn^{2+} . Ca^{2+} is highly effective at dissociating $\alpha\text{V}\beta 3$ -FN complexes, obtaining a K_d 50-fold higher than complexes in running buffer alone on in running buffer with no bivalent cations. Ca^{2+} was also very effective at dissociating other IN-FN complexes, while addition of buffer with Mg^{2+} had little or no effect on K_d . Varying the concentration of Ca^{2+} it is possible to assess the concentration required for half-maximal inhibition and the apparent affinity of Ca^{2+} binding: the results suggested that Ca^{2+} acts through a site of moderate affinity (apparent K_d circa 0.4 mM). Similar results were obtained for $\alpha 5\beta 1$, but the results are not so accurate because of a marked increase in the K_d in absence of Mn^{2+} .

- Blocking of the effect of Ca^{2+} by Mn^{2+} , but not Mg^{2+}

Then it was investigated if the effect of Ca^{2+} could be overcome by co-addition of Mn^{2+} or Mg^{2+} during dissociation phase. Both $\alpha 5\beta 1$ and $\alpha\text{V}\beta 3$ complexes were tested and resulted in the fact that Mn^{2+} , but not Mg^{2+} , is able to nullify the effect of Ca^{2+} on dissociation rate.

To determine the concentration of Mn^{2+} required for obtaining the 50% recovery of the dissociation rate in Mn^{2+} alone, different concentrations of Mn^{2+} were added with 0.2 mM Ca^{2+} , and then the dissociation rates were fitted against the concentration of Mn^{2+} to obtain the IC_{50} value (concentration of Mn^{2+} at which the effect of Ca^{2+} is inhibited by 50%). Similar results were obtained for both complexes, but for $\alpha 5\beta 1$ the ability of Mn^{2+} to replace the gradually lost MIDAS ion may contribute to recovery of the dissociation rate. Higher concentration of Ca^{2+} required higher concentration of Mn^{2+} to block the effects. The simplest explanation is that the two ions compete for binding to the same site on IN: this site can bind both with similar moderate affinity, but not Mg^{2+} .

- Stimulatory mAB directed against β subunit

On the hypothesis that Ca^{2+} could have an allosteric effect on IN-FN complexes, because it is known to inhibit Mn^{2+} supported ligand binding by favouring the closed form of the β -I domain, it was postulated that a mAB which binding favours the open form of β -I could overcome the ability of Ca^{2+} to dissociate IN-Fn complexes. It was tested mAB TS2/16, that binds in the $\alpha 2$ helix region of the $\beta 1$ subunit β -I domain. The results showed its ability to greatly reduce the ability of Ca^{2+} to dissociate complexes. Trying to prove also the inverse effect, it was tested a function-blocking mAB 13, allosteric inhibitor of $\beta 1$ integrins with contrary effect to TS2/16, and that also binds in the $\alpha 2$ helix region of the β -I domain. mAB 13 was able to increase the dissociation rate by circa 10 times fold, much like the second inhibitory mAB, 4B4.

A study on the structure of extracellular $\alpha\text{V}\beta 3$ with a cyclic pentapeptide ligand Arg-Gly-Asp-[D-Phe]-[N-methyl-Val-] and the pro-adhesive cation Mn^{2+} reveals that the pentagonal peptide inserts into a crevice between the propeller and β -I domains on the integrin head. The RGD sequence makes the main contact area with the integrin, and each residue participates in the interaction. The Arg and Asp side chains point in the opposite directions, exclusively contacting the propeller and β I domains respectively. The five C α atoms of the cyclic peptide form a slightly distorted pentagon, that takes a more regular without integrin-binding. Thus, the distortion of the peptide ring is likely related to the

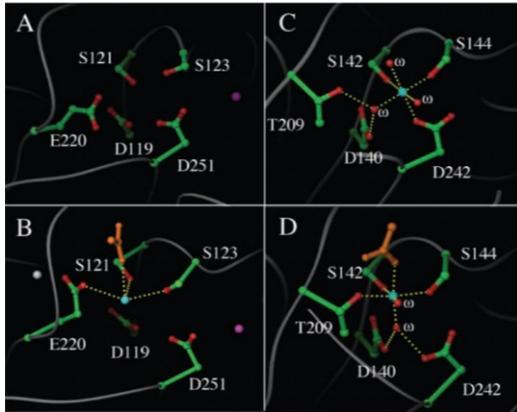


Figure 43: Diagram of the MIDAS motif in βI (A and B) and αI from CD11b (C and D). (A) and (B) MIDAS residues (single letter abbreviations: S, Ser; E, Glu; D, Asp; T, Thr) in unliganded (A) and liganded (B) βI . Coordinating side chains are shown in ball-and-stick representations with oxygen atoms in red, carbon in green; the ligand aspartate is in gold. In addition to the ligand aspartate, the Mn^{2+} (cyan) in the βI MIDAS is coordinated directly with the hydroxyl oxygens of Ser121 and Ser123 and with one carboxylate oxygen from Glu220. The carboxyl oxygens of Asp119 and Asp251 of βI lie within 6Å of the metal ion and likely mediate additional contacts through water molecules similar to the liganded forms of αI (D). The Mn^{2+} ion at ADMIDAS (magenta) is present in (A) and (B). The Mn^{2+} ions at MIDAS and at LIMBS (cyan and gray, respectively) are only present in (B). (C) and (D) MIDAS residues in unliganded (C) and liganded (D) αI from CD11b. The metal ion (cyan) is present in both. Water molecules are labeled ω ; the pseudo-ligand glutamate is in gold. Hydrogen bonds and metal ion coordination are represented with dotted yellow lines (J. P. Xiong et al., 2002).

The one difference between the sites is the replacement of a conserved Thr, which contacts the cation in liganded αI , with Glu220 in βI .

In the unliganded $\alpha V\beta 3$ -Mn structure, the Glu220 side chain intrudes into the MIDAS site, approaching the space where a cation would bind. Thus, it appears to reduce affinity for cations at MIDAS through steric hindrance. In the liganded $\alpha V\beta 3$ -RGD-mn structure, the Glu220 side chain occupies a different position, allowing accommodation of a cation at MIDAS.

In addition to incorporating Mn^{2+} at MIDAS when liganded, βI also incorporates a second Mn^{2+} ion in a region 6 Å away from MIDAS. This site is defined as ligand-associated metal binding site (LIMBS), and is formed by the other carboxylate oxygen of Glu220, the side chains of Asp158, Asn215 and Asp217, and the carbonyl oxygens of Asp217 and Pro219. Although the LIMBS Mn^{2+} does not contact the ligand, coordination of Mn^{2+} depends on it. Asp158 and Glu220 occupy different positions in the unliganded structure and thus the coordination sphere for LIMBS does not exist.

Most likely the role of LIMBS is to stabilize the reoriented Glu220 and to add conformational stability and structural rigidity to the ligand-binding surface.

Binding of the pentapeptide ligand is associated with tertiary and quaternary changes in $\alpha V\beta 3$ -Mn:

- Changes in the tertiary structure involve βI , affecting mainly its $\alpha 1$ - $\alpha 2$ loops and helices and the $\alpha 2$ -C', F- $\alpha 7$ and B-C ("ligand-specificity") loops. The movements seem to be casually linked to the top of helix $\alpha 1$ which approaches MIDAS, permitting contacts with both MIDAS cation and ligand through Ser121, Tyr122 and Ser123. In the complex, the backbone amide and carbonyl oxygens of Tyr122 directly contact the ligand Asp, and both serine side chains coordinate the MIDAS cation. Thus, $\alpha 1$ is fastened to the ligand-MIDAS assembly within the complexes. The ADMIDAS (adjacent to MIDAS) cation moves in synchrony with $\alpha 1$, because it is primarily coordinated by $\alpha 1$ residue Asp126 and Asp127; this changes its coordination sphere slightly from that of the unliganded structure, since its coordination by the carbonyl oxygen of Met335 is replaced by a carboxylate oxygen from Asp251. Most of the remaining

structural changes can be seen as indirectly caused by the shift of $\alpha 1$: $\alpha 1'$ directly follows $\alpha 1$, and $\alpha 2$ and the top flank of $\alpha 7$ flank $\alpha 1'$.

The ligand-specificity region also approaches the ligand. This movement can be related to a salt bridge in this region, between Asp179 and Arg214. Arg214 is near the ligand Asp, and it does not form a salt bridge to Asp179 in the unliganded structure. The functional implications of these changes are reflected by the location in the $\alpha 1$ - $\alpha 2$ segment of βI if epitopes both for activation and inhibitory monoclonal antibodies.

These tertiary changes in the liganded form of βI resemble those seen in liganded αI . In αI distinguishing feature of its transition to the liganded state is a 10 Å downward shift of the COOH-terminal $\alpha 7$ helix with realignment of its hydrophobic contacts. However, the position of $\alpha 7$ helix in liganded βI does not change. Likely, the activation in αI and βI is achieved by different mechanisms. Reorientation of the COOH-terminal $\alpha 7$ helix, perhaps in response to inside-out signaling, makes αI ligand-competent in an allosteric manner. In βI , where the movement of the $\alpha 7$ helix is more improbable, reorientation of the MIDAS Glu220 residue results in a ligand-competent form by unblocking MIDAS. A second interpretation could be that the conformation of βI in the unliganded $\alpha V\beta 3$ -Mn and $\alpha V\beta 3$ -Ca structures represents a ligand-competent state of the I-type domain, captured in the context of an integrin heterodimer, and thus the tertiary changes in βI would be ligand-induced.

- Quaternary rearrangements in the integrin head region are also observed in the complex. The interface between βI and the αV propeller undergoes a small change, with the two domains moving closer together at the peptide-binding site. In addition, the propeller undergoes a small rotation at the propeller-thigh interface, with βI moving in concert. As in the case of G-proteins, ligand binding to βI alters its orientation relative to the propeller.

It's important to notice that both tertiary and quaternary changes are observed in an integrin in the presence of its smallest recognition unit, even within the constrained crystal lattice. Natural integrin ligands are significantly larger, structurally diverse and often multivalent, thus the conformational rearrangement could represent a minimalist view of the changes in the receptor that take place during integrin-ligand interactions (J. Xiong et al., 2001).

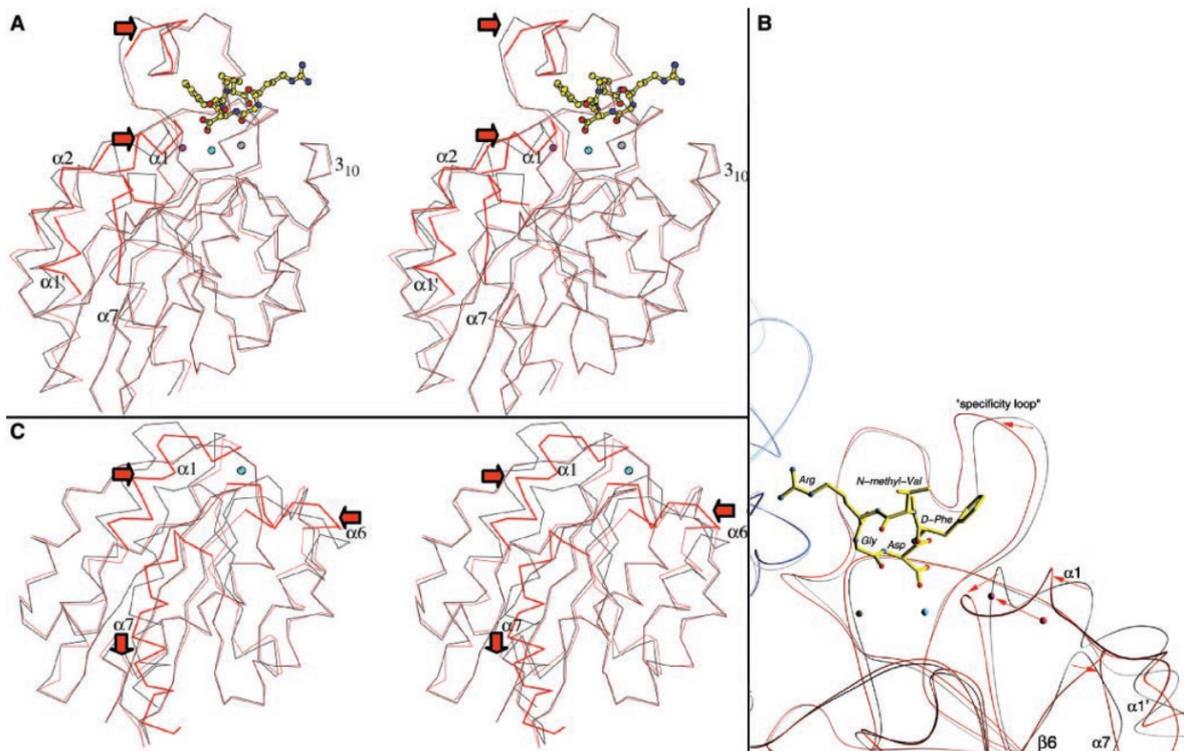


Figure 44: Ligand-induced structural changes in $\beta 1$ in comparison with those of $\alpha 1$ (from CD11b). (A) Superposition, in stereo, of the $\alpha V\beta 3$ -Mn (gray) and $\alpha V\beta 3$ -RGD-Mn (red) structures. The superposition is based on the C α atoms of the central β -sheet [43 atoms per structure; root mean square deviation (RMSD), 0.42 Å]. Residues of $\alpha V\beta 3$ -RGD-Mn with a distance of more than 1.5 Å to corresponding residues of $\alpha V\beta 3$ -Mn are shown with thicker red lines. The major structural changes in $\beta 1$ involve helices $\alpha 1$, $\alpha 1'$, $\alpha 2$, the F- $\alpha 7$ loop, and the ligand-specificity region. (B) Magnified view of the rearrangements at the ligand-binding site in $\beta 1$. Superposition of the propeller and βA domains of $\alpha V\beta 3$ -Mn (gray) and $\alpha V\beta 3$ -RGD-Mn (αV , blue; $\beta 3$, red) is based on the C α atoms of the αV propeller domain. The directions of protein movements (including the 4 Å displacement of Mn 2^+ at ADMIDAS) are indicated by red arrows. This view differs from (A) by a rotation of 180° around a vertical axis. (C) Superposition, in stereo, of the "liganded" (red) and "unliganded" forms of αI from the CD11b integrin. The metal ion sphere at MIDAS is in cyan. The superposition is based on the C α atoms of the central β -sheet (43 atoms; RMSD // 0.43 Å). Residues of liganded αI with a distance of more than 1.5 Å to corresponding residues of unliganded αI are shown with thicker red lines. The major structural changes in αI involve helices $\alpha 1$, $\alpha 7$, the F- $\alpha 7$, and E- $\alpha 6$ loops. Arrows (red) indicate the direction of the major protein movements in each case (J. P. Xiong et al., 2002).

In another study surface plasmon resonance was used to characterize the affinity of Fab LM609 to the ectodomain of human $\alpha V\beta 3$ integrin (Borst et al., 2017). The LM609 antibody specifically recognizes $\alpha V\beta 3$ integrin and inhibits angiogenesis, bone resorption, and viral infections in an arginine-glycine-aspartate- independent manner. LM609 entered phase II clinical trials for the treatment of several cancers and was also used for $\alpha V\beta 3$ -targeted radioimmunotherapy. LM609 binds at the interface between the β -propeller domain of the αV chain and the $\beta 1$ domain of the $\beta 3$ chain, near the RGD-binding site, of all observed integrin conformational states; this result suggests complex formation with Fab LM609 does not require the $\alpha V\beta 3$ ectodomain to be in a defined conformational state. Integrating these data with fluorescence size-exclusion chromatography, was demonstrated that LM609 sterically hinders access of large ligands to the RGD-binding pocket, without obstructing it. In all resolved conformations, LM609 interacts with the $\alpha V\beta 3$ headpiece domain, probably at the interface between the β -propeller domain of the αV chain and the $\beta 1$ domain of the $\beta 3$ chain, site recognized by the typical triangular shape of the headpiece apex region. Attempted complex formation between Fab LM609 and $\alpha IIb\beta 3$ and visualization using negative-staining EM yielded no detectable complexes, corroborating the statement that LM609 specifically binds to $\alpha V\beta 3$ but not to $\alpha IIb\beta 3$ (Wu et al., 1998). LM609 Fab does not appear to overlap with the RGD-binding pocket in any

of the integrin conformations observed, and as expected fluorescein isothiocyanate-tagged RGD peptide, FITC-GRGDSPK (FITC-RGD) could also bind to $\alpha\text{V}\beta\text{3}$, with the same magnitude, following pre-incubation of $\alpha\text{V}\beta\text{3}$ with a saturating amount of Fab LM609: both ligands could simultaneously bind to $\alpha\text{V}\beta\text{3}$, ruling out competitive binding to the same epitope. Although both RGD-independent (LM609) and RGD-dependent (e.g., fibronectin) binding involve contacts with the β propeller of the $\alpha\text{V}\beta\text{3}$ subunit and the βI domain of the $\alpha\text{V}\beta\text{3}$ subunit, distinct sets of interactions mediate attachment in each case. However, previous in vitro studies demonstrated LM609 inhibits fibronectin to $\alpha\text{V}\beta\text{3}$. Superposition of type 7-10FnIII modules (3T1W, 1FNF (Leahy et al., 1996)) onto the corresponding domain of $\alpha\text{V}\beta\text{3}$ -10FnIII complex structure (4MMX PDB entry, (Van Aghthoven et al., 2014b)) reveals that binding of LM609 to the integrin headpiece would likely sterically hinder subsequent attachment of fibronectin due to the expected clashes with 8FnIII and 9FnIII (**Figure 46**). Previous reports also showed LM609 inhibition on the binding of fibrinogen to $\alpha\text{V}\beta\text{3}$ integrin in vitro (Wu et al., 1998). These considerations suggest that this is mediated through a mechanism of steric hindrance due to the comparably large sizes of fibrinogen and fibronectin (Mosesson, 2005), and the ability of the FITC-RGD peptide to bind to LM609-bound integrin $\alpha\text{V}\beta\text{3}$. This inhibition mechanism allows rationalization of the properties of this antibody that has been shown to inhibit angiogenesis and bone resorption in vivo, as well as infection of RGD-containing viruses such as parechovirus and Kaposi's sarcoma-associated herpesvirus (Garrigues et al., 2008). Because Fab LM609 could bind to all observed integrin conformational states, it has the potential to interact with $\alpha\text{V}\beta\text{3}$ receptors prior to inside-out signal activation, which could be leveraged for future therapeutic strategies. Similar to what was found in a previous study on 17E6 anti- body (Mahalingam et al., 2014), the full-length LM609 antibody might also interfere with integrin clustering or promote integrin internalization, therefore enhancing therapeutic effects relative to the monovalent Fab LM609. The specificity of LM609 for $\alpha\text{V}\beta\text{3}$ raises the possibility of developing antibody- drug conjugates directing therapeutic compounds to $\alpha\text{V}\beta\text{3}$ integrin-expressing cell-types (Borst et al., 2017).

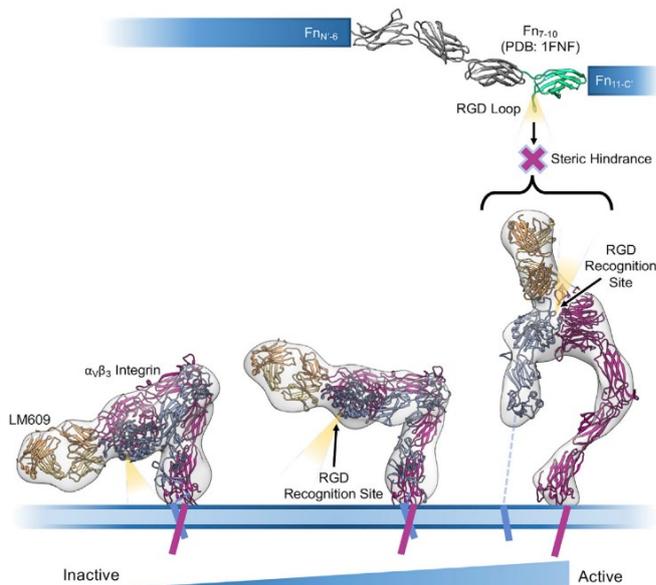


Figure 45: Model different conformational states of $\alpha\text{V}\beta\text{3}$ integrin bound to LM609, and the latter sterical hindrance that could inhibit fibronectin binding in the RGD binding site (Borst et al., 2017)

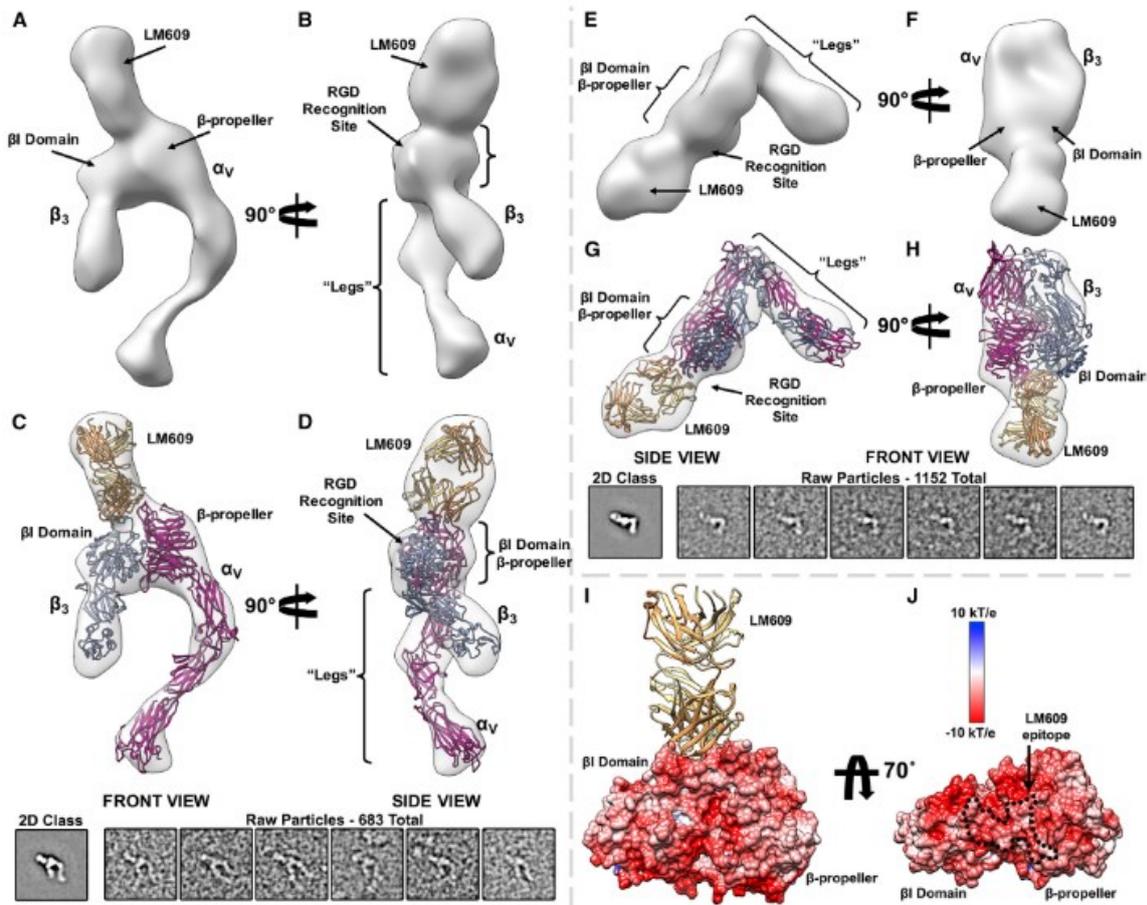


Figure 35: Structure of the Fab LM609/αVβ3 Integrin Complex Determined by Single-Particle EM (A and B) Two orthogonal views of a random conical tilt 3D reconstruction featuring an αVβ3 integrin extended state. (C and D) Corresponding views related to (A) and (B) showing the fit of the pseudo-atomic model obtained with Rosetta (ribbon) into the reconstruction (transparent surface). The lower panel depicts the associated 2D class average and a few raw particles (low-pass filtered to 20-Å resolution) used for the RCT reconstruction of the integrin/LM609 extended state. (E and F) Two orthogonal views of a random conical tilt 3D reconstruction featuring an αVβ3 integrin bent state. (G and H) Corresponding views related to (E) and (F) showing the fit of the pseudo-atomic model obtained with Rosetta (ribbon) into the reconstruction (transparent surface). The lower panel depicts the associated 2D class average and a few raw particles (low-pass filtered to 20-Å resolution) used for the RCT reconstruction of the integrin/LM609 bent state. (I) Ribbon diagram of the Fab LM609 bound to the integrin headpiece. Only the β propeller of the αV subunit and the β1 domain of the β3 subunit are shown in surface representation colored by electrostatic potential. LM609 is represented as yellow and orange ribbons corresponding to the heavy and light chains, respectively. (J) Corresponding view related to (I) rotated by 70°. The approximate epitope of the LM609 Fab is indicated with black dashed lines and features a pronounced negative electrostatic potential. αVβ3 is represented as fuchsia (αV) and light blue (β3) ribbons in panels (C), (D), (G), and (H).

From observations on binding of a high-affinity form of 10FNIII (h10FNIII) to αVβ3 (Van Agthoven et al., 2014b), it can be noted that unlike wild-type 10FNIII and cyclic RGD-based peptides, h10FNIII did not induce activation-specific conformational LIBS (ligand induced binding site) mAb epitopes in the integrin N- and C-terminal domains, and it reduced LIBS expression induced by both the activating cation Mn^{2+} and the constitutive (mutational) activation of αVβ3. Furthermore, it didn't substantially alter the hydrodynamic behavior of the soluble αVβ3 ectodomain. The interface between αVβ3-hFN10 interface was surprisingly modest even relative to the αVβ3-cilengitide interface and was distinguished by contacts with the glycan at Asn266 of the α-subunit propeller domain. These contacts significantly contributed to the adhesion function of cellular αVβ3. The glycan at Asn266 is conserved in the fibronectin receptor α5β1, and mutation of the equivalent residue in α5 (N275Q) impaired α5β1-mediated cell adhesion, suggesting that the glycan contact may also be in the α5β1-FN interface.

This interface should also be stronger than the $\alpha\text{V}\beta\text{3}$ -FN interface because of an interaction of FN-type III domain 9 with the α -subunit propeller, an interaction that is not used by $\alpha\text{V}\beta\text{3}$ (Bowditch et al., 1994). This could also explain the greater susceptibility of the smaller $\alpha\text{V}\beta\text{3}$ -FN interface to force-induced binding or unbinding events, which would make it more suitable than the more extensive $\alpha\text{5}\beta\text{1}$ -FN interface for mediating dynamic outside-in signal transduction (Roca-Cusachs et al., 2009). Structural and mutational studies support a critical role for the novel Trp1496-Tyr122 π - π interaction in 'locking' the integrin in an inactive conformation: first, removing the Trp1496 side chain from hFN10 resulted in a domain that acted as wtFN10. LIBS were also induced by hFN10 binding to cellular $\alpha\text{V}\beta\text{3}$ lacking the Tyr122 side chain; second, changing the orientation of the Trp1496 side chain in hFN10 so that it no longer faced also led to induction of LIBS when cellular $\alpha\text{V}\beta\text{3}$ bound hFN10/B. These data strongly argue that blocking the inward movement of the α1 helix towards MIDAS is sufficient to halt the associated tertiary changes in the βI domain that led to outside-in signaling. Thus, altering the side chain orientation of Trp1496 by design or selection of its local environment can substantially affect the tertiary or quaternary changes induced by binding of RGD-based ligands. For example, a cyclized form of the RGD-containing loop of hFN10 could be modified by changing the orientation of the tryptophan side chain in eptifibatide or by replacing d-phenylalanine with d-tryptophan in a modified form of cilengitide. The critical βI Tyr122 is also conserved in $\alpha\text{5}\beta\text{1}$ and β2 integrins, which, like $\alpha\text{V}\beta\text{3}$, are drug targets (Van Agthoven et al., 2014b).

In a study on $\alpha\text{5}\beta\text{1}$ (Xia & Springer, 2014) soaking experiments were performed using two different RGD ligands: a linear RGD peptide, a cyclic GRGDSP peptide, and a cyclic ACRGDGWCG peptide. The results show that binding of cyclic RGD peptide presents a 20-fold higher affinity than a linear RGD peptide and induces conformational change in the β1 -subunit βI domain to a state that is intermediate between closed (low affinity) and open (high affinity), while binding of a linear RGD peptide induces no shape shifting. The high affinity of cyclic peptide may be because cyclization together with backbone hydrogen bonds stabilizes a favorable conformation for binding to $\alpha\text{5}\beta\text{1}$. Alteration of the cyclic peptide Trp to Ala or mutation of α5 Trp-157 to Ala greatly diminished cell attachment, successfully predicting the interaction of the peptide Trp with the α5 Trp-157. We see a T-shaped edge-to-center interaction known to stabilize aromatic interactions between the two Trp residues. Despite the importance of α5 Trp-157 in binding the cyclic peptide, it is not important in binding fibronectin, which has a Pro residue in the equivalent position. In the closed conformation the SyMBS binds Ca^{2+} , the MIDAS binds Mg^{2+} , and the ADMIDAS binds Ca^{2+} . Linear peptide binding induces shape shifting when Ca^{2+} is depleted during soaking: Ca^{2+} bound to the adjacent to metal ion-dependent adhesion site (ADMIDAS), at the locus of shape shifting, moves and decreases in occupancy, correlating with an increase in affinity for RGD measured when Ca^{2+} is depleted. Comparisons in affinity between four-domain and six-domain headpiece constructs suggest that flexible integrin leg domains contribute to conformational equilibria. High resolution views of the hybrid domain interface with the plexin–semaphorin–integrin (PSI) domain in different orientations show a ball-and-socket joint with a hybrid domain Arg side chain that rocks in a PSI domain socket lined with carbonyl oxygens (Xia & Springer, 2014). The four-domain fragment of the $\alpha\text{5}\beta\text{1}$ headpiece containing the α5 -subunit β -propeller domain and the β1 -subunit plexin–semaphorin–integrin (PSI), hybrid, and βI domains in closed conformation shows that the orientation between the β1 -subunit βI and hybrid domains is very similar to that in the five-domain $\alpha\text{5}\beta\text{1}$ headpiece-SG/19 Fab complex (Nagae et al., 2012) and the $\alpha\text{IIb}\beta\text{3}$ closed conformation in absence of ligand (Jieqing Zhu Bing-Hao Luo, Patrick Barth, Jack Schonbrun, David Baker, 2009; Zhu et al., 2010), but is distinct from that in open $\alpha\text{IIb}\beta\text{3}$ conformation found in the presence of ligand (Xiao et al., 2004). Sg/19 binds to the α βI -hybrid domain interface and inhibits ligand binding allosterically by stabilizing this interface in its close, low affinity conformation. Sg/19 does not enforce an artificial orientation between the βI -hybrid domains but binds to the same

overall closed hybrid- β I domain orientation as seen in the absence of Fab. The closed conformation at the ligand binding site, the α 1-helix, and the N-terminal end of the α 1'-helix are identical in the presence and absence of SG/19. However, it does induce a small conformational changes in its epitope: the hybrid domain β W- β I loop is flexible in absence of SG/19, while in this case it becomes ordered and forms a substantial part of the epitope. SG/19 binds to the C-terminal end of the α 1'-helix, a C-terminal continuation of the α 1-helix. Sg/19 slightly pushes α 1' inward to enable the α 1' Arg154 and Arg155 side chains to reorient and form part of the epitope. In the GRGDSP soaked structure the RGD Arg guanido group formed strong, bidentate, charged hydrogen bonds to α 5 Asp-227; in contrast, hydrogen bonding to α 5 Gln-221 was through two intermediate water molecules. The RGD Asp side chain coordinated the MIDAS Mg^{2+} ion using one oxygen and formed a weak, 3.4-Å hydrogen bond to the β I domain β 1- α 1 loop backbone through the other oxygen. RGD binding induced no movement in the β 1- α 1 loop or α 1-helix, which were completely superimposable on the native structure, namely in closed "state 1" as defined by α IIb β 3 structures. Moreover, there was no loss of Ca^{2+} at the ADMIDAS; Ca^{2+} electron density after soaking with RGD peptide was as strong as before. In absence of added Ca^{2+} , RGD induced substantial movement of the β 1- α 1 loop and α 1-helix, enabling the non-MIDAS-binding Asp carboxyl oxygen to form a stronger, 3.0-Å hydrogen bond to the β 1- α 1 loop backbone. Accompanying the movement of the β 1- α 1 loop, the side chain of Ser-134 moved to directly coordinate the MIDAS Mg^{2+} ion, displacing a water at the MIDAS seen with RGD in the presence of Ca^{2+} . Slight differences were also seen at the RGD Arg guanido group, which now formed a direct hydrogen bond to α 5 Gln-221. Moreover, density for Ca^{2+} at the ADMIDAS was diminished, whereas there was little change in Ca^{2+} density at the SyMBS or Mg^{2+} density at the MIDAS. These differences and the shape shifting in the presence and absence of Ca^{2+} directly demonstrate that Ca^{2+} binding to the ADMIDAS stabilizes integrins in the low-affinity, closed conformation, prevents shifting from the closed conformation and Ca^{2+} increases affinity for RGD peptide by 10-fold (Xia & Springer, 2014) (**Figure 51**). The intermediate state and diminishment of ADMIDAS Ca^{2+} electron density obtained soaking with RGD peptide in the absence of Ca^{2+} were essentially identical to that obtained by Nagae et al. (PDB entry 3VI3, 3VI4) when both Mg^{2+} and Ca^{2+} were omitted from the RGD solution. Then, it was proven that the disulfide cyclized peptide ACRGDGWCW binds α 5 β 1 strongly and specifically: binding of the cyclic peptide in the presence of Ca^{2+} induced shape shifting of the β 1- α 1 loop and α 1 helix similar to the linear peptide in the absence of Ca^{2+} . The cyclic peptide also induced movement of the Ser-134 side chain into direct coordination with the MIDAS Mg^{2+} . Importantly, the ADMIDAS Ca^{2+} ion moved with the β 1- α 1 loop backbone and with its Ser-134 carbonyl oxygen, which forms one ADMIDAS coordination. Also the α 1-helix with its Asp-137 and Asp-138 side chains moved with the ADMIDAS Ca^{2+} ion, forming two further ADMIDAS coordinations. Electron density for the moved ADMIDAS Ca^{2+} was good, but weaker than the SyMBS Ca^{2+} electron density. Shape shifting with the cyclic peptide and Ca^{2+} , the linear peptide without Ca^{2+} depletion, and RGD in the SG/19 complex were similar in extent to one another, and approximated intermediate state 4 of α IIb β 3. The residues between the two disulfide-bonded cysteines in ACRGDGWCW were well-defined and show the basis for high-affinity binding. Cyclization decreases the conformational space accessible to a solvated peptide and thus lowers the entropic cost of becoming ordered when binding to a receptor. Moreover, the cyclic peptide conformation was stabilized by three backbone hydrogen bonds, including a 2.8-Å hydrogen bond with excellent geometry between the Arg-3 carboxyl oxygen and Trp-7 amide nitrogen. The cyclic peptide RGD moiety assumed a bound conformation very similar to that in linear RGD peptides. Importantly, the face of the cyclic peptide Trp side chain formed a T-shaped interaction with the edge of the α 5 Trp-157 side chain (T-shaped interactions are energetically favored by electrostatic interactions between the electron-rich edges and electron-poor centers of aromatic rings) (**Figure 48-49-50-52**).

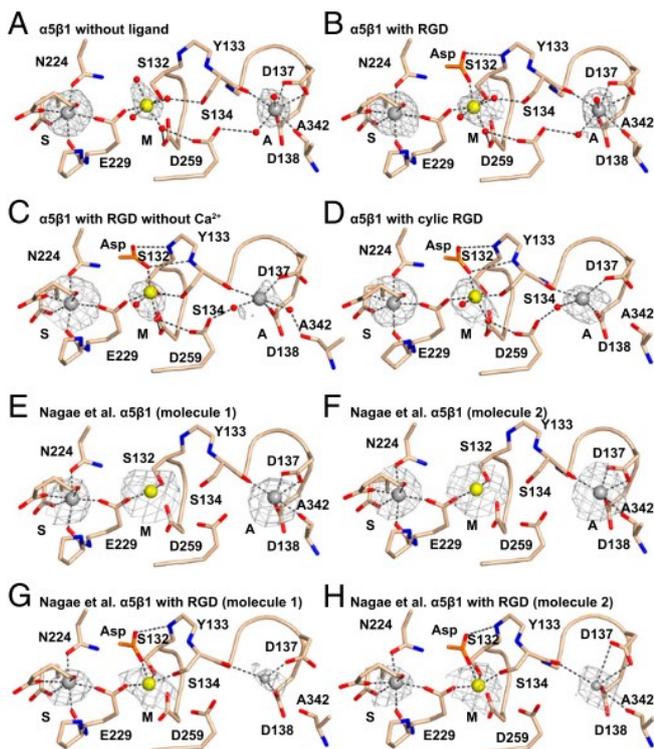


Figure 48: $\beta 1$ domain metal binding sites. (A–H) The indicated structures are in identical orientations. The backbones and side chains of metal-coordinating residues are colored wheat with red oxygens and blue nitrogens. Waters are shown as small red spheres. Ca^{2+} (silver) and Mg^{2+} (gold) are shown as large spheres. Putative ADMIDAS Ca^{2+} ions not included in molecular models in G and H are shown as small spheres. Simulated-annealing omit map $F_o - F_c$ electron density contoured at 2.5σ is shown as mesh around metal ions.

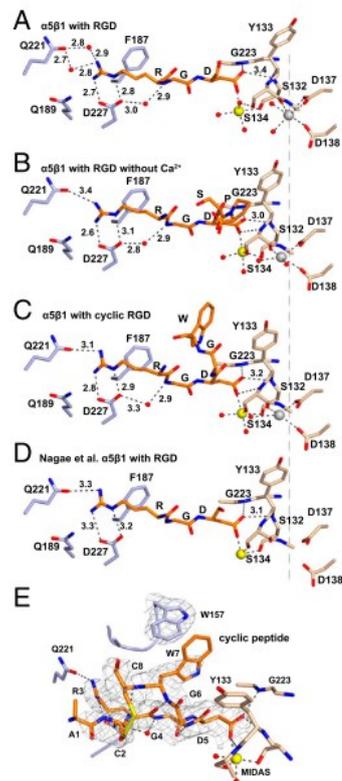


Figure 49: RGD-binding pocket. (A–D) Vertical alignment in identical orientations with a dashed line for comparison. Residues of $\alpha 5$ (light blue), $\beta 1$ (wheat), and ligand (RGD and two more C-terminal residues if built in the model) (orange) are shown in stick representation. Ca^{2+} (silver; ADMIDAS) and Mg^{2+} (gold; MIDAS) are shown as spheres. Waters are shown as small red spheres. Selected distances are shown in Å. (E) Binding of the cyclic peptide. The color scheme is as in A–D. All cyclic peptide residues visible in electron density are shown. $2F_o - F_c$ density at 1σ is shown as mesh for the cyclic peptide and $\alpha 5$ Trp-157.

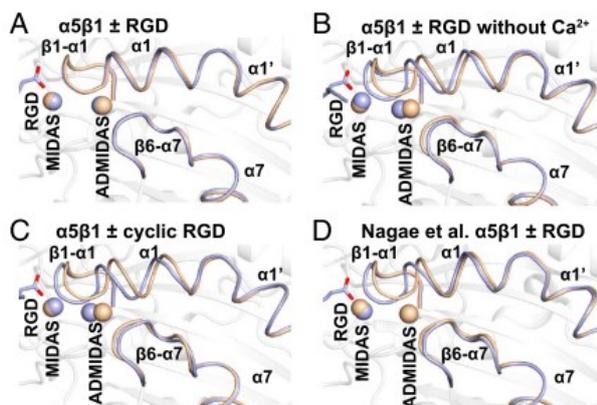


Figure 50: RGD-induced conformational rearrangements in the $\beta 1$ domain. (A–D) The indicated unliganded (wheat) and RGD-soaked (light blue) structures are shown in identical orientations. Arg and Asp residues and disulfide bonds are shown in stick representation. Metals at MIDAS and AdMIDAS are shown as spheres.

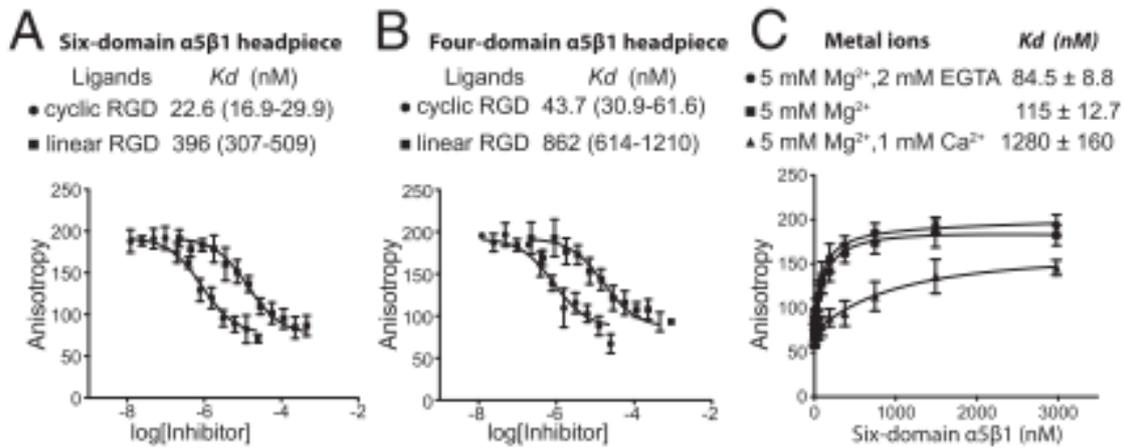


Figure 51: RGD binding affinity. (A and B) Binding of six-domain $\alpha 5 \beta 1$ (A) and four-domain $\alpha 5 \beta 1$ (B) headpieces (200 nM) to linear and cyclic RGD peptides measured by competition with 2 nM fluorescent cyclic peptide using fluorescence anisotropy in 1 mM Mn^{2+} and 0.1 mM Ca^{2+} (C) Binding of the six-domain $\alpha 5 \beta 1$ headpiece to 2 nM fluorescent cyclic RGD peptide in different divalent cations. Data show mean \pm SD of triplicate samples.

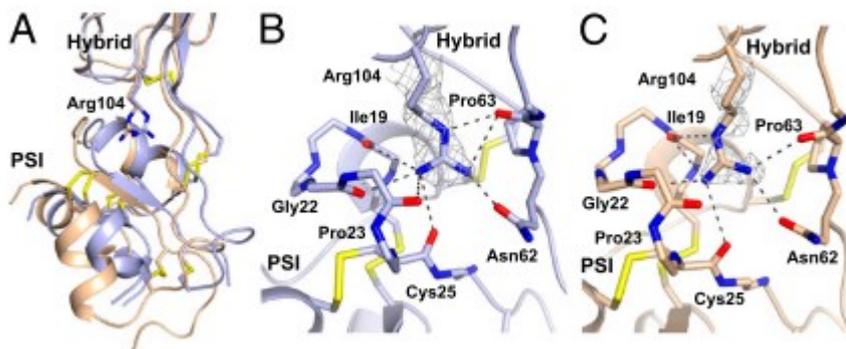


Figure 12: Rocking at the hybrid–PSI interface. The change in PSI orientation is shown after superimposition on the hybrid domain. The rocking Arg-104 side chain in the hybrid domain shown in stick representation binds to a socket in PSI lined with carbonyl groups. Dashed lines indicate hydrogen bonds. (A) Combined view. (B and C) Individual detailed views, with $2F_o - F_c$ density for Arg-104 contoured at 1σ shown in mesh. Unliganded and linear RGD peptide + Ca^{2+} $\alpha 5 \beta 1$ structures have carbons in light blue and wheat, respectively, and disulfide bonds in yellow.

Headpiece structures of the PDB entries (4WJK, 4WK0, 4WK2, 4WK4) provide insights into the flexibility of knee-proximal thigh, PSI, and EGF1 domains in the extended conformation of integrins. These domains show some flexibility in crystal structures of intact integrin ectodomains (Dong et al., 2012; Sen et al., 2013; Xie et al., 2010; Zhu et al., 2008), in which the headpiece is bent at α -subunit thigh–calf-1 and β -subunit EGF1–EGF2 junctions and is in contact with the lower legs. Upper leg–lower leg interfaces and the closeness of the α -subunit and β -subunit knees help to stabilize the orientation of knee-proximal domains in the bent conformation. However, these constraints are removed upon integrin extension. PSI and EGF1 extend from the same end of the hybrid domain.

Interaction between PSI and EGF1 in $\beta 2$ integrins maintains similar hybrid–PSI and hybrid–EGF1 orientations in different $\alpha \beta 2$ ectodomain crystal forms and $\beta 2$ leg fragments (Sen et al., 2013; M. L. Shi et al., 2005, 2007; Xie et al., 2010). In contrast, in $\alpha 4 \beta 7$ headpiece 3.1-Å structures, both PSI and EGF1 domains were missing from electron density (Yu et al., 2012). Electron density for the PSI domain is generally poorer than for other integrin domains in both bent ectodomain and head-piece structures. Although this is true in Xia & al structures as well, the overall higher-resolution (1.78 and

1.85 Å) $\alpha 5\beta 1$ headpiece structures show clear density at the hybrid–PSI interface and reveal the basis for flexibility.

When the PSI domain moves relative to the hybrid domain, the side chain of hybrid domain residue Arg-104 rocks and changes rotamer to maintain an extensive network of hydrogen bonds to a socket in the PSI domain that is lined with backbone carbonyl oxygens. The Arg functions as a ball in a ball-and-socket joint between two tandem domains (“rocking Arg in a carbonyl socket” (Xia & Springer, 2014)). The closest analogous interface is found in selectins between the lectin and EGF domains, where the α -Imino group of the N-terminal Trp of the lectin domain pivots in an EGF domain carbonyl cage in selectin allostery (Springer, 2009). The thigh domain in integrins shows considerable flexibility at its interface with the β -propeller domain (Sen et al., 2013; Xie et al., 2010; Yu et al., 2012, 2013). The hybrid, PSI, and EGF1 domains come much closer to the thigh domain in the closed than open headpiece conformation and thus limit the range of orientations available to the thigh domain (figure ,B). Therefore, the presence of the thigh domain will entropically favor the open headpiece conformation. Similarly, the presence of the EGF1 domain in our six-domain but not four-domain headpiece will limit thigh domain flexibility in the closed conformation and favor the open headpiece conformation (**Figure 53**).

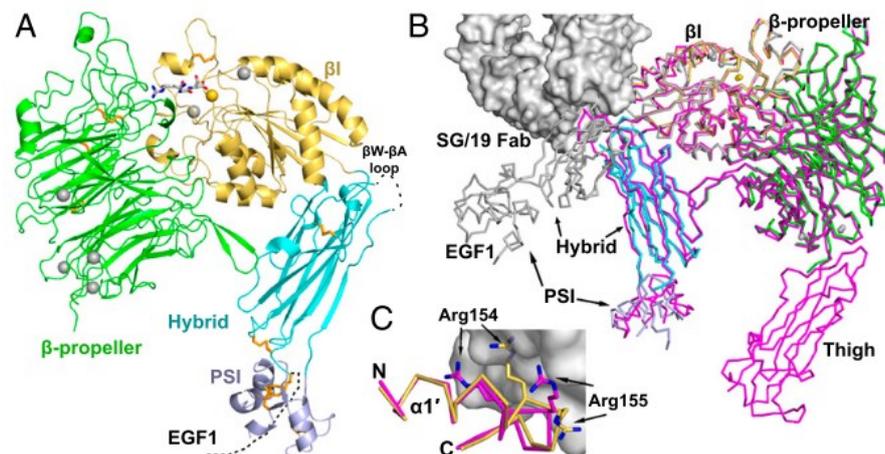


Figure 53: Crystal structure of the $\alpha 5\beta 1$ integrin head-piece. (A) Ribbon diagram of $\alpha 5\beta 1$. The β -propeller, βI , hybrid, and PSI domains and RGD are in green, yellow orange, cyan, light blue, and gray, respectively. Ca^{2+} (silver) and Mg^{2+} (gold) are shown as spheres. Disulfide bonds (orange) and RGD are shown in stick representation. (B) Integrin headpiece comparisons. Four-domain $\alpha 5\beta 1$, color-coded as in A, five-domain $\alpha 5\beta 1$ (magenta) ((Nagae et al., 2012)), and open $\alpha IIb\beta 3$ (silver) ((Springer et al., 2008)) are superimposed and shown in Ca ribbon; for clarity, only $\beta 3$ of $\alpha IIb\beta 3$ is shown. SG/19 Fab bound to five-domain $\alpha 5\beta 1$ ((Nagae et al., 2012)) is shown as a surface. (C) Binding of SG/19 to the βI domain $\alpha 1'$ -helix. Structures are colored and shown as in B, with the Arg-154 and -155 side chains also shown and the view rotated $\sim 180^\circ$ around a vertical axis relative to B (Xia & Springer, 2014).

The reason why cyclic RGD peptide induce $\beta 1$ - $\alpha 1$ loop and $\alpha 1$ -helix movement toward the open conformation whereas the linear peptide doesn't could not be inferred by occupancy of RGD, because it is essentially complete in both cases, and differing occupancy cannot explain the differences here. The causes could be various and mainly three: first, although the bound conformations of RGD appear similar, there may be a significant difference in binding enthalpy to drive the shift; second, RGD slides slightly in its groove from $\beta 1$ toward $\alpha 5$ between closed state 1 and intermediate state 4 (same as $\alpha IIb\beta 3$ (Zhu et al., 2013)), with a magnitude that is greater for cyclic RGD than linear RGD in state 4, particularly at the RGD Gly and Asp, and may be enhanced by the peptide Trp interaction with $\alpha 5$ Trp-157, consistent with greater movement in this direction in the cyclic peptide of the backbone following the peptide Asp; finally in the linear peptide, the C-terminal Ser and Pro are disordered in the presence

of Ca^{2+} , and have a spatial probability distribution that includes space that the side chain of Tyr-133 in the $\beta 1$ - $\alpha 1$ loop must occupy in intermediate state 4. This entropic barrier to $\beta 1$ - $\alpha 1$ loop shape shifting by the peptide C terminus is removed by cyclization. Notably, the peptide C terminus becomes ordered by the close approach of the $\beta 1$ Tyr-133 side chain in the Ca^{2+} -deficient complex, despite lower resolution compared with the Ca^{2+} -replete linear peptide complex.

Good density for the AdMIDAS Ca^{2+} in the intermediate state with the cyclic peptide shows that movement of AdMIDAS-coordinating residues does not abolish their interaction with Ca^{2+} ; indeed, in our Ca^{2+} -depleted and -replete intermediate structures, as well as in models with Ca^{2+} added to the Nagae et al. structures, coordination to the Ser-134 backbone and Asp-137 and Asp-138 side chains is present in all cases. However, it is very probable the validity of the suggestion that breakage of the AdMIDAS coordination to the Ala-342 backbone in the $\beta 6$ - $\alpha 7$ loop, the first consequence of AdMIDAS movement, is associated with weaker AdMIDAS Ca^{2+} density (Nagae et al., 2012). According to the law of mass action, Ca^{2+} at the AdMIDAS can only favor the closed conformation over the open conformation if it is bound with higher affinity in the closed than open conformation, demonstrated by the fact that Ca^{2+} removal destabilizes the closed conformation and enables RGD to induce integrin movement toward the open state. It appears that the AdMIDAS Ca^{2+} coordination in the closed state is closest to pentagonal bipyramidal, with bidentate coordination in the closed state by the carboxyl oxygens of Asp-137 in the plane of the pentagon. Whereas there are seven AdMIDAS Ca^{2+} coordinations in our 1.78- and 1.85-Å closed states, there are only four or five AdMIDAS Ca^{2+} coordinations in the 2.5-Å intermediate-state structures, correlating with lower Ca^{2+} occupancy. Furthermore, Asp-137 coordination is monodentate.

In open conformation structures of $\alpha 1\beta 3$, coordination of Ca^{2+} or Mn^{2+} at the AdMIDAS is more octahedral, with monodentate coordination by the $\beta 3$ Asp corresponding to $\beta 1$ Asp-137 (Springer et al., 2008; Zhu et al., 2013). The open AdMIDAS coordination also differs from closed because AdMIDAS coordination by a $\beta 6$ - $\alpha 7$ loop backbone carbonyl is replaced by an Asp side chain that coordinates through a water to the MIDAS Mg^{2+} ion (Xiao et al., 2004; Zhu et al., 2013). The preference of Ca^{2+} for pentagonal bipyramidal over octahedral coordination, together with changes in metal ligands between closed and open conformations, may explain why Ca^{2+} removal increases affinity for RGD and enables RGD to induce $\beta 1$ - $\alpha 1$ loop and $\alpha 1$ -helix shape shifting. Conversely, Mg^{2+} and Mn^{2+} favor octahedral over pentagonal bipyramidal coordination, and their substitution for Ca^{2+} at the AdMIDAS may explain why Mn^{2+} and Mg^{2+} /EGTA (EGTA selectively chelates Ca^{2+}) are general activating agents for integrins (Dransfield et al., 1992; Gailit & Ruoslahti, 1988; Harding, 2001). These results are in agreement with previous findings that Ca^{2+} exerts an inhibitory effect on $\beta 1$, $\beta 2$, and $\beta 7$ integrins by binding to the AdMIDAS (Chen et al., 2003; Dransfield et al., 1992; Kamata et al., 2002; Mould, Barton, Askari, Craig, et al., 2003). However, the ideas on how Mn^{2+} activates integrins are controversial, particularly for $\alpha 5\beta 1$, for which has been suggested that activation is achieved by binding of Mn^{2+} to the MIDAS (Mould et al., 2002; Mould, Barton, Askari, Craig, et al., 2003). It is possible that Mn^{2+} activates $\alpha 5\beta 1$ by replacing both Mg^{2+} at the MIDAS and Ca^{2+} at the AdMIDAS; however, this fails to explain why, for most integrins, removal of Ca^{2+} by EGTA in the presence of excess Mg^{2+} activates as well as Mn^{2+} . An early $\alpha \beta 3$ crystal structure showed an intermediate RGD-bound state similar to Xia and Springer (Xia & Springer, 2014), with separation between the MIDAS and AdMIDAS metal ions, and reported that in contrast to integrin $\alpha 1$ domains, $\beta 1$ domains did not convert between open and closed states (J. P. Xiong et al., 2002), and so $\alpha 5\beta 1$ mutational studies were interpreted with the assumption that the MIDAS and AdMIDAS were independent of one another. However, it is known now that integrin $\beta 1$ domains undergo opening to a high-affinity state, and that the AdMIDAS metal ion moves closer to the MIDAS metal ion. In the open state, the side chain of $\beta 3$ Asp-251, equivalent

to $\beta 1$ Asp-259, forms a direct coordination to the ADMIDAS metal ion and an indirect coordination through a water to the MIDAS metal ion (Xiao et al., 2004).

Mutation of $\beta 1$ MIDAS residue Asp-130 abolishes the ability of Mn^{2+} to activate, but not the ability of Ca^{2+} to inhibit, binding of 12G10 antibody to an activation epitope on the $\beta 1$ domain $\alpha 1'$ -helix involving residues Arg-154 and Arg-155 (Mould et al., 2002). Considering the current structural knowledge, this suggests either that Mn^{2+} activates by binding to the MIDAS, or that Mn^{2+} activates at the AdMIDAS in a MIDAS- dependent manner, for example, that AdMIDAS movement into close proximity to the MIDAS with 12G10 epitope exposure requires a metal at the MIDAS. In further support for Mn^{2+} activation at the MIDAS, mutation of ADMIDAS residues Asp-138 and Asp-139 to Ala was found to still permit activation by Mn^{2+} and inhibition by Ca^{2+} (Mould, Barton, Askari, Craig, et al., 2003). However, the D137A mutation lowered ligand binding affinity more than D138A, raising the possibility that D138A might not completely inhibit metal binding at the AdMIDAS; furthermore, ligand binding experiments showing stimulation by Mn^{2+} used the partially active D138A mutant and the stimulatory antibody 12G10, which might have acted like crutches to enable a crippled AdMIDAS to bind to Mn^{2+} when in its active, MIDAS-proximal conformation. A definitive understanding of how Mn^{2+} and Mg^{2+} /EGTA activate will require further work (Xia & Springer, 2014)

A study dated to 1996 (Paul Mould et al., 1996) was focused on the effect of ligand recognition by $\alpha 5\beta 1$ on the binding of a mAb13, that inhibits $\beta 1$ integrin function. Ligand (fibronectin fragment or GRGDS peptide) decreased the binding of mAb 13 to $\alpha 5\beta 1$: at high ligand concentrations, approximately 50% of the total integrin bound mAb 13 with >50-fold lower affinity than in the absence of ligand. The concentration of ligand required for half-maximal inhibition of antibody binding was independent of anti- body concentration, suggesting that ligand acts as an allosteric inhibitor of mAb 13 binding. Analysis adding the activating 9GE7 found that it increases the maximum level of ligand binding by 2-fold, indicating that up to 50% of the total integrin could not bind ligand without 9EG7 stimulation. Further addition of RGD ligand induced the effect that essentially all of the integrin bound mAb 13 with very low or zero affinity, demonstrating that mAb 13 recognizes an epitope that is dramatically attenuated by ligand occupancy in the ligand-occupied form of $\alpha 5\beta 1$. So, since mAb 13 preferentially recognizes the unoccupied conformation of the integrin, the antibody may inhibit ligand binding by stabilizing the unoccupied state of $\alpha 5\beta 1$. Ligand behaves as an allosteric inhibitor of antibody binding, suggesting that mAb 13 does not perturb integrin function by direct competition for the ligand binding site. Since ligand behaved as an allosteric inhibitor of antibody binding (mAb 13 and ligand appeared to recognize non-overlapping sites on the $\beta 1$ subunit) it seemed possible that ligand and mAb 13 binding could occur simultaneously, but antibody and ligand binding were inversely correlated and mutually exclusive, suggesting that mAb 13 may induce a conformational change that results in displacement of ligand from the integrin. An alternative explanation of the exclusive binding could be that the binding of mAb 13 to $\alpha 5\beta 1$ inactivates the integrin, rendering it incapable of ligand recognition. However, in experiments in which integrin was preincubated with mAb13, antibody binding could be reversed by the CCBD fragment or GRGDS peptide, demonstrating that antibody-occupied integrin was still capable of binding ligand. Based on the model of the allosteric but mutually exclusive in binding between oxygen and 2,3-diphosphoglycerate to hemoglobin, that hypothesize that binding of an antibody that prefers a conformation could shift the equilibrium of polypeptides in conformational equilibria toward a specific state (Baldwin & Chothia, 1979; BENESCH & BENESCH, 1969; Sachs et al., 1972). Hence, it has been suggested that some anti-LIBS mAbs stimulate integrin function by shifting the conformational equilibrium between inactive and active states of the integrin in favor of the active state (Frelinger et al., 1991). By evidence that the conformations of the active (competent to bind ligand) and ligand-occupied states are not identical and the fact that apart from the exposure of LIBS on the ligand-occupied state, at least one site on $\alpha 5\beta 1$ is strongly attenuated after

ligand recognition, it must take in account the existence of at least three conformational states of an integrin (inactive, active, and ligand-occupied), and it is more accurate to consider that anti-LIBS mAbs can stimulate ligand binding by shifting a conformational equilibrium between the active (unoccupied) state and the ligand-occupied state in favor of the ligand-occupied state. Nevertheless, the conformation of the ligand-occupied state appears to be closely related to that of the active state, and hence some anti-LIBS antibodies may also be able to shift a conformational equilibrium between inactive and active states of the integrin in favor of the active state.

The observation that the short (5-mer) peptide was capable of blocking mAb 13 binding implied that this inhibition was not due to long-range steric hindrance of anti-body binding by ligand. About 50% of the total integrin showed low affinity binding of mAb 13 in the presence of ligand, whereas the affinity of the remainder appeared unaltered. We interpreted the first of these two populations as integrin that is capable of attaining an active conformation (and is, therefore, competent to bind ligand), whereas the second population is locked in an inactive conformation.

Interestingly, 9EG7 only had a small effect on the apparent affinity of ligand binding; in contrast, other stimulatory anti- β 1 mAbs such as 8A2 appear to increase the apparent affinity of ligand binding but have little effect on the maximal level of ligand binding (Faull et al., 1993). It is possible that the population of ligand-occupied integrin that fails to bind mAb 13 corresponds to an additional conformational state of α 5 β 1, perhaps with ligand irreversibly bound to the integrin (Muller et al., 1993).

An important corollary of the observations done is that antibodies whose epitopes are preferentially expressed on the unoccupied state may be able to inhibit integrin function by shifting a conformational equilibrium in favor of the unoccupied state (provided that ligand binding is reversible). The epitope recognized by the inhibitory anti- β 1 mAb 13 is attenuated by ligand occupancy, i.e. the antibody binds with a much lower affinity to the ligand-occupied state than to the unoccupied state of the integrin. Since the receptors in the ligand-occupied state may still weakly express these epitopes, the converse of LIBS epitopes are called ligand-attenuated binding site epitopes and mAb 13 is designated an anti-ligand-attenuated binding site mAb; the perturbation induced by binding of mAb13 acts in the opposing manner to anti-LIBS mAbs . 13 inhibits ligand binding by stabilizing the conformation of the unoccupied state, rather than by sterically blocking a ligand binding site. Indeed, it is difficult to envisage how mAb 13 could act as a direct competitive inhibitor of ligand binding (i.e. block a site directly involved in ligand binding), whereas antibodies with epitopes very close to that of mAb 13 (such as 12G10, TS2/16, and 8A2) strongly stimulate integrin function (Kovach et al., 1992; Luque et al., 1994; Mould et al., 1995). Since all known inhibitory mAbs recognize the same region of the β 1 subunit (Takada & Puzon, 1993), other inhibitory mAbs may perturb ligand binding by the same mechanism as mAb 13; in agreement with this suggestion, it was also proven that ligand also appears to act as an allosteric inhibitor of P4C10 binding (Stewart et al., 1995).

It will be important to determine if these antibodies inhibit ligand binding directly, or like mAb 13, recognize epitopes that are attenuated by ligand occupancy. Data for mAb 13 highlight the difficulty of attempting to localize ligand binding sites on integrins by epitope mapping of inhibitory mAbs, since some of these antibodies may recognize sequences that regulate integrin activity, rather than sites that are directly involved in ligand recognition. It also remains to be determined if activating and inhibitory anti- β 1 mAbs mimic the function of biological activators or inhibitors of integrin function. Because the region of the β 1 subunit that contains the epitopes for these mAbs is crucially involved in the regulation of integrin-ligand interactions, it is important the focus studies on mapping epitopes in this region (Paul Mould et al., 1996).

Similar to other integrin types, $\alpha\text{IIb}\beta\text{3}$ headpiece was crystallized in the two conformational state, open and closed. The closed headpiece conformation is the same as the one found for the crystals of the complete $\alpha\text{IIb}\beta\text{3}$ ectodomain without Fab (Zhu et al., 2008) and in crystals of the $\alpha\text{IIb}\beta\text{3}$ headpiece complex with 10E5 Fab (Zhu et al., 2010). Similar, the open conformation of $\alpha\text{IIb}\beta\text{3}$ headpiece is found in cacodylate pseudo-ligand or RGD mimetics, with or without bound 10E5 Fab (Springer et al., 2008; Xiao et al., 2004). Fabs have no carbohydrate and minimal flexibility compared with integrins and can stabilize crystal lattices and improve resolution. 10E5 Fab binds to the β -propeller domain, far from shape-shifting portions of the β subunit and has no influence on β -propeller or integrin conformation as shown by comparisons among many views of the $\alpha\text{IIb}\beta\text{3}$ structure. In the work of Zhu et al. (Zhu et al., 2013) integrin $\alpha\text{IIb}\beta\text{3}$ is analyzed and crystallized in 8 different conformational state, obtained with different RGD and metal ion concentrations, to evaluate the mechanism of the headpiece opening in integrin activation. In the table below there is a list with the different conditions relative to conformational state.

By comparison with the previous crystallography structure of the closed conformations of $\alpha\text{V}\beta\text{3}$ ectodomain ((J. P. Xiong et al., 2002), PDB entry 1L5G) and $\alpha\text{5}\beta\text{1}$ headpiece ((Nagae et al., 2012), PDB entries 3VI3 and 3VI4) it can be noted that RGD/ $\alpha\text{V}\beta\text{3}$ structure in Mn^{2+} (PDB 1L5G) is closest to state 5. The RGD/ $\alpha\text{5}\beta\text{1}$ structure in Mg/Ca is closest to states 3 and 4. These structures showed one intermediate trapped at a particular point in the shape-shifting pathway, in contrast to the eight RGD-bound states in a conformational continuum studied here (**Figure 54**).

GRGDSP	Metal ions		Soaking time	Resolution	Chains A + B (molecule 1)	Chains C + D (molecule 2)	PDB accession no.
	Mn or Mg	Ca					
<i>mM</i>	<i>mM</i>	<i>mM</i>	<i>h</i>	Å			
0	Mn, 2	0.1	4	2.45	State 1	State 1	3ZDX
10	Mg, 5	1.0	24	2.45	State 2°	State 1°	3ZDY
1	Mn, 2	0.1	4	2.75	State 7°	State 3°	3ZDZ
3	Mn, 2	0.1	4	2.95	State 7	State 4°	3ZEO
5	Mn, 2	0.1	4	3.00	State 7	State 5°	3ZE1
10	Mn, 2	0.1	4	2.35	State 6°	State 8°	3ZE2

Figure 54: Relation of crystal soaking condition to conformational state of $\alpha\text{IIb}\beta\text{3}$ PDB entries (Zhu et al., 2013).

Through an observation of RGD interactions it can be noted that RGD binds at the interface between the αIIb β -propeller and β3 βI domains (**Figure 55**). The Arg and Asp side chains of RGD extend linearly in opposite directions toward βIIb and β3 , respectively (**Figure 55, B–J**). A binding pocket is formed by aliphatic and aromatic side chains, water-mediated interaction with the Arg back-bone carbonyl oxygen, and specific interactions with the Arg and Asp side chains. The Arg's positively charged guanidino moiety forms a salt bridge and, in states 4–8, also forms hydrogen bonds to the side chain of αIIb residue Asp-224 (**Figure 55**). One Asp carboxyl oxygen coordinates to the MIDAS metal ion, and depending on the conformation, the two Asp carboxyl oxygens hydrogen bond to one to three backbone nitrogens of βI domain residues Tyr-122 and Ser-123 in the $\beta\text{1-}\alpha\text{1}$ loop and Arg-214 (**Figure 55, B–J**).

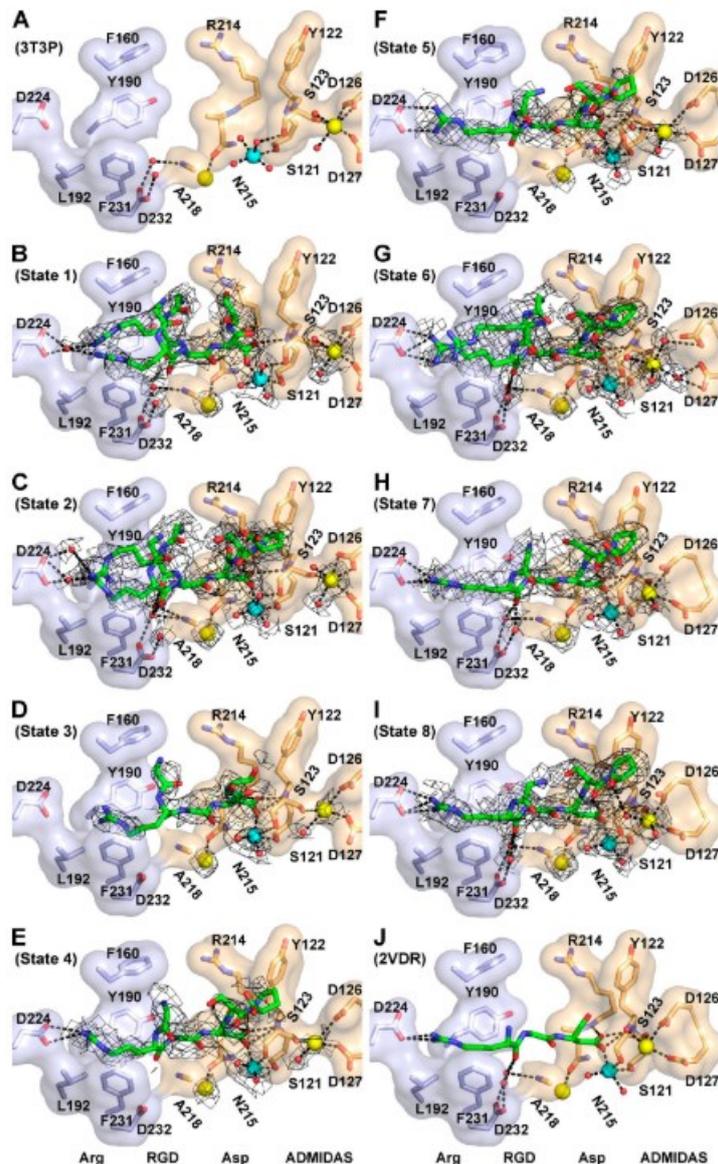


Figure 55: RGD binding pocket; A) PDB entry 3T3P native closed; B-I) stated from 1 to 8 as listed in the table, from closed to intermediate to open; J) PDB entry 2VDR native open (Zhu et al., 2013). Residues that contribute to the RGD-binding pocket are shown both as sticks and transparent surfaces in light blue (allb) and wheat ($\beta 3$). Metal ions are shown as yellow (SyMBS and AdMIDAS) or cyan (MIDAS) spheres. The waters are smaller red spheres. GRGDSP peptides are shown in stick with green carbons. Oxygens and nitrogens are red and blue, respectively. Composite omit simulated-annealing electron density is in black mesh contoured at 3σ for SyMBS and MIDAS metal ions, 1σ (except 0.5σ in E and F) for AdMIDAS metal ion, and 0.5σ for waters and GRGDSP peptide. Hydrogen bonds and metal ion coordination bonds are dashed.

The real space correlation coefficient (RSCC) of electron density for Arg, Gly, and Asp residues in RGD is an estimate of occupancy and order of the ligand at the ligand binding site. The overall trend in increase of RSCC for each of Arg, Gly, and Asp in molecule 2 in states 3, 4, 5, and 8 shows that occupancy by RGD increases over this concentration range. For comparison, RSCC for open $\alpha 1 \beta 3$ headpiece crystals formed with cacodylate ion bound to the $\beta 1$ MIDAS; the cacodylate was replaced by soaking with 0.05 mM RGD peptide for 96 h in Mg/Ca (Fig. 3, state 8). The RSCC values for Asp and Gly of the latter are similar to those for molecule 1 in state 6 and molecule 2 in state 8, which suggests that saturation with RGD is nearly complete after soaking with 10 mM RGD for 4 h in Mn/Ca (Fig. 56). Among the residues of the ligand, the order of RSCC is Arg < Gly < Asp for all molecules. This is consistent with Asp as the primary driver of RGD binding and disorder or multiple conformations of the Arg side chain. Two alternative Arg side chain conformations were evident in molecules 1 and 2

after soaking with 10 mM peptide for 24 h in Mg/Ca. Molecule 1 has greater accessibility of its ligand binding pocket in the crystal lattice than molecule 2, and more completely bound ligand when soaking was limited to 4 h.

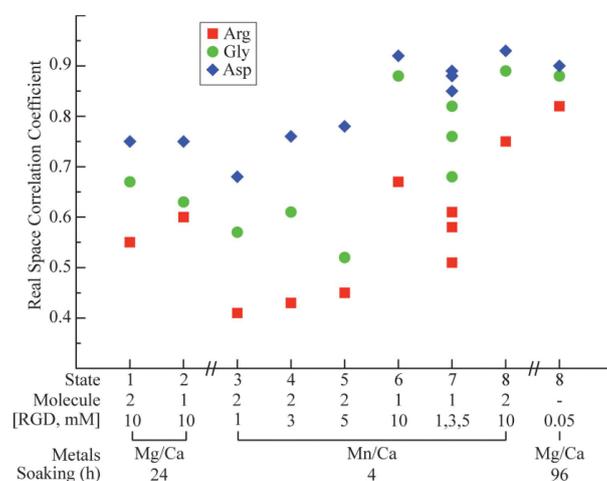


Figure 56: Occupation of the ligand binding site by RGD. As an estimate of binding of each residue of RGD, and their order, it was measured the real space cross-correlation between composite omit simulated-annealing electron density and the molecular model of bound RGD, as explained in the Materials and methods (Zhu et al., 2013).

Strong electron densities were present for the metals at the three β I domain metal ion binding sites. When Mn^{2+} was present, it largely replaced the Mg^{2+} at the MIDAS and the Ca^{2+} at the synergistic metal binding site (SyMBS) and adjacent to MIDAS (AdMIDAS); electron density was fit best when metals at all three sites were modeled as Mn^{2+} (Fig. 55, (Zhu et al., 2013)). In crystals soaked with Mn/Ca alone, the RGD-binding pockets of both molecules 1 and 2 were occupied with solvent molecules, which often occupied the same positions as polar atoms of the ligand.

The structures of Zhu et al. also reveal movements in position of the bound RGD. One of the most important movements during opening is the strengthening of hydrogen bonds between the RGD Asp side chain and the β I β 1- α 1 loop backbone as the distance between these elements decreases (Fig. 55 and Fig. 57). However, both elements also move together toward the α subunit, with the Asp side chain and β 1- α 1 loop moving 1.3 and 2.2 Å, respectively (Fig. 57). Thus, during the opening process, the entire RGD backbone slides in its groove closer to Asp-224 in the β -propeller domain (Fig. 55). The Arg side chain forms strong hydrogen bonds to α IIb Asp-224 only in the final stages of RGD backbone sliding. In states 1–3, the distances are too great for hydrogen bonds, and intervening water molecules are explicitly visible in states 1 and 2 in which RGD density is strong (Fig. 55, B and C). Arg density is weak in states 3–5 (Fig. 55, D–F), and two Arg conformations are present in state 6 (Fig. 55 G). It is only by state 7 that two strong hydrogen bonds develop between the RGD Arg and α IIb Asp-224 side chains (Fig. 55 H).

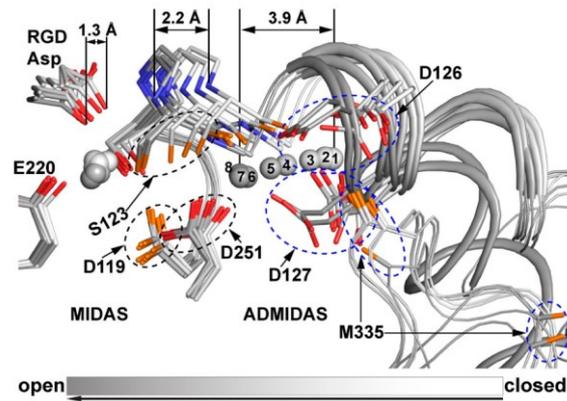
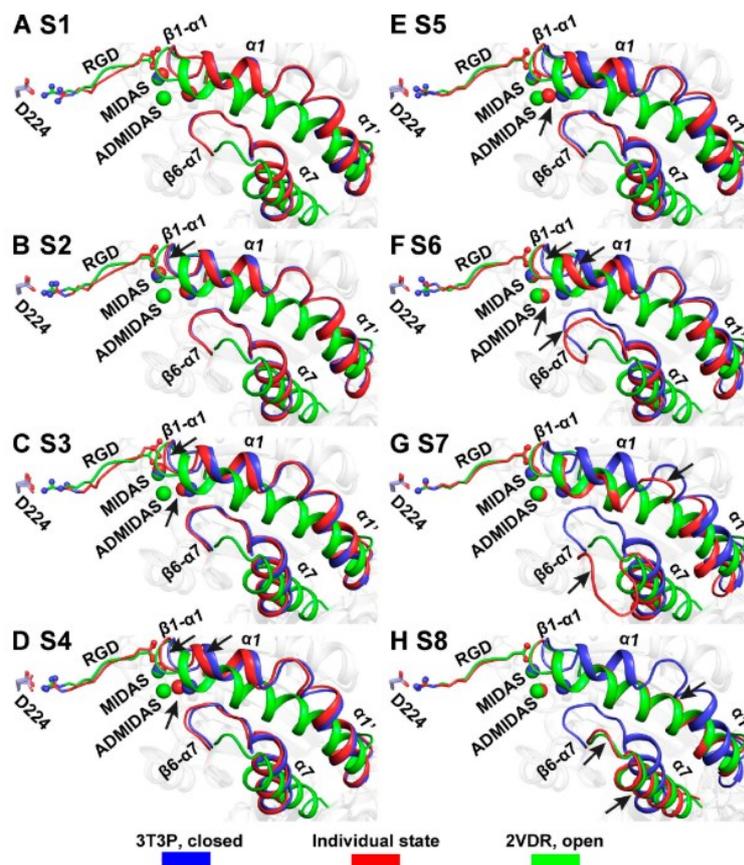


Figure 27: Conformational transition from closed to open around the MIDAS and AdMIDAS. States 1–8 are shown superimposed and shaded on their carbons and metal ions over a grayscale from closed state 1 (white) to open state 8 (dark gray). Key side chain, RGD Asp, and $\beta 1$ - $\alpha 1$ loop backbone atoms are shown in sticks. The remaining backbone is shown as a wormlike trace, with the $\alpha 1$ and merged $\alpha 1/\alpha 1'$ helices thicker. Distances show over- all movements. MIDAS and AdMIDAS metal ions are spheres with states numbered for the AdMIDAS. Some side chains and the Met-335 carbonyl group are circled, and their oxygens are shown in orange or red to tell them apart. Nitrogens are shown in blue (Zhu et al., 2013).



Downloaded from jcb.rupress.org on June 6, 2015

Figure 58: Overview of the moving portions of the $\beta 1$ domain between three conformational states. A-H) States 1 to 8; $\beta 1$ domain regions that undergo the largest movements are shown in cartoon. Asp-224 of the $\alpha 1b$ subunit and RGD are shown in stick. MIDAS and AdMIDAS metal ions are shown as spheres (Zhu et al., 2013).

Although the term hinge opening was used in molecular dynamics (MD) simulations of $\alpha v \beta 3$ (Puklin-Faucher et al., 2006; Puklin-Faucher & Vogel, 2009) the simulations did not result in the open $\beta 1$ domain or open headpiece conformations seen here or previously (Xiao et al., 2004). Compared with

a 62° increase in angle between β I and hybrid domains in the open headpiece crystal structure, MD showed a 23° increase starting with a model of 10FnIII bound to a RGD-liganded state 5–like structure and a 13° increase starting with an unliganded state 1 structure lacking metal ions at the MIDAS and SyMBS. The α 1 and α 1' helices merged and approached the α 7 helix; however, the merger occurred in the absence of α 7 helix pistoning and displacement of β 6- α 7 loop residue Val-340 from its ratchet pocket and occupation of this pocket by Leu-134 in the α 1/ α 1' helix. Because the simulations began with state 5, they were silent on RGD sliding and β 1- α 1 loop movement, which are largely complete by states 5–6. Dependence on RGD was not demonstrated by omitting RGD from the state 5–like structure. RGD was not sufficient for α 1 and α 1' helix merger because it required a modeled interaction between the body of Fn3 domain 10 and the Trp-129 side chain in the α 1 helix itself. Large changes in position and rotamer of Trp-129 between our states 4, 6, 7, and 8 (**Fig. 59 B**) are incompatible with maintenance of a side chain hydrogen bond to the 10FnIII in the MD simulations.

MD simulations of integrins are challenging because the force fields lack descriptions of the highly directional nature of octahedral metal coordination and hydrogen bonds. MD lost physiological coordinations at the MIDAS and SyMBS (Craig et al., 2004) and an invariant water at the MIDAS (Puklin-Faucher et al., 2006). Additionally, the α V β 3 structures used in simulations lacked cis-Pro at β I domain residues 163 and 169 and contained a sequence-to-structure frameshift at specificity-determining loop residues 168–176 (Dong et al., 2012) that lie near the FnIII docking site and α 1 helix and could have affected MD results.

Surprisingly, β I domain reshaping is related to a change in position of RGD in the ligand-binding pocket. Movement of the β 1- α 1 loop toward the Asp of RGD permitted the entire RGD moiety to slide in its groove away from the β 3 subunit and toward the α IIb subunit. Sliding enabled a water-mediated interaction between the RGD Arg side chain and α IIb Asp-224 to be converted to a much stronger direct, charged interaction through multiple hydrogen bonds. Sliding also correlated with a shift from multiple Arg side chain conformations to a single conformation. Sliding and disorder or multiple conformations of the Arg side chain show that binding of the Asp of RGD to the MIDAS metal ion and the β 1- α 1 loop backbone is energetically more important than Arg interactions with the α IIb subunit.

State 1 structure shows that RGD can bind to α IIb β 3 with no appreciable change in structure. The early parts of the pathway from state 1 to 6 are captured in atomic detail as movements of <1 Å, whereas those in states 7 and 8 involve large concerted changes. Ligand-induced integrin shape shifting begins in the β 1- α 1 loop and then works its way more C-terminally in the α 1 and α 1' helices until they finally merge and push the β 6- α 7 loop and α 7 helix out of the way toward the hybrid domain.

Linkage between the MIDAS and ADMIDAS coordination shells grows between states 3 and 6. In state 3, the Ser-123 side chain replaces a water molecule in the inner MIDAS coordination sphere; the Ser-123 backbone oxygen directly coordinates the ADMIDAS through states 1–8. In state 5, the ADMIDAS metal ion directly coordinates the Asp-251 side chain. In state 6, the same Asp-251 oxygen indirectly coordinates the MIDAS through the Ser-123 side chain; the other Asp-251 oxygen indirectly coordinates the MIDAS through a water molecule throughout states 1–8 (**Fig. 59**). By state 6, interaction between the MIDAS and ADMIDAS coordination shells is strong enough to leave behind the direct ADMIDAS coordinations to the Asp-126 and Asp-127 side chains, which are reformed in state 7. Saltatory motions between states 1 and 7 thus include numerous steps of removal, addition, and removal yet again of waters that provide indirect MIDAS and ADMIDAS coordinations to side chains. Another movement not predicted by structural interpolation is that of the β 6- α 7 loop. Its coordinating Met-335 carbonyl oxygen moves toward the fleeing ADMIDAS in state 3 and stays in this new position until state 6, when it moves in the opposite direction.

Presumably, the movement of the $\beta 1$ - $\alpha 1$ loop places strain on the $\alpha 1$ helix, which contains ADMIDAS-coordinating residues Asp-126 and Asp-127. In state 7, there is a large rigid-body movement of the $\alpha 1$ helix, as Asp-126 and Asp-127 catch up with and reform direct coordinations to the ADMIDAS. In turn, the movement of the $\alpha 1$ helix is likely to place strain on the $\alpha 1'$ helix. In final state 8, the $\alpha 1'$ helix catches up with and joins to the $\alpha 1$ helix. Alignment of the $\alpha 1$ and $\alpha 1'$ helices squeezes at their junction the $\beta 6$ - $\alpha 7$ loop. The hydrophobic ratchet pocket occupied by $\beta 6$ - $\alpha 7$ residue Val-340 in states 1–6 is occupied by $\alpha 1$ helix residue Leu-134 in state 8, with Val-340 in an intermediate position in state 7 (**Fig. 59**). Merging of the $\alpha 1$ and $\alpha 1'$ helices appears to be the final straw that breaks the camel's back and pushes the 7 helix toward the hybrid domain, causing it to swing out. The side chain of Trp-129 near the end of the $\alpha 1$ helix appears to buffer the large $\alpha 1$ and $\alpha 1'$ conformational movements between states 6 and 8. Between states 1 and 6, the Trp-129 side chain can occupy either of the two rotamers shown in **Fig. 59 B** (image 1). In state 7, it adopts a very different buried rotamer and acts as a placeholder for the $\alpha 1'$ helix. Then, in state 8, Trp-129 moves outward again and adopts yet another rotamer to make way for the merge of the $\alpha 1$ and $\alpha 1'$ helices and the side chain of Leu-134.

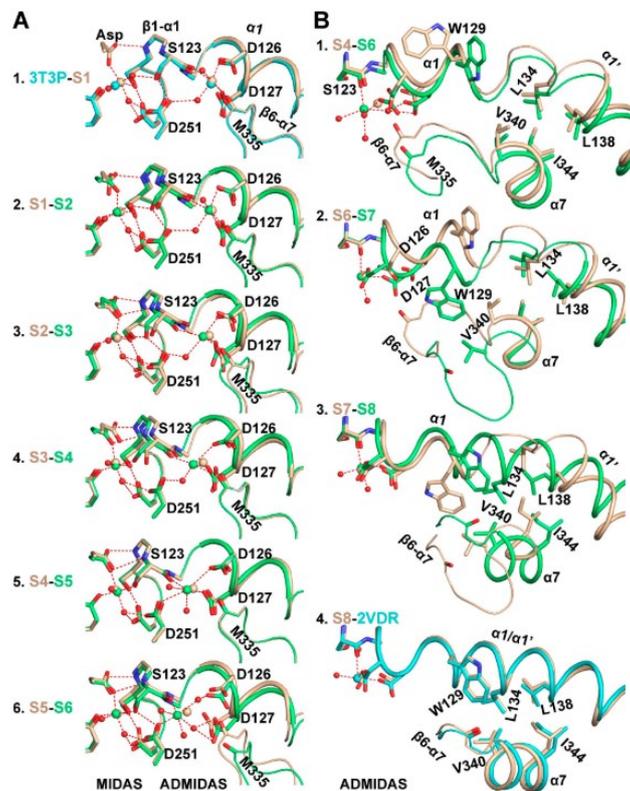
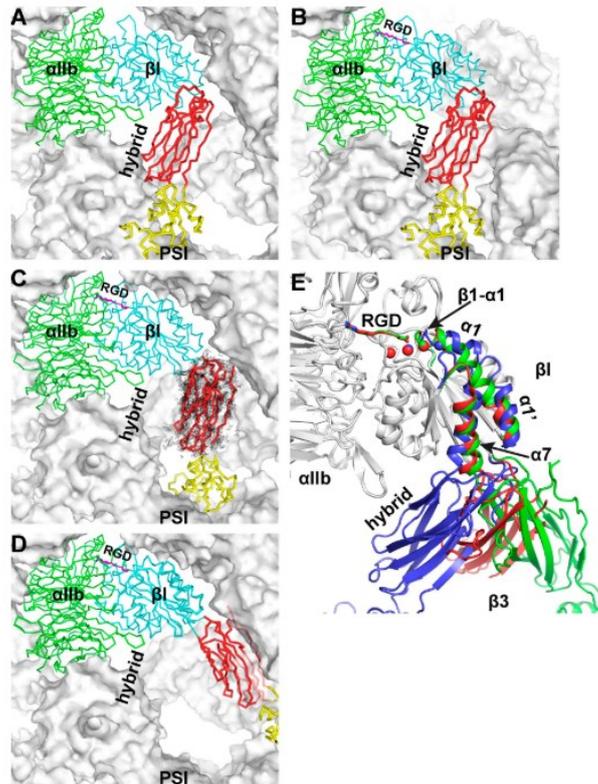


Figure 59: Detailed comparisons between nearest-neighbor states. (A) The region around the ligand Asp, $\alpha 1$ - $\beta 1$ loop, $\alpha 1$ -helix, and $\beta 6$ - $\alpha 7$ loop where movement is greatest between states 1 and 6 (S1–S6). (B) The region around the $\alpha 1$ helix, $\alpha 1'$ helix, $\beta 6$ - $\alpha 7$ loop, and $\alpha 7$ helix where movement is greatest between states 6 and 8. Each panel compares two nearest-neighbor states. For economy, and to compare the two rotamers of W129 in states 1–5, image 1 in B compares states 4 and 6. The carbons and metal ions of each state are in the same colors as the names of each state or the reference structures 3T3P (closed) and 2VDR (open). For clarity, water molecules as spheres and metal coordination bonds as red dashes are shown only for the second named structure in each image. Nitrogens and oxygens are shown in blue and red, respectively.



Downloaded from jcb.rupress.org on June 6, 2015

Figure 60: Hybrid domain swing out. (A–D) One integrin molecule is shown as a Ca trace, with different colors for each domain. The hybrid domain (red) and PSI and I-EGF-1 domains (yellow) are shown as thicker traces for emphasis. Other integrin molecules and all Fabs in the crystal lattice are shown as white, semitransparent, solvent accessible surfaces. The label hybrid is placed in identical positions in A–D. (A) Molecule 2 before soaking (3T3P closed structure). (B) Molecule 1 after soaking with 10 mM RGD and Mn/Ca. (C) Molecule 2 after soaking with 10 mM RGD and Mn/Ca. Composite omit simulated-annealing electron density contoured at 0.5σ around the hybrid domain is shown as purple mesh. PSI and I-EGF-1 domains are missing in density, and superposition on the hybrid domain is used to show their approximate location in the lattice. (D) The native open headpiece (Protein Data Bank [PDB] accession no. 2VDR) superimposed based on the β -propeller and β I domains in C and shown in the same lattice as in C. Severe clashes are evident. (E) Superposition of α IIb β 3 headpieces. Similar regions in gray and colored shape-shifting portions in cartoon; metal ions are shown as spheres, and RGD is shown in stick. Structures are molecule 2 after soaking with 10 mM RGD and Mn/Ca (red), native closed (PDB accession no. 3T3P; blue), and native open (PDB accession no. 2VDR; green).

These movements illustrate the separability yet interdependence of the MIDAS and ADMIDAS coordination shells, the β 1- α 1 loop, α 1 helix, α 1' helix, and β 6- α 7 loop. The final movements of the α 1' helix and α 7 helix are highly concerted. Compared with the earlier movements in states 1–6, those of the α 1 helix and β 6- α 7 loop in state 7 are much larger, whereas the movements in the α 1' helix, β 6- α 7 loop, and α 7 helix in state 8 are cataclysmic. The extraordinary movements that follow state 7 likely correspond to a large decrease in free energy and suggest that state 7 must be close to the transition state for conformational change.

A limitation of such study could be that the states seen here are trapped in crystals and may be perturbed by lattice contacts, but changes were observed in two molecules in different lattice environments. Furthermore, despite the presence of states 2, 6, and 7 in molecule 1 and states 1, 3, 4, 5, and 8 in molecule 2, these states fall on a single pathway of headpiece opening. Constraints limiting conformational change differ in crystals and on cell surfaces. Nonetheless, hybrid domain swing out is the largest motion in headpiece opening and thus may be rate limiting for conformational change on cell surfaces as well as in crystals. Furthermore, the different moving elements, i.e., α 1 helix, α 1' helix, β 6- α 7 loop, α 7 helix, and hybrid domain are key on cell surfaces as well, as shown by mutational studies in these elements (Barton et al., 2004; Kamata et al., 2010; B. H. Luo et al., 2003,

2009; B. H. Luo, Springer, et al., 2004; B. H. Luo, Takagi, et al., 2004; Mould, Barton, Askari, McEwan, et al., 2003; Yang et al., 2004)

The principle of the reversibility of chemical reactions implies that inside-out activation of integrins may proceed by a similar pathway as outside-in activation studied here but in the opposite direction.

Two alternative mechanisms for achieving conformational change in proteins, selection by ligand of preexisting conformational states and ligand-induced fit (Henzler-Wildman & Kern, 2007), are linked in a thermodynamic cycle (**Fig. 61**). In the structure of Zhu et al. (Zhu et al., 2013) studied here, no conformational change is required for RGD binding to state 1, and thus, induced fit (**Fig. 61, k1 and k2**) is the mechanism driving conformational change. However, on cell surfaces, interconversion between conformational states driven by thermal motion (**Fig. 61, k3 and k-3**) will be far faster than in crystals, and inside-out signals may also increase k3. Therefore, the relative fluxes of integrins to ligand-bound states on cell surfaces through induced fit (k1 and k2), and selection of preexisting states (k3 and k4) will depend on their relative rates, as well as ligand concentration, and remains to be determined.

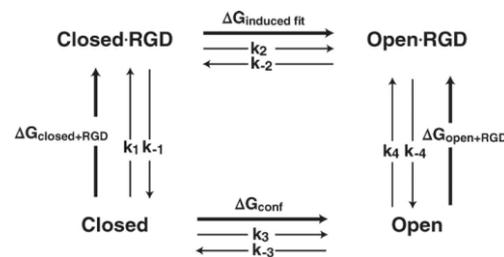


Figure 61: The thermodynamic cycle for ligand binding and conformational change in integrins. The ΔG values are for reactions in the direction shown by the arrows. From the unliganded closed state in the bottom left, the induced-fit mechanism proceeds clockwise, and the preexisting conformational change mechanism with selection by ligand proceeds counterclockwise, to the liganded open state in the top right. The cycle for headpiece fragments can be modified for intact integrins on cell surfaces by adding additional conformational states.

RGD must bind with higher affinity to the open than closed headpiece to drive conversion to the higher energy open headpiece conformation. This may be formally demonstrated using the thermodynamic cycle (**Fig. 61**). The difference in energy between any two states in this cycle is identical, whether conversion occurs by clockwise or counterclockwise routes. Therefore, $\Delta G_{\text{open+RGD}} = \Delta G_{\text{closed+RGD}} + \Delta G_{\text{induced fit}} - \Delta G_{\text{conf}}$.

Crystal soaking experiments demonstrate that $\Delta G_{\text{induced fit}}$ is highly negative. A large number of experiments referenced in the Introduction demonstrate that in absence of ligand, $[\text{Closed}] \gg [\text{Open}]$; therefore, ΔG_{conf} is highly positive. It follows from the aforementioned equation that $\Delta G_{\text{open+RGD}} \ll \Delta G_{\text{closed+RGD}}$; i.e., that the affinity of the open headpiece is much higher than that of the closed headpiece for RGD.

Conclusions based on soaking experiments in Mn/Ca and the thermodynamic cycle are supported by soaking results in Mg/Ca (**Fig. 56**): Based on soaking with 10 mM RGD in Mn/Ca for different times, near-equilibrium concentrations of RGD should have been reached inside the crystal lattice after soaking with 10 mM RGD in Mg/Ca for 24 h. Incomplete occupation of the ligand binding site in Mg/Ca as shown by the RSCC values in **Fig. 56** suggests an affinity insufficient to saturate binding in 10 mM RGD, consistent with lack of detectable binding after soaking for 72 h with 0.34 mM RGD in Mg/Ca. In contrast, when crystals with the open headpiece were soaked for 96 h with 0.05 mM RGD in Mg/Ca, RGD peptide completely replaced a cacodylate pseudo-ligand (**Fig. 3**; Springer et al., 2008). Complete

saturation of the open headpiece at 0.05 mM RGD and incomplete saturation of the closed headpiece at 10 mM RGD support the conclusion that the affinity of the open α IIb β 3 headpiece for RGD peptide is ≥ 200 -fold higher than that of the closed headpiece in Mg/Ca.

The ability of the lattice around molecule 2 to hold the integrin headpiece in the closed conformation after binding RGD in Mg/Ca gives the possibility to estimate the affinity of the closed headpiece for RGD. This affinity is between 0.34 and 10 mM and closer to 10 mM. It is not possible to estimate the affinity of the closed headpiece for RGD from measurements of binding to integrins in solution or on cell surfaces because, as already discussed, the higher affinity for the open headpiece is more than sufficient to pay the energetic penalty of inducing or selecting the open headpiece conformation. Conversion to the open headpiece is consistent with 50% effective concentration estimates in the range of 7 μ M to 1 mM RGD in Mg/Ca for binding of antibodies to ligand-induced binding sites or increased protease sensitivity using isolated, intact α IIb β 3 or α IIb β 3 and α V β 3 on cell surfaces (Frelinger et al., 1988, 1990; Parise et al., 1987). 50% inhibitory concentration values for inhibition by RGD of activated platelet binding to ligands are in the range of 0.01–0.2 mM (Plow et al., 1985), suggesting that the headpiece has been opened.

The thermodynamic cycle characterized here for headpiece fragments has wide applicability. In extended integrins, as in headpiece fragments, none of the headpiece domains are in buried interfaces, and the lower β leg is highly flexible (Fig. 34, dashed lines). Therefore, the relative energies of the four states in the cycle (Fig. 61) are expected to be very similar in headpiece fragments and in extended integrins on cell surfaces. Bent integrins on cell surfaces have extensive interfaces that are exposed upon extension; the hybrid domain is in one of these buried interfaces (Fig. 34). Also, the α and β subunit C termini are close to one another in the bent conformation, and the TM domains associate (Luo et al., 2007; Springer and Dustin, 2012). Because headpiece opening exposes most of the same interfaces that become exposed upon integrin extension, $\Delta G_{\text{induced fit}}$ and ΔG_{conf} in Fig. 8 (which each require headpiece opening) are higher (by similar amounts) for cell surface integrins than for headpiece fragments. If one is willing to accept that ligand-induced binding site epitopes in β 3 integrins measure either integrin extension induced by headpiece opening or headpiece opening itself, one could conclude that $\Delta G_{\text{induced fit}}$ is negative and ΔG_{conf} is positive (as for integrin headpieces) and deduce that the open headpiece conformation has higher affinity for RGD than the closed headpiece conformation in intact integrins, just as demonstrated here for headpiece fragments.

It remains to be elucidated if headpiece opening upon ligand binding is general for integrins or is dependent on the integrin or the ligand. MAdCAM-1 binding to the α 4 β 7 headpiece can yield either an intermediate or open state as seen by EM (Yu et al., 2012), which appear to mediate rolling adhesion (intermediate affinity) and firm adhesion (high affinity), respectively (Yang et al., 2004)(Chen et al., 2004). Thus, there may be differences among integrins. However, headpiece opening has been demonstrated for basically integrins that had been tested, including those containing the β 1, β 2, β 3, β 6, and β 7 subunits (M. Shi et al., 2011; Springer & Dustin, 2012; Yu et al., 2012). An obvious exception would be when a carboxyl group is absent in the ligand because this is the moiety that interacts with the integrin β subunit in which allostery occurs. Thus, antagonists that mimic the Arg moiety of RGD and do not bind the MIDAS, or instead displace Mg^{2+} from the MIDAS, do not activate opening and stabilize the closed headpiece against opening, respectively (Zhu et al., 2010, 2012).

The primary role of the Asp shown here in allostery suggests that Arg of RGD might not be required to pull the β I domain β 1- α 1 loop toward the α subunit to open the headpiece. Some small molecules based on RGD, selected by the pharmaceutical industry for their ability to bind equally well to the low and high affinity states of α IIb β 3, are reported not to induce reactivity with antibodies to ligand-induced binding sites (Aga et al., 2004) and, therefore, appear not to induce headpiece opening. These

mimetics have unique chemical features that may set them apart from the Asp and Glu side chains present in physiological integrin ligands.

The large movement at the integrin knees upon headpiece opening is thought to be important for transmission of allostery through long legs that are flexible except when elongational force is applied in cell adhesion (Springer & Dustin, 2012; Zhu et al., 2008). Similar to what is already seen in other integrin types activation, a movement of only 2 Å at the $\beta 1$ - $\alpha 1$ loop in the ligand binding site is transmitted through an intricate shape-shifting pathway a distance of 40 Å across the βI domain. A 10-Å $\alpha 7$ helix movement like that of a connecting rod in the βI domain causes the hybrid domain to swing out by pivoting at its other connection to the βI domain. The length of the hybrid domain is 40 Å, and the PSI and I-EGF1 domains attached at the end opposite the βI domain make the total length of the upper integrin β leg circa 70 Å. The leverlike swing of the upper leg amplifies the 2-Å movement of the $\beta 1$ - $\alpha 1$ loop to a 75-Å increase in separation at the integrin knees (**Fig. 34**). The integrin headpiece appears to be designed for the purpose of transmitting allostery in extracellular environments. If in general the open headpiece is the high affinity state of integrins, induction by inside-out signals of headpiece opening can be a general mechanism for integrin activation on cell surfaces (Zhu et al., 2013)

2. Methods

2.1. Crystallography structures

The existing PDB entries of the structures, modules and integrins of our interest will be analyzed on their respective paper and on MOE, using the FASTA of fibronectin and various integrins as standard primary sequence. The sequence with best resolution of the residues involved in FN-IN complexes will be privileged and mutations will be avoided. In our previous work the crystallographic structure found in PDB entries 3T1W and 4LXO were used in the construction of the homology model of FN in these residues. Another research was started to find other structures that could better visualize the structure and conformation of FN and IN and their complexes. 4LXO represents 9FnIII, 10FnIII-elegantin chimera FN, thus a mutated form created to assess the effects of such mutations on FN behavior. Besides the respective paper is not available. Observing the entry on MOE mutations from the original FASTA are found in various residues, including the RGDS sequence. For all these reasons this entry will be discarded, and 3T1W will be used to assess the possibility of steric or allosteric inhibition of some bound integrin states.

Various crystallographic structures of INs are available, some of them portraying the interactions with FN and RGD ligands. The majority is also represented by INs fixated by monoclonal antibodies, in order to achieve a specific conformational state. Some of these mAb are known to be candidate as potential medical treatment due to their ability to induce conformational changes in IN and acting as steric or allosteric inhibitor. As already repeated in this paper, 8FnIII, 9FnIII, 10FnIII modules are the ones involved interactions with FN.

In this work special attention is given to integrin type $\alpha V\beta 3$, $\alpha IIb\beta 3$, $\alpha 5\beta 1$, given their importance in physiological and pathological interactions involving angiogenesis, thrombosis, wound healing and cancer progression linked to such effects.

In general, all the entries already used in the study on FN homology model were analyzed anew. FASTA sequence for integrin αIIb , αV , $\alpha 5$, $\beta 1$ and $\beta 3$ were obtained from UniProt, and a search on UniProt brought to the attention the PDB entries listed in the table alike. The following table reports all the entries, their content, modules and/or integrins represented, conformational state in the known

degree (open/extended, intermediate/half-bent, closed/bent) and resolution reported. All entries represent structures for the Homo Sapiens species.

PDB	Paper content	Chains	State (head/leg)	Resolution Å
1L5G	Crystal structure of the extracellular segment of integrin $\alpha V\beta 3$ in complex with an arg-gly-asp ligand (J. P. Xiong et al., 2002)	$\alpha V, \beta 3$	Closed	3.20
4MMX	Integrin $\alpha V\beta 3$ ectodomain bound to 10FnIII of FN (Van Agthoven et al., 2014b)	$\alpha V, \beta 3, 10FnIII$	Closed	3.32
6NAJ	Integrin $\alpha V\beta 3$ ectodomain bound to Hr10 variant of 10FnIII (Adair et al., 2020)	$\alpha V, \beta 3, Hr10FnIII$	Closed	3.10
6AVQ	The Therapeutic Antibody LM609 Selectively Inhibits Ligand Binding to Human alpha-V beta-3 Integrin via Steric Hindrance (Borst et al., 2017)	$\alpha V, \beta 3$	Closed	35
6AVR	Human alpha-V beta-3 Integrin (intermediate conformation) in complex with the therapeutic antibody LM609 (Borst et al., 2017)	$\alpha V, \beta 3$	Intermediate	35
6AVU	Human alpha-V beta-3 Integrin (open conformation) in complex with the therapeutic antibody LM609 (Borst et al., 2017)	$\alpha V, \beta 3$	Open	35
4G1M	Re-refinement of alpha V beta 3 structure (Dong et al., 2012)	$\alpha V, \beta 3$	Closed	2.90
1M1X	Crystal structure of the extracellular segment of integrin alpha vbeta3 bound to Mn2+ (J. P. Xiong et al., 2002)	$\alpha V, \beta 3$	Closed	3.30
3VI3	Crystal structure of alpha5beta1 integrin headpiece (ligand-free form) (Nagae et al., 2012)	$\alpha 5, \beta 1$	Closed	2.90
3VI4	Crystal structure of alpha5beta1 integrin headpiece in complex with RGD peptide (Nagae et al., 2012)	$\alpha 5, \beta 1$	Closed	2.90
4WJK	Metal Ion and Ligand Binding of Integrin $\alpha 5\beta 1$ (Xia & Springer, 2014)	$\alpha 5, \beta 1$	Closed	1.85
4WK0	Metal Ion and Ligand Binding of Integrin $\alpha 5\beta 1$ (Xia & Springer, 2014)	$\alpha 5, \beta 1$	Closed	1.76
4WK2	Metal Ion and Ligand Binding of Integrin $\alpha 5\beta 1$ (Xia & Springer, 2014)	$\alpha 5, \beta 1$	Closed	2.50
4WK4	Metal Ion and Ligand Binding of Integrin $\alpha 5\beta 1$ (Xia & Springer, 2014)	$\alpha 5, \beta 1$	Closed	2.50
7NWL	Cryo-EM structure of human integrin $\alpha 5\beta 1$ (open form) in complex with FN and TS2/16 Fv-clasp (Schumacher et al., 2021)	$\alpha 5, \beta 1, 10FnIII, TS2/16$	Open	3.10 Å
7NXD	Cryo-EM structure of human integrin alpha5beta1 in the half-bent conformation (Schumacher et al., 2021)	$\alpha 5, \beta 1$	Intermediate	4.60

7U60	Integrin alpha IIB beta 3 headpiece with Fab complex with Eptifibatide (Zhu, Lin to be published)	α IIB, β 3	Closed	2.55
2VDL	Re-refinement of Integrin AlphaIIBBeta3 Headpiece (Springer et al., 2008)	α IIB, β 3	Open	2.75
2VDP	Integrin AlphaIIBBeta3 Headpiece Bound to Fibrinogen Gamma chain peptide, LGGAKQAGDV (Springer et al., 2008)	α IIB, β 3	Open	2.80
2VDO	Integrin AlphaIIBBeta3 Headpiece Bound to Fibrinogen Gamma chain peptide, HHLGGAKQAGDV (Springer et al., 2008)	α IIB, β 3	Open	2.51
2VDM	Re-refinement of Integrin AlphaIIBBeta3 Headpiece Bound to Antagonist Tirofiban (Springer et al., 2008)	α IIB, β 3	Open	2.90
2VDN	Re-refinement of Integrin AlphaIIBBeta3 Headpiece Bound to Antagonist Eptifibatide (Springer et al., 2008)	α IIB, β 3	Open	2.90
2VDQ	Integrin AlphaIIBBeta3 Headpiece Bound to a Chimeric Fibrinogen Gamma chain peptide, HHLGGAKQRGDV (Springer et al., 2008)	α IIB, β 3	Open	2.59
2VDK	Re-refinement of Integrin AlphaIIBBeta3 Headpiece (Springer et al., 2008)	α IIB, β 3	Open	2.60
2VDR	Integrin AlphaIIBBeta3 Headpiece Bound to a chimeric Fibrinogen Gamma chain peptide, LGGAKQRGDV (Springer et al., 2008)	α IIB, β 3	Open	2.40
2VC2	Re-refinement of Integrin AlphaIIBBeta3 Headpiece Bound to Antagonist L-739758 (Springer et al., 2008)	α IIB, β 3	open	3.10
4Z7N	Integrin alphaIIBbeta3 in complex with AGDV peptide (Lin et al., 2016)	α IIB, β 3	Closed	2.60
3FCS	Structure of complete ectodomain of integrin alphaIIBbeta3 (Zhu et al., 2008)	α IIB, β 3	Closed (much more than entries of same pape)	2.55
3FCU	Structure of headpiece of integrin alphaIIBbeta3 in open conformation (Zhu et al., 2008)	α IIB, β 3	Open (not completely)	2.90
3T3P	A Novel High Affinity Integrin alphaIIBbeta3 Receptor Antagonist That Unexpectedly Displaces Mg ²⁺ from the beta3 MIDAS (Zhu et al., 2012)	α IIB, β 3	Native Closed	2.20
3NID	The Closed Headpiece of Integrin alphaIIB beta3 and its Complex with an alphaIIB beta3 -Specific Antagonist That Does Not Induce Opening (Zhu et al., 2010)	α IIB, β 3	Closed	2.30
3ZDX	Integrin alphaIIB beta3 headpiece and RGD peptide complex (Zhu et al., 2013).	α IIB, β 3	Closed	2.45

3ZDY	Integrin alphaIIb beta3 headpiece and RGD peptide complex (Zhu et al., 2013)	α IIb, β 3	Closed	2.45
3ZDZ	Integrin alphaIIb beta3 headpiece and RGD peptide complex	α IIb, β 3	Closed	2.45
3ZE0	Integrin alphaIIb beta3 headpiece and RGD peptide complex	α IIb, β 3	Closed	2.95
3ZE1	Integrin alphaIIb beta3 headpiece and RGD peptide complex	α IIb, β 3	Closed	3.00
3ZE2	Integrin alphaIIb beta3 headpiece and RGD peptide complex (Zhu et al., 2013).	α IIb, β 3	Closed	2.35
3T1W	Structure of the four domain fragment 7FnIII, EDB, 8FnIII, 9FnIII (Schiefner et al., 2012)	7-9FnIII		2.40

Every single structure was downloaded, and structure preparation was performed on MOE. The standard procedure consisted in correction of structural issues with the “Structure preparation tool”, a QuickPrep and an operation of Protonate3D with salt concentration of 0.15 and dielectric constant of 1.

PDB entries referring to α IIb β 3 3ZDX, 3ZDY, 3ZDZ, 3ZE0, 3ZE1, 3ZE2, represent the integrin in two conformational states each. In work it will be used the state 1 (3ZDX), state 2 (3ZDY), state 3 (3ZDZ), state 4 (3ZE0), state 5 (3ZE1) and state 8 (3ZE2), for each entry respectively (Zhu et al., 2013).

The structures of each integrin type were aligned and superposed, in order to highlight conformational changes between different integrin states. Each set of a integrin typeThe superposition are saved under *alpha5beta1_structsuper.moe*, *alphaVbeta3_structsuper.moe*, *alphaIIbbeta3_structsuper.moe*.

Exploiting the presence of crystallography structures containing RGD ligands or FN fragments, the ligands-IN, FN-IN contacts were analyzed and compared for different integrin type and ligand molecule. Then, a first docking analysis to refine the pose of RGD ligands and calculate the S score of the binding in the crystallography structures was performed, to form a standard to which compare the new results: RGD ligands from known structure of the analyzed integrin subunits were taken as standard which will be used in the refinement docking: 1L5C.C was used as standard for the α V β 3 integrins, 7U60.M for α IIb β 3, 3VI4.I, 4WK0.C, 4WK2.C and 4WK4.C for α 5 β 1.

The docking procedure was performed using induced fit as refinement method, returning a total of 10 refined poses from 50. The first poses of each structure with relative scores are listed in the table below. The MDB database files were saved as “docking_RGDstructTagchainChain_ReceptorTaginchainChain.mdb”. The evaluate the S score of crystallography structures that lacked the RGD ligand binding site, the RGD binding site is simulated by RGD ligand structures used as standard in the previous operation (e.g. 6NAJ is aligned and superposed to 1L5G, and docking simulation is performed by the RGD ligand 1L5G.C with receptor 6NAJ chains and as ligand site the own RGD ligand 1L5G.C). This simulation could also help to understand if the computational simulation returns affinity results that confirm the consideration that the bent state of IN presents lower affinity to RGD interactions, and that some monoclonal antibodies could act as inhibitor of such interactions. For structures with mAb present in crystallography, simulation were done in presence and absence of the mAb, to evaluate if their presence could affect in same way the S-score.

The RGD ligand of different PDB entries were docked between each other to see the affinity of this ligand for different structures and subunits.

To confirm the assumption that bigger ligands have higher affinity it was performed the docking between the ligands of entries 4WK2 and 4WK4 to see an eventual increase in S score.

Following, cilengitide (EMD-121974), a known RGD-based peptid with promising therapeutic use, was tested as ligand in the RGD ligand binding site through docking simulation. Docking setting used remain unchanged. The site chosen for docking was the interface that interacts with the RGD ligand, so the chains representing these ligands were chosen, as done previously in the refinement. PDB entry 1L5C.C was used as ligand site for integrins $\alpha V\beta 3$, 7U60.M for $\alpha IIb\beta 3$, 3VI4.I and 4WK0.C for $\alpha 5\beta 1$ integrins. The results are saved as "docking_cilengitide_ReceptorTaginstructSiteTagchainsitechain.mdb"

To analyze the interactions involved in cilengitide-IN interactions, the contacts between the Cilengitide ligand and the integrins subunits were calculated, to observe the mechanism of binding and its possible similarity to interactions with known RGD-ligands, 9FnIII and 10FnIII.

Cilengitide structure was modified trying to find other conformation or structure that could possibly be a better inhibitor for the integrin $\alpha V\beta 3$ receptor. These changes should enhance the interaction between the ligands and the interaction regions in both integrin subunits. The ideal would be to enhance the polar and hydrophobic properties of the chain (Asp with βA and Gly with αV , respectively) that interacts with the two-integrin subunit, making affinity for RGD-mimetics stronger than the pathological interaction. A total of 11 new compounds were built, and their S score analyzed through docking simulation using the already stated settings.

The compounds used are totally computationally built, so it's not even known if they are synthesizable or safe for health. First a search of possible pre-existing patents was performed on PubChem, Wipo and Google Patents. The compounds ADMET was analyzed thanks to Domiziano Doria, of University of Alberta, through ADMET software.

3. Results

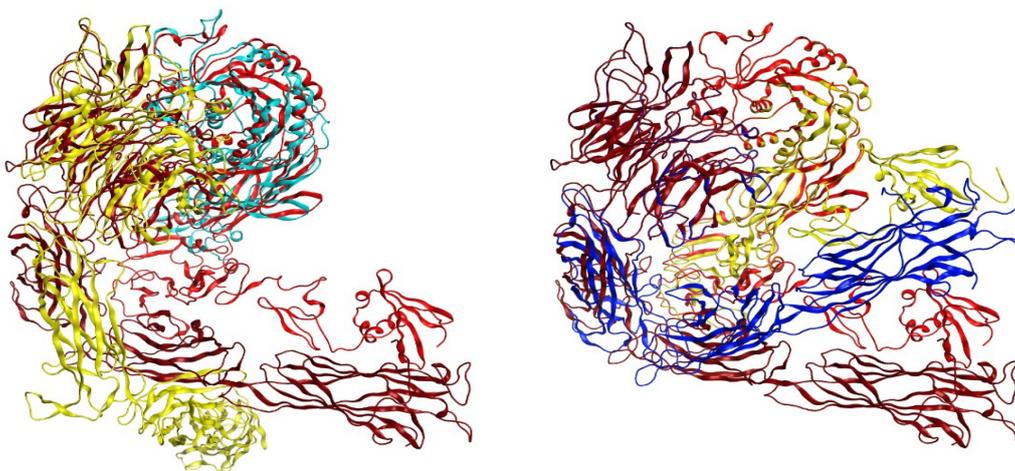


Figure 62: Observing the superposition of PDB entries representing $\alpha V\beta 3$ a sequence of the conformational state of integrin can be extracted. In particular entries 6AVR, 6AVQ and 6AVU represent integrin $\alpha V\beta 3$ in complex with mAb LM609 in the three conformational state known for integrin (open/extended, intermediate, closed/bent). a) Open extended conformation compared to intermediate conformation; b) intermediate conformation compared to close conformation.

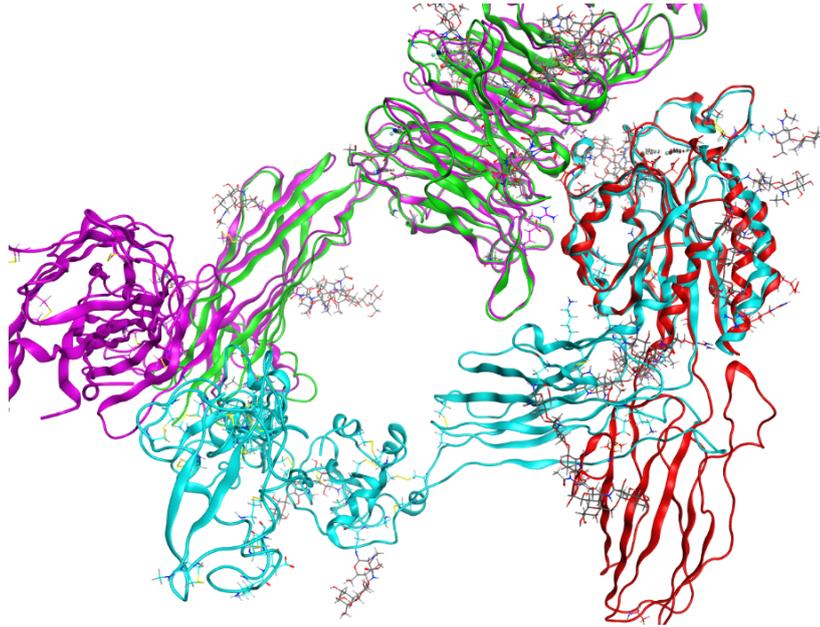


Figure 63: superposition of PDB entries representing $\alpha 5 \beta 1$. In particular entries represent integrin $\alpha 5 \beta 1$ in two conformational state (PDB entries: 7NWL (open/extended), 7NXD (half-bent)).

From observations on the superposition of the different PDB entries of the three integrin types, it is easily noticed the transition between the three conformational state: the swing of the β -I and hybrid domain to open the headpiece can be easily observed (**figure 62-63-64**).

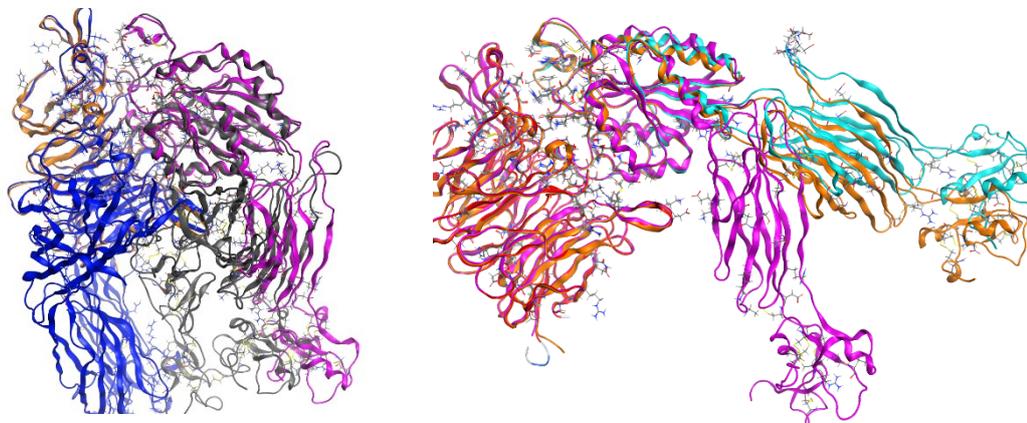


Figure 64: superposition of PDB entries representing $\alpha IIb \beta 3$. The two figures try to sum up the differences between the different conformational states and the movements that in particular the subunit β performs. a) comparison between the bent-state (PDB 3FCS) and a less closed state (PDB 7U60) b) Differences between PDB entries 7U60 and other two entries representing two open state of IN $\alpha IIb \beta 3$ (PDB entry 3FCU: less opened; PDB entry 2VDO: completely opened)

Conformational changes in the top face of β -I and β -propeller are more difficult to observe, but differences in the loops involved RGD-ligand binding are highlighted by slight differences in the superposition. Confirming the literature, the open-high affinity state of integrin is achieved by the α helix 7 and α helix 1 in the β I-like domain moving downward on ligand bindings and the hybrid domain swinging outward from the β -propeller upon integrin activation (**Figure 62-64**).

From analysis on contacts between ligands and integrin subunits it is confirmed the involvement of the integrin residues in β -I and β -propeller in the RGD binding, together with the help of the three coordinating metal ion binding sites (presence of Mg^{2+} in $\alpha5\beta1$ and Mn^{2+} in $\alphaV\beta3$).

Analyzing the contacts between the RGD ligand and the integrin subunits in $\alphaV\beta3$ it can be confirmed that the RGD peptide creates interface using the Asp150 and Asp218 of the αV chain with the ligand Arg, and the Asn215 and Arg216 with Asp5003, finding also a contact with a Mn^{2+} in MIDAS, for $\beta3$. This is consistent with the literature (**Figure 65a**).

a)

Type	ChainA	PosA	SetA	ChainB	PosB	SetB	Energy	Dist	BB	Freq
IH	5-1L5G_1.C	1	Arg5001	1-1L5G_1.A	150	Asp150	-13.36	2.76	--	2
IH	5-1L5G_1.C	1	Arg5001	1-1L5G_1.A	218	Asp218	-34.31	3.03	--	6
IM	5-1L5G_1.C	3	Asp5003	4-1L5G_1.B	3	MN4003	-19.54	2.17	--	2
H	5-1L5G_1.C	3	Asp5003	3-1L5G_1.B	162	Asn215	-9.10	2.79	-*	2
H	5-1L5G_1.C	3	Asp5003	3-1L5G_1.B	163	Arg216	-0.90	3.02	bb	1

b)

Type	ChainA	PosA	SetA	ChainB	PosB	SetB	Energy	Dist	BB	Freq
H	6-4MMX_1.C	28	Val1444	9-4MMX_1.F	5	MANS	-1.50	3.15	b-	1
H	6-4MMX_1.C	29	Arg1445	9-4MMX_1.F	5	MANS	-2.40	2.65	b-	1
H	6-4MMX_1.C	52	Ser1468	9-4MMX_1.F	5	MANS	-1.50	2.89	--	1
I	6-4MMX_1.C	77	Arg1493	1-4MMX_1.A	150	Asp150	-2.36	3.39	--	1
IH	6-4MMX_1.C	77	Arg1493	1-4MMX_1.A	218	Asp218	-21.52	2.81	--	3
IH	6-4MMX_1.C	79	Asp1495	4-4MMX_1.B	3	MN700	-16.26	2.14	--	2
H	6-4MMX_1.C	79	Asp1495	3-4MMX_1.B	121	Ser121	-0.50	3.22	--	1
H	6-4MMX_1.C	79	Asp1495	3-4MMX_1.B	122	Tyr122	-3.40	2.82	b-	1
H	6-4MMX_1.C	79	Asp1495	3-4MMX_1.B	123	Ser123	-5.40	2.62	b-	1
H	6-4MMX_1.C	79	Asp1495	3-4MMX_1.B	215	Asn215	-2.10	3.45	b-	1

Figure 65: contacts analysis between: a) the RGD ligand of 1L5G and the integrin subunits of $\alphaV\beta3$; b) the 10FNIII module in 4MMX.C and the integrin subunit of $\alphaV\beta3$

By comparison with the contacts in 4MMX between fibronectin and integrin, it is confirmed that the interfaces between the FN and αV fall in these residues, while between $\beta3$ and FN it seems that the interaction are more complex, involving with Asp1495 always Mn^{2+} and Asn215, but lacking the Arg216, that sees its place taken by Ser121, Tyr122 and Ser123 (**Figure 65b**).

This could be due to the fact that the 10FNIII presents a more extended interface area which to interact with the β subunit and interactions are more focused on the exterior of the head-piece, while RGD ligands tend to bury themselves giving life to different interactions.

Contacts between the RGD ligands were also analyzed for the $\alpha5\beta1$ structure relatives to the RGD ligands present (**Figure 66**). The role of the involved residues found in literature is confirmed.

a)

Type	ChainA	PosA	SetA	ChainB	PosB	SetB	Energy	Dist	BB	Freq
IH	4-3VI4.B	2	MG592	7-3VI4.I	3	Asp5003	-25.73	1.81	--	2
H	3-3VI4.B	132	Ser132	7-3VI4.I	3	Asp5003	-0.70	3.15	--	1
H	3-3VI4.B	133	Tyr133	7-3VI4.I	3	Asp5003	-4.00	2.89	b-	1
H	3-3VI4.B	134	Ser134	7-3VI4.I	3	Asp5003	-0.60	3.51	b-	1
H	1-3VI4.A	221	Gln221	7-3VI4.I	1	Arg5001	-8.10	3.01	--	2
H	3-3VI4.B	224	Asn224	7-3VI4.I	3	Asp5003	-6.20	2.93	*	2
H	3-3VI4.B	225	Leu225	7-3VI4.I	3	Asp5003	-1.20	3.14	bb	1
IH	1-3VI4.A	227	Asp227	7-3VI4.I	1	Arg5001	-34.62	2.97	--	6
H	3-3VI4.B	227	Ser227	7-3VI4.I	2	Gly5002	-1.40	2.86	b-	1

b)

Type	ChainA	PosA	SetA	ChainB	PosB	SetB	Energy	Dist	BB	Freq
H	7-4WK0_1.C	1	Arg1524	8-4WK0_1.C	3	HOH1603	-1.60	2.70	b-	1
H	7-4WK0_1.C	1	Arg1524	8-4WK0_1.C	4	HOH1604	-2.50	2.81	b-	1
H	7-4WK0_1.C	1	Arg1524	3-4WK0_1.A	117	HOH717	-6.40	2.89	--	2
IH	7-4WK0_1.C	1	Arg1524	1-4WK0_1.A	227	Asp227	-34.44	2.99	--	6
H	8-4WK0_1.C	1	HOH1601	6-4WK0_1.B	64	HOH664	-1.00	2.93	--	1
H	8-4WK0_1.C	1	HOH1601	3-4WK0_1.A	221	HOH821	-1.50	2.75	--	1
H	8-4WK0_1.C	1	HOH1601	1-4WK0_1.A	227	Asp227	-2.50	2.75	--	1
H	7-4WK0_1.C	2	Gly1525	8-4WK0_1.C	1	HOH1601	-2.90	2.91	b-	1
H	7-4WK0_1.C	2	Gly1525	8-4WK0_1.C	2	HOH1602	-1.00	3.22	b-	1
H	7-4WK0_1.C	2	Gly1525	8-4WK0_1.C	4	HOH1604	-2.30	2.63	b-	1
H	8-4WK0_1.C	2	HOH1602	6-4WK0_1.B	37	HOH637	-1.30	2.74	--	1
IH	7-4WK0_1.C	3	Asp1526	5-4WK0_1.B	1	MG501	-26.73	1.85	--	2
H	7-4WK0_1.C	3	Asp1526	8-4WK0_1.C	5	HOH1605	-2.90	2.70	*	2
H	7-4WK0_1.C	3	Asp1526	4-4WK0_1.B	123	Ser132	-0.80	3.37	--	1
H	7-4WK0_1.C	3	Asp1526	4-4WK0_1.B	124	Tyr133	-0.50	3.57	b-	1
H	7-4WK0_1.C	3	Asp1526	4-4WK0_1.B	214	Gly223	-0.50	3.34	b-	1
H	7-4WK0_1.C	3	Asp1526	4-4WK0_1.B	215	Asn224	-4.80	2.87	b-	1
H	7-4WK0_1.C	3	Asp1526	4-4WK0_1.B	216	Leu225	-2.10	2.87	bb	1
H	8-4WK0_1.C	3	HOH1603	6-4WK0_1.B	172	HOH772	-1.30	2.79	--	1
H	8-4WK0_1.C	6	HOH1606	6-4WK0_1.B	128	HOH728	-1.60	2.67	--	1
H	8-4WK0_1.C	6	HOH1606	6-4WK0_1.B	237	HOH837	-1.50	2.70	--	1

c)

Type	ChainA	PosA	SetA	ChainB	PosB	SetB	Energy	Dist	BB	Freq	Cons
H	3-4WK2_1.C	2	Arg5001	1-4WK2_1.A	221	Gln221	-7.20	2.84	--	1	0
IH	3-4WK2_1.C	2	Arg5001	1-4WK2_1.A	227	Asp227	-33.35	3.02	--	6	0
H	3-4WK2_1.C	3	Gly5002	13-4WK2_1.B	13	HOH613	-2.40	2.67	b-	1	0
H	3-4WK2_1.C	3	Gly5002	12-4WK2_1.A	61	HOH661	-2.80	2.93	b-	1	0
IM	3-4WK2_1.C	4	Asp5003	11-4WK2_1.B	1	MG501	-23.87	1.87	--	2	0
H	3-4WK2_1.C	4	Asp5003	2-4WK2_1.B	134	Ser132	-0.60	3.20	--	1	0
H	3-4WK2_1.C	4	Asp5003	2-4WK2_1.B	125	Tyr133	-4.80	2.82	b-	1	0
H	3-4WK2_1.C	4	Asp5003	2-4WK2_1.B	126	Ser134	-0.90	3.43	b-	1	0
H	3-4WK2_1.C	4	Asp5003	2-4WK2_1.B	215	Gly223	-0.50	3.47	b-	1	0
H	3-4WK2_1.C	4	Asp5003	2-4WK2_1.B	216	Asn224	-4.00	2.96	*	2	0
H	3-4WK2_1.C	4	Asp5003	2-4WK2_1.B	217	Leu225	-1.60	3.06	bb	1	0
H	3-4WK2_1.C	5	Ser5004	13-4WK2_1.B	13	HOH613	-1.30	2.73	b-	1	0

d)

Type	ChainA	PosA	SetA	ChainB	PosB	SetB	Energy	Dist	BB	Freq
H	15-4WK4_1.C	1	HOH101	1-4WK4_1.A	227	Asp227	-3.50	2.62	--	1
H	15-4WK4_1.C	2	HOH102	14-4WK4_1.B	55	HOH655	-0.70	2.94	--	1
H	15-4WK4_1.C	2	HOH102	2-4WK4_1.B	225	Ser227	-1.70	2.70	--	1
H	3-4WK4_1.C	3	Arg3	1-4WK4_1.A	221	Gln221	-8.20	2.94	--	2
IH	3-4WK4_1.C	3	Arg3	1-4WK4_1.A	227	Asp227	-33.40	3.01	--	6
H	15-4WK4_1.C	3	HOH103	14-4WK4_1.B	55	HOH655	-1.30	2.91	--	1
H	15-4WK4_1.C	3	HOH103	2-4WK4_1.B	257	Asp259	-0.60	2.84	b-	1
H	3-4WK4_1.C	4	Gly4	15-4WK4_1.C	1	HOH101	-3.10	2.86	b-	1
H	3-4WK4_1.C	4	Gly4	15-4WK4_1.C	2	HOH102	-1.20	2.79	b-	1
H	15-4WK4_1.C	4	HOH104	13-4WK4_1.A	108	HOH709	-1.40	2.79	--	1
IH	3-4WK4_1.C	5	Asp5	12-4WK4_1.B	1	MG501	-23.78	1.86	--	2
H	3-4WK4_1.C	5	Asp5	2-4WK4_1.B	130	Ser132	-0.80	3.21	--	1
H	3-4WK4_1.C	5	Asp5	2-4WK4_1.B	131	Tyr133	-4.60	2.89	b-	1
H	3-4WK4_1.C	5	Asp5	2-4WK4_1.B	132	Ser134	-1.10	3.39	b-	1
H	3-4WK4_1.C	5	Asp5	2-4WK4_1.B	222	Asn224	-2.20	3.13	*	2
H	3-4WK4_1.C	5	Asp5	2-4WK4_1.B	223	Leu225	-2.00	3.03	bb	1
H	3-4WK4_1.C	7	Trp7	2-4WK4_1.B	222	Asn224	-3.80	2.87	b-	1

Figure 66: contacts analysis between RGD ligands and their relative $\alpha5\beta1$ entry : a) RGD linear peptide in 3VI4.I; b) RGD linear peptide in 4WK0.C, c) GRGDSP linear peptide in 4WK2.C, d) ACRGDWC cyclic peptide in 4WK4.C

The results of refine docking simulations of the RGD ligands to evaluate the S score and the involved affinity of the ligands are listed below (**figure 67-68-69-70-71-72-73**). As stated in the

methods, S score of crystallography structures lacking an RGD ligand were evaluated by using the RGD ligand/ligands of the respective RGD-bound structures. S score is used as standard to evaluate the next docking simulations.

	mol	receptor	rseq	mseq	S	rmsd_refine	E_conf	E_place	E_score1	E_refine	E_score2
1			1	1	-10.0126	2.4446	-226.2545	-56.5720	-9.4495	-81.7096	-10.0126
2			1	1	-70.1308	0.8752	6406.7432	-23.9450		-70.1308	
3			1	1	-9.3249	2.9172	-224.7239	-22.3909	-14.0158	-71.3899	-9.3249
4			1	1	-9.6543	2.1283	-233.8920	-93.6254	-11.7149	-77.5580	-9.6543
5			1	1	-7.8715	3.4069	-223.2934	-49.8671	-18.7617	-52.1091	-7.8715
6			1	1	-8.5768	2.1687	-217.6971	-74.2958	-10.8267	-57.1554	-8.5768
7			1	1	-10.5259	2.2491	-218.3247	-9.1131	-6.3768	-77.2436	-10.5259
8			1	1	-8.2309	2.5691	-218.8483	-48.5399	-9.9032	-51.7775	-8.2309
9			1	1	-8.2156	1.4646	-223.5677	-86.8762	-11.4126	-51.6510	-8.2156
10			1	1	-8.4274	1.9166	-220.8239	-63.2286	-11.4396	-54.6952	-8.4274
11			1	1	-7.6058	2.2464	-223.5357	-27.9393	-12.9482	-47.3351	-7.6058
12			1	1	-7.6797	4.0949	-228.1146	-14.1352	-11.3234	-43.8717	-7.6797

Figure 67: docking score of the cyclic RGD peptide 1L5G.C to all the entries representing $\alpha V\beta 3$ in ligand binding site 1L5G.C; in entry 2 it is also scored the docking between 10FnIII (4MMX.C) and 4MMX: 1) RGD-1L5G; 2) 10FnIII-4MMX; 3) RGD-4MMX; 4) RGD-6NAJ; 5) RGD-4G1M; 6) RGD-1M1X; 7) RGD-6AVU_withLM609; 8) RGD-6AVU_withoutLM609; 9) RGD-6AVR_withLM609; 10) RGD-6AVR_withoutLM609; 11) RGD-6AV_withLM609; 12) RGD-6AVQ_withoutLM609;

	mol	receptor	rseq	mseq	S	rmsd_refine	E_conf	E_place	E_score1	E_refine	E_score2
1			1	1	-10.2235	1.5210	-390.3882	-93.6715	-14.5580	-104.5107	-10.2235
2			1	1	-10.2171	1.5194	-390.3998	-93.6715	-14.5580	-104.3033	-10.2171
3			1	1	-10.7840	1.5159	-387.3074	-64.0915	-13.8059	-106.0788	-10.7840
4			1	1	-10.7957	1.5390	-387.3019	-64.0915	-13.8059	-106.2811	-10.7957
5			1	1	-5.9043	1.9441	-397.2172	-54.5909	-15.2309	-36.3303	-5.9043
6			1	1	-10.8917	2.0786	-398.0578	-97.2134	-16.5717	-96.3231	-10.8917
7			1	1	-10.9742	1.4525	-400.0625	-46.8421	-15.1668	-103.3062	-10.9742
8			1	1	-10.6358	1.9162	-398.7374	-76.2546	-13.0035	-97.7392	-10.6358
9			1	1	-8.5670	2.5148	-399.1753	-50.3843	-13.2908	-69.7175	-8.5670
10			1	1	-8.5730	2.5174	-399.1621	-50.3843	-13.2908	-69.7390	-8.5730
11			1	1	-10.7573	2.3772	-387.3031	-60.2951	-18.2684	-109.0351	-10.7573

Figure 68: docking score of the linear RGD peptide 3VI4.I to all the entries representing $\alpha 5\beta 1$ in ligand binding site 3VI4.I: 1) RGD-3VI3_withSG19Fab; 2) RGD-3VI3_withoutSG19Fab; 3) RGD_3VI4_withSG19Fab; 4) RGD_3VI4_withoutSG19Fab; 5) RGD-4WKJ; 6) RGD-4WK0; 7) RGD-4WK2; 8) RGD-4WK4; 9) RGD-7NWL_withTS2/16Fv-clasp; 10) RGD-7NWL_withoutTS2/16; 11) RGD-7NXD.

	mol	receptor	rseq	mseq	S	rmsd_refine	E_conf	E_place	E_score1	E_refine	E_score2
1			1	1	-10.6760	1.8630	-384.2348	-60.1948	-14.1618	-114.8062	-10.6760
2			1	1	-10.7928	2.4068	-381.9468	-62.8479	-13.8475	-115.3498	-10.7928
3			1	1	-10.7728	2.0060	-382.4599	-77.1091	-13.2672	-105.5525	-10.7728
4			1	1	-10.7797	2.0052	-382.4702	-77.1091	-13.2672	-105.5985	-10.7797
5			1	1	-5.9766	1.7222	-398.4121	-61.8425	-13.2398	-34.1571	-5.9766
6			1	1	-11.2313	2.3824	-402.1307	-62.6152	-17.4291	-104.5695	-11.2313
7			1	1	-10.8026	2.3828	-402.5845	-85.3173	-13.7165	-101.9678	-10.8026
8			1	1	-10.1723	1.8533	-394.6674	-72.8380	-11.2188	-93.9184	-10.1723
9			1	1	-8.3821	1.4051	-398.3482	-91.0755	-12.3628	-64.8842	-8.3821
10			1	1	-8.3930	1.3925	-398.3585	-91.0755	-12.3628	-64.9881	-8.3930
11			1	1	-10.5091	1.9685	-393.1290	-80.2950	-14.9268	-97.8957	-10.5091

Figure 69: docking score of the linear RGD peptide 4WK0.C to all the entries representing $\alpha 5\beta 1$ in ligand binding site 4WK0.C: 1) RGD-3VI3_withSG19Fab; 2) RGD-3VI3_withoutSG19Fab; 3) RGD_3VI4_withSG19Fab; 4) RGD_3VI4_withoutSG19Fab; 5) RGD-4WJK; 6) RGD-4WK0; 7) RGD-4WK2; 8) RGD-4WK4; 9) RGD-7NWL_withTS2/16Fv-clasp; 10) RGD-7NWL_withoutTS2/16; 11) RGD-7NXD.

	mol	receptor	rseq	mseq	S	rmsd_refine	E_conf	E_place	E_score1	E_refine	E_score2
1			1	1	-12.6186	3.2896	-276.0728	-53.3138	-14.8206	-131.6065	-12.6186
2			1	1	-12.6123	3.2801	-276.0380	-53.3138	-14.8206	-131.6930	-12.6123
3			1	1	-13.2264	2.3645	-263.5567	-47.4364	-17.1497	-138.5297	-13.2264
4			1	1	-13.2096	2.3636	-263.5956	-47.4364	-17.1497	-138.4431	-13.2096
5			1	1	-8.4555	2.7513	-278.8760	-81.4227	-15.7487	-60.0983	-8.4555
6			1	1	-12.5445	2.3852	-273.2415	-46.3423	-16.2423	-105.7914	-12.5445
7			1	1	-12.1097	3.1502	-282.9539	-59.3148	-14.6244	-109.6061	-12.1097
8			1	1	-12.1649	2.5540	-276.3649	-33.7377	-12.9917	-106.7612	-12.1649
9			1	1	-10.0389	3.7194	-273.6851	-68.8705	-18.8827	-82.3044	-10.0389
10			1	1	-10.5881	3.3841	-279.8319	-57.5442	-13.1884	-91.9320	-10.5881
11			1	1	-12.5861	1.8702	-261.7253	-52.2822	-14.8809	-123.9760	-12.5861

Figure 70: docking score of the linear RGD peptide 4WK2.C to all the entries representing $\alpha 5\beta 1$ in ligand binding site 4WK2.C: 1) RGD-3VI3_withSG19Fab; 2) RGD-3VI3_withoutSG19Fab; 3) RGD_3VI4_withSG19Fab; 4) RGD_3VI4_withoutSG19Fab; 5) RGD-4WJK; 6) RGD-4WK0; 7) RGD-4WK2; 8) RGD-4WK4; 9) RGD-7NWL_withTS2/16Fv-clasp; 10) RGD-7NWL_withoutTS2/16; 11) RGD-7NXD.

	mol	receptor	rseq	mseq	S	rmsd_refine	E_conf	E_place	E_score1	E_refine	E_score2
1			1	1	-11.4142	2.3048	-240.7027	-33.0021	-12.6002	-96.2792	-11.4142
2			1	1	-11.6251	2.7503	-266.0939	-58.3806	-10.7061	-102.6357	-11.6251
3			1	1	-11.6216	2.7497	-266.0905	-58.3806	-10.7061	-102.6318	-11.6216
4			1	1	-7.3882	3.5198	-258.2656	-37.2361	-11.9295	-44.0043	-7.3882
5			1	1	-12.8557	3.0399	-256.3327	-24.6369	-20.5936	-111.2589	-12.8557
6			1	1	-12.3986	3.0015	-251.2587	-67.9521	-11.4475	-110.0353	-12.3986
7			1	1	-12.2280	2.1089	-248.7252	-106.1627	-18.4317	-103.3925	-12.2280
8			1	1	-9.8187	2.5287	-244.9957	-33.1214	-13.3542	-74.7637	-9.8187
9			1	1	-9.7490	2.5101	-240.2829	-75.4118	-12.4216	-74.1134	-9.7490
10			1	1	-12.0688	2.7424	-251.1523	-67.7788	-10.2686	-112.0730	-12.0688

Figure 71: docking score of the cyclic RGD peptide 4WK4.C to all the entries representing $\alpha 5\beta 1$ in ligand binding site 4WK4.C: 1) RGD-3VI3_withSG19Fab; 2) RGD_3VI4_withSG19Fab; 3) RGD_3VI4_withoutSG19Fab; 4) RGD-4WJK; 5) RGD-4WK0; 6) RGD-4WK2; 7) RGD-4WK4; 8) RGD-7NWL_withTS2/16Fv-clasp; 9) RGD-7NWL_withoutTS2/16; 10) RGD-7NXD.

	mol	receptor	rseq	mseq	S	rmsd_refine	E_conf	E_place	E_score1	E_refine	E_score2
1	7U60	2VC2	1	1	-12.7513	4.3384	-315.1890	-77.6459	-11.8195	-109.9025	-12.7513
2	7U60	2VC2	1	1	-11.0792	3.2799	-309.0137	-70.2176	-12.8463	-83.1555	-11.0792
3	7U60	2VDM	1	1	-11.0093	3.9637	-310.9868	-84.7645	-11.8391	-92.4334	-11.0093
4	7U60	2VDM	1	1	-11.7422	1.8010	-315.3935	-99.8014	-18.6050	-102.5324	-11.7422
5	7U60	2VDN	1	1	-12.3125	2.0429	-321.3913	-96.9064	-12.4669	-107.4741	-12.3125
6	7U60	2VDN	1	1	-12.0318	1.9055	-314.6323	-77.7074	-13.3840	-104.8748	-12.0318
7	7U60	2VDO	1	1	-11.6857	2.4215	-314.0317	-79.1420	-17.3163	-94.3870	-11.6857
8	7U60	2VDO	1	1	-9.0443	1.7557	-311.2931	-46.4889	-16.2748	-63.8448	-9.0443
9	7U60	2VDP	1	1	-11.2398	3.3803	-321.0634	-34.1681	-12.9796	-87.8634	-11.2398
10	7U60	2VDP	1	1	-11.9664	3.4462	-312.6049	-63.9800	-12.5108	-102.0934	-11.9664
11	7U60	2VDQ	1	1	-10.0567	1.6580	-311.3283	-28.5735	-14.9556	-77.7250	-10.0567
12	7U60	2VDQ	1	1	-10.6496	3.2613	-317.1350	-40.1707	-12.8933	-93.7871	-10.6496
13	7U60	2VDR	1	1	-12.6355	3.5553	-318.0149	-38.9906	-15.2629	-107.6013	-12.6355
14	7U60	2VDR	1	1	-12.8839	3.6833	-318.5924	-38.9906	-15.2629	-108.8014	-12.8839
15	7U60	3FCS	1	1	-8.7305	2.2642	-309.1219	-88.5728	-11.1677	-53.0228	-8.7305
16	7U60	3FCU	1	1	-11.1277	1.8346	-315.1455	-85.2755	-17.4202	-96.3597	-11.1277
17	7U60	3NID	1	1	-7.2596	1.8372	-316.4763	-48.0275	-14.4549	-41.2167	-7.2596
18	7U60	3NID	1	1	-7.4228	3.5490	-317.3265	-32.8912	-12.7338	-45.0729	-7.4228
19	7U60	3T3P	1	1	-7.3797	2.6106	-316.0271	-93.5425	-13.8832	-46.8023	-7.3797
20	7U60	3T3P	1	1	-7.5623	2.4767	-316.6924	-61.4786	-13.7578	-47.3284	-7.5623
21	7U60	3ZDX	1	1	-7.4287	2.8273	-318.4836	-69.4304	-13.0041	-43.4500	-7.4287
22	7U60	3ZDX	1	1	-7.6894	2.2350	-316.5328	-90.8883	-13.1717	-46.8078	-7.6894
23	7U60	3ZDY	1	1	-10.6238	2.2082	-314.4067	-63.5299	-14.2509	-83.8696	-10.6238
24	7U60	3ZDY	1	1	-10.8511	2.0396	-314.1940	-74.7071	-14.2541	-91.4030	-10.8511
25	7U60	4Z7N	1	1	-6.8289	5.5256	-315.4823	-20.4385	-18.4995	-35.8767	-6.8289
26	7U60	7U60	1	1	-9.8026	2.1597	-300.9470	-74.9612	-13.2127	-75.6470	-9.8026

Figure 72: docking score of the linear RGD peptide 7U60.M to all the entries representing *alpha*IIb β 3 in ligand binding site 7U60.M: 1) RGD-2VC2_with10E5mAb (or Fab); 2) RGD-2VC2_without10E5mAb; 3) RGD-2VDM_with10E5mAb; 4) RGD-2VDM_without10E5mAb; 5) RGD-2VDN_with10E5mAb; 6) RGD-2VDN_without10E5mAb; 7) RGD-2VDO_with10E5mAb; 8) RGD-2VDO_without10E5mAb; 9) RGD-2VDP_with10E5mAb; 10) RGD-2VDP_without10E5mAb; 11) RGD-2VDQ_with10E5mAb; 12) RGD-2VDQ_without10E5mAb; 13) RGD-2VDR_with10E5mAb; 14) RGD-2VDR_without10E5mAb; 15) RGD-3FCS; 16)) RGD-3FCU; 17) RGD-3NID_with10E5mAb; 18) RGD-3T3P_without10E5mAb; 19) RGD-3NID_with10E5mAb; 20) RGD-3NID_without10E5mAb; 21) RGD-3ZDX_with10E5mAb; 22) RGD-3ZDX_without10E5mAb; 23) RGD-3ZDY_without10E5mAb; 24) RGD-3ZDY_without10E5mAb; 25) RGD-4Z7N; 26) RGD-7U60;

	mol	receptor	rseq	mseq	S	rmsd_refine	E_conf	E_place	E_score1	E_refine	E_score2
1	3ZDY	3ZDY	1	1	-12.4951	3.0622	-279.2842	-61.2763	-18.3039	-105.3374	-12.4951
2	3ZDZ	3ZDZ	1	1	-10.5782	2.1420	-222.7878	-78.6479	-14.6854	-83.7818	-10.5782
3	3ZE0	3ZE0	1	1	-12.2022	2.1072	-288.0935	-151.3517	-14.8838	-110.5224	-12.2022
4	3ZE1	3ZE1	1	1	-12.0645	2.2949	-281.1248	-80.8277	-13.4246	-106.7544	-12.0645
5	3ZE2	3ZE2	1	1	-10.3646	3.0315	-210.8210	-70.1529	-14.4799	-81.9705	-10.3646

Figure 73: refine docking score for PDB entries representing *alpha*IIb β 3.

Docking between various RGD ligands and different integrin structures are listed in the table below (Figure 74)

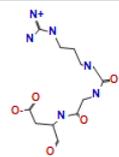
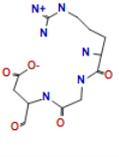
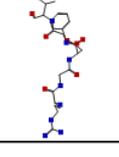
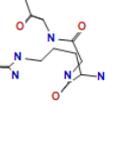
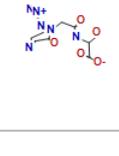
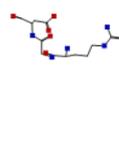
	mol	receptor	rseq	mseq	S	rmsd_refine	E_conf	E_place	E_score1	E_refine	E_score2
a)			1	1	-8.1707	1.8706	-382.8947	-50.7949	-12.2351	-60.5760	-8.1707
b)			1	1	-7.5437	2.5135	-385.3893	-50.7086	-11.0985	-42.4346	-7.5437
c)			1	1	-8.6067	2.1661	-312.3312	-79.3234	-7.3395	-56.1592	-8.6067
d)			1	1	-9.0930	3.5762	-393.3517	-62.3424	-15.5597	-62.4052	-9.0930
e)			1	1	-10.5091	1.9685	-393.1290	-80.2950	-14.9268	-97.8957	-10.5091
f)			1	1	-11.5499	1.2181	-388.3779	-85.5410	-12.3989	-121.9513	-11.5499

Figure 74: docking between different RGD ligands and integrin structures in the binding site 1L5G.C: a)3VI4.I-1L5G; b) 4WK0.C-1L5G, c)7U60.M-1L5G; d) 4WK0.C-7NWL_withoutantibody; e)4WK0.C-7NXD; f) 3VI4.I-3VI3.

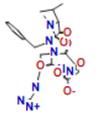
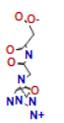
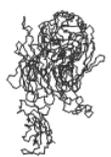
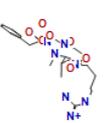
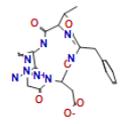
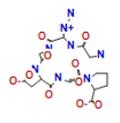
	mol	receptor	rseq	mseq	S	rmsd_refine	E_conf	E_place	E_score1	E_refine	E_score2
a)			1	1	-6.6121	5.1980	-227.0968	-1.0292	-8.2106	-33.8680	-6.6121
b)			1	1	-6.5469	1.9068	-399.2349	-75.8749	-15.2707	-39.5660	-6.5469
c)			1	1	-9.5408	1.7008	-398.2750	-90.7569	-14.2313	-74.5886	-9.5408

Figure 75: docking between different RGD ligands and integrin structures in the binding site 7U60.M: a) 1L5G.C-7U60; b) 3VI4.I-7U60; c) 4WK0.C-7U60.

	mol	receptor	rseq	mseq	S	rmsd_refine	E_conf	E_place	E_score1	E_refine	E_score2
a)			1	1	-10.5154	2.3363	-225.7230	-82.7882	-17.9470	-93.6239	-10.5154
b)			1	1	-11.9016	4.3841	-223.2744	-69.4324	-18.8917	-122.4144	-11.9016
c)			1	1	-8.9991	2.0079	-225.4294	-99.4771	-11.7298	-69.2792	-8.9991
d)			1	1	-10.7751	2.3145	-228.9534	-55.9703	-13.4449	-102.8135	-10.7751
e)			1	1	-10.3769	3.5049	-268.8851	-45.8404	-11.9070	-89.9467	-10.3769

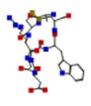
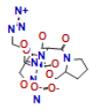
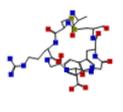
f)			1	1	-9.5078	1.5323	-261.8124	-83.4127	-17.5006	-65.1405	-9.5078
g)			1	1	-13.2597	2.2841	-258.9433	-32.6271	-13.4447	-122.5224	-13.2597
h)			1	1	-11.5746	3.2925	-253.1171	-34.2325	-14.0306	-105.4210	-11.5746

Figure 76: docking between different RGD ligands and integrin structures in the binding site 3VI4.I: a) 1L5G.C-3VI3; b) 1L5G.C-3VI4, c) 1L5G-7NWL_without_mAb; d) 1L5G.C-7NXD; e) 4WK2.C-7NWL_without_mAb; f) 4WK4.C-7NWL_withoutmAb; g) 4WK2-7NXD; h) 4WK4-7NXD.

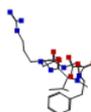
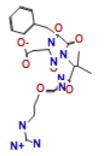
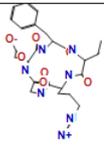
	mol	receptor	rseq	mseq	S	rmsd_refine	E_conf	E_place	E_score1	E_refine	E_score2
a)			1	1	-7.7338	7.6799	-189.9264	-16.1392	-18.6176	-29.2164	-7.7338
b)			1	1	-10.7876	1.7635	-219.5034	-53.6094	-14.7449	-99.4597	-10.7876
c)			1	1	-12.1222	3.4239	-226.9114	-1.9621	-18.4766	-106.1630	-12.1222

Figure 77: docking between different RGD ligands and integrin structures in the binding site 4WK0.C: a) 1L5G.C-7NWL_without_mAb; b) 1L5G.C-7NXD, c) 1L5G-4WK0.

mAbs found in the structures were also studied for their role in inactivating integrins and lowering affinity for ligands, but in some cases it seem that performing docking analysis without the presence of mAb gives better affinity, probably due to steric hindrance by the presence of the mAb.

It is also confirmed that bigger RGD ligands presents higher affinity for the RGD binding site. Also cyclic peptide tend to present a higher affinity than linear peptide, for the reasons already explained. Protein-protein docking between 10FnIII and $\alpha v \beta 3$ gives an S score of circa -70, a high value that confirms the fact that bigger RGD ligands present much more interactions with the integrin headpiece; in particular it is also known the mechanism of synergy binding to the PHSRN motif in 9FnIII, that further enhance affinity (Figure 74-75-76-77).

Results of docking between cilengitide and the three integrin types are listed in the figure below (Figure 78-79-80-81).

	mol	receptor	rseq	mseq	S	rmsd_refine	E_conf	E_place	E_score1	E_refine	E_score2
1			1	1	-9.8963	3.3568	-237.8019	-41.3170	-15.9673	-91.3001	-9.8963
2			1	1	-7.8847	6.7121	-155.2718	-58.2066	-9.2456	-43.5359	-7.8847
3			1	1	20.4409	3.7107	245.9406	-58.7136	-11.6291	349.8085	20.4409
4			1	1	-8.3259	2.2380	-157.6622	-46.9713	-6.9862	-48.4722	-8.3259
5			1	1	-6.8738	2.3118	-163.7704	-77.9771	-8.9529	-35.7895	-6.8738
6			1	1	-16.0935	2.0291	-166.8613	-109.9130	-8.5366	-62.4291	-16.0935
7			1	1	-7.2601	2.1431	-141.4510	-65.4573	-10.2273	-44.0810	-7.2601
8			1	7	-6.4864	1.9733	-480.0711	-146.9180	-13.7543	-43.5312	-6.4864

Figure 78: results of the docking between cilengitide and the $\alpha V\beta 3$ PDB entries, performed in the 1L5G.C ligand site. Legend of entries of the database: 1) 1L5G; 2) 4MMX; 3) 6NAJ; 4) 6AVU; 5) 6AVR; 6) 6AVQ; 7) 1M1X; 8) 4G1M.

	mol	receptor	rseq	mseq	S	rmsd_refine	E_conf	E_place	E_score1	E_refine	E_score2
1			1	1	-6.8964	2.4928	-168.6979	-67.5273	-9.9733	-38.4796	-6.8964
2			1	1	-6.6542	1.5212	-173.4963	-45.5112	-6.2526	-42.7069	-6.6542
3			1	1	-7.1843	1.8341	-166.9434	-59.0014	-9.0875	-42.9708	-7.1843
4			1	1	-10.3292	2.2511	-136.9420	-80.8404	-9.1781	-81.5594	-10.3292
5			1	2	-7.3661	2.0347	-173.3127	-48.1700	-16.3882	-41.0391	-7.3661
6			1	1	-5.9018	2.2493	-163.1268	-29.5792	-14.1864	-29.7494	-5.9018
7			1	1	-7.5033	2.2240	-165.5514	-32.5964	-7.5715	-43.3360	-7.5033
8			1	1	-6.5629	2.2894	-154.3282	-65.8929	-8.2142	-39.4207	-6.5629

Figure 79: results of the docking between cilengitide and the $\alpha 5\beta 1$ PDB entries, performed in the 3VI4.I ligand site. Legend of entries of the database: 1) 3VI3; 2) 3VI4; 3) 7NWL ; 4) 7NXD ; 5) 4WK0; 6) 4WJK ; 7) 4WK2; 8) 4WK4

	mol	receptor	rseq	mseq	S	rmsd_refine	E_conf	E_place	E_score1	E_refine	E_score2
1			1	1	-6.3724	3.3323	-168.7993	-66.6763	-9.4516	-28.7626	-6.3724
2			1	7	-7.1555	2.5534	-478.7958	-33.9223	-15.0154	-103.4393	-7.1555
3			1	1	-6.9669	1.7034	-145.9025	-89.3638	-9.6562	-42.1884	-6.9669
4			1	1	-7.3071	1.9283	-162.5375	-59.0809	-11.9020	-43.7989	-7.3071
5			1	1	-6.9356	2.6768	-172.9757	-13.2476	-15.2269	-49.7169	-6.9356
6			1	1	-6.0854	1.6401	-164.7439	-44.8826	-14.1010	-28.6908	-6.0854
7			1	1	-6.3276	2.2845	-163.5715	-36.8670	-7.0854	-34.3892	-6.3276
8			1	1	-7.2586	3.7489	-162.9308	-19.3981	-8.3022	-40.7753	-7.2586

Figure 80: results of the docking between cilengitide and the $\alpha 5b1$ PDB entries, performed in the 4WK0.C ligand site. Legend of entries of the database: 1) 3VI3; 2) 3VI4; 3) 7NWL ; 4)7NXD ; 5) 4WK0; 6) 4WJK ; 7) 4WK2; 8) 4WK4

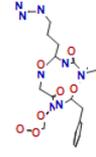
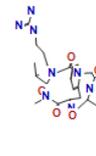
	mol	receptor	rseq	mseq	S	rmsd_refine	E_conf	E_place	E_score1	E_refine	E_score2
a)			1	1	-6.3008	1.8879	-165.6257	-59.5879	-8.7118	-34.3603	-6.3008
b)			1	1	-6.4485	1.9508	-164.2384	-101.3658	-10.0819	-22.8035	-6.4485

Figure 81: results of the docking between Cilengitide and 4WK4 PDB entry, performed in a) the 4WK2.C ligand site and b) the 4WK4.C ligand site, respectively.

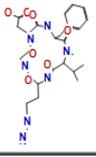
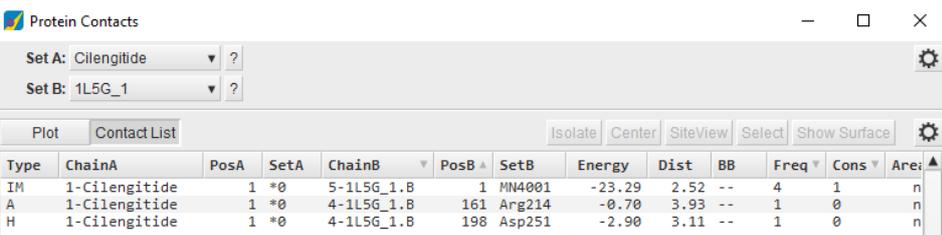
	mol	receptor	rseq	mseq	S	rmsd_refine	E_conf	E_place	E_score1	E_refine	E_score2
a)			1	1	-7.0185	3.2317	-168.2741	-46.4829	-13.6983	-37.2889	-7.0185
b)			1	1	-6.2499	1.8955	-160.6521	-56.7488	-9.0920	-34.2889	-6.2499

Figure 82: results of the docking between cilengitide and the $\alpha\text{Ib}\beta\text{3}$ PDB entries a) 7U60 and b) 4Z7N, in the 7U60.M binding site.

6NAJ results is highly positive because the docking run was performed without deleting the Hr10FnIII present: the RGD ligand presents very low affinity because the site is already occupied by a high affinity form of 10FnIII (**Figure 78-3**).

To assess if the inhibitor cilengitide bound to integrin in a similar manner to native RGD compounds, the contacts between the compound and the receptor after the docking procedure were calculated and are listed in the figures below (**Figure 83**).



Type	ChainA	PosA	SetA	ChainB	PosB	SetB	Energy	Dist	BB	Freq	Cons	Area
IM	1-Cilengitide	1	*0	5-1L5G_1.8	1	MN4001	-23.29	2.52	--	4	1	n
A	1-Cilengitide	1	*0	4-1L5G_1.8	161	Arg214	-0.70	3.93	--	1	0	n
H	1-Cilengitide	1	*0	4-1L5G_1.8	198	Asp251	-2.90	3.11	--	1	0	n

Figure 83: Contacts between Cilengitide and PDB entry 1L5G ($\alpha\text{V}\beta\text{3}$)

Seven different compounds were built with the molecule builder tool in MOE. The new compounds are represented in the image below (**Figure 84**) and docking score with PDB entry 1L5G ($\alpha\text{V}\beta\text{3}$) are listed in **Figure 85**. Compounds labelled *cilengitide_mod7*, the seventh molecule built, is tested with docking to different PDB entries, to see if there are differences between different types and conformational states (**Figure 86**).

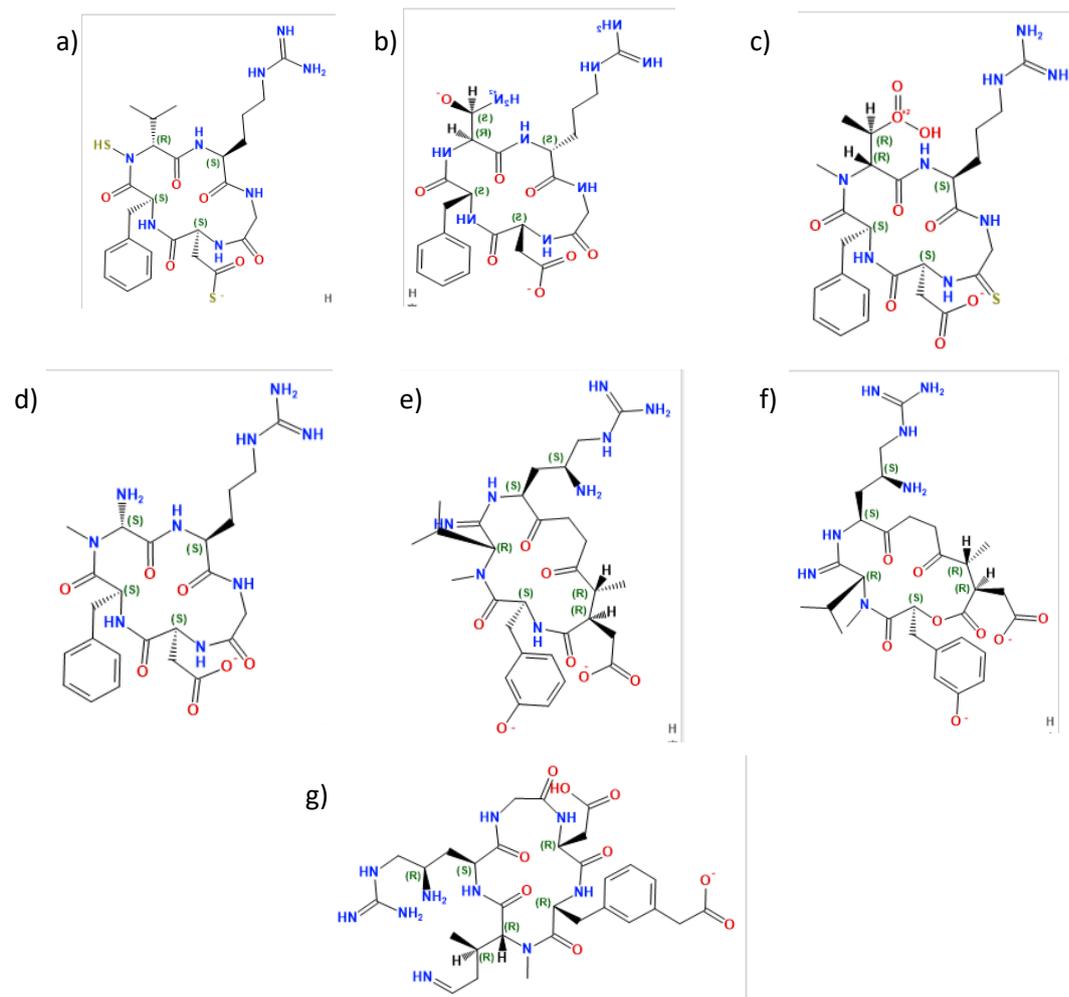


Figure 84: compounds built starting from the cilengitide formula. New molecules are labelled as cilengitide_mod_*, with * going from 1 to 7 (a-g).

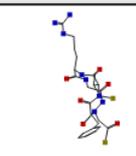
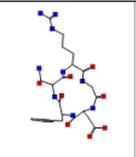
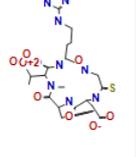
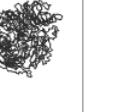
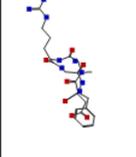
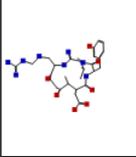
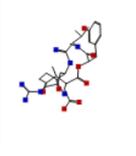
	mol	receptor	rseq	mseq	S	rmsd_refine	E_conf	E_place	E_score1	E_refine	E_score2
a)			1	1	-9.2364	2.9156	-242.5235	-44.5302	-10.6745	-71.3693	-9.2364
b)			1	1	-9.4616	1.8911	-516.4954	-66.8926	-11.7652	-73.0008	-9.4616
c)			1	1	-9.6451	3.4350	-358.7808	-57.5488	-13.3252	-73.8387	-9.6451
d)			1	1	-9.2678	1.7849	-255.2493	-87.4819	-12.4870	-73.8440	-9.2678
e)			1	1	-9.7493	2.4263	-397.8318	-78.7935	-13.8694	-71.5197	-9.7493
f)			1	1	-9.0150	2.6186	-395.0823	-86.6962	-12.3166	-64.8193	-9.0150

Figure 85: Docking scores of cilengitide_mod_(1-6) (corresponding to a-f).

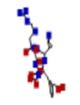
	mol	receptor	rseq	mseq	S	rmsd_refine	E_conf	E_place	E_score1	E_refine	E_score2
a)			1	1	-7.6419	2.3388	-252.9410	-74.0988	-8.3842	-38.4752	-7.6419
b)			1	1	-7.4228	1.9158	-273.3043	-43.9185	-8.2838	-48.9598	-7.4228
c)			1	1	-9.0690	2.0626	-259.5785	-77.7413	-14.6851	-65.0553	-9.0690
d)			1	1	-12.8927	2.0603	-267.5837	-110.0118	-12.6404	-121.0497	-12.8927

Figure 86: Docking result of cilengitide_mod_7 with different integrin structures in their RGD ligand site or in simulated RGD ligand site, like in refine docking; a) 1L5G; b) 3VI4; c) 7NWL; d) 7NXD

Furthermore, five additional compounds originating from the Cilengitide molecule were built (Figure 87):

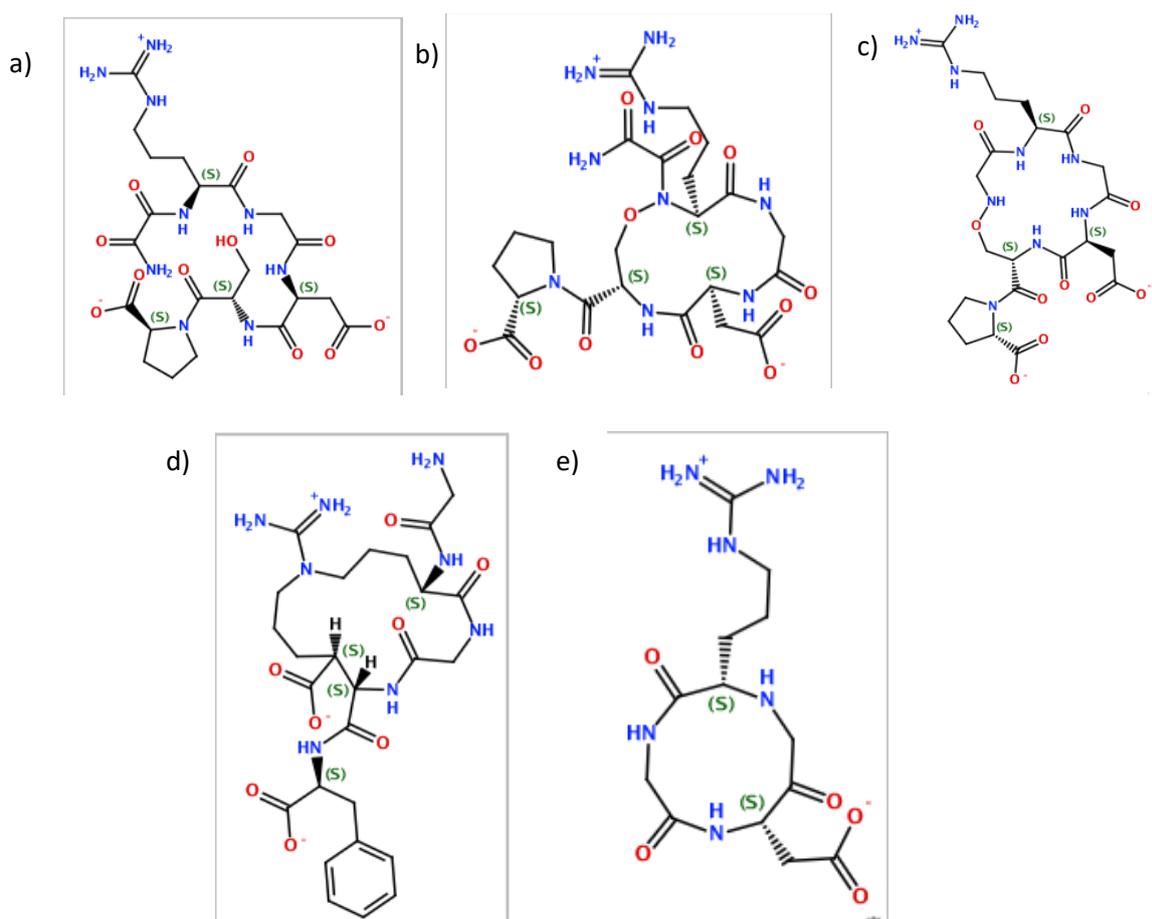


Figure 87: Chemical structures of *n* compounds. The first three (a-c) were obtained by cyclization 4WK2.C (compound_mod_a, compound_mod_b, compound_mod_c), the fourth (d) from cilengitide (compound_mod_d), the fifth (e) from cyclization of 3VI4.I (compound_mod_e).

These new compounds were tested on PDB entry 4G1M ($\alpha V\beta 3$), together with the previous built structures, docking the compounds in a simulated binding site using superposition with 1L5G. Results are listed in **Figure 88**.

	mol	receptor	rseq	mseq	S	rmsd_refine	E_conf	E_place	E_score1	E_refine	E_score2
1	Cilengitide (2)	4G1M	1	1	-6.5309	1.6556	-293.7867	-81.3212	-11.3159	-31.6333	-6.5309
2	Cilengitide (2)	4G1M	1	2	-6.1443	1.3900	-581.0028	-59.5469	-16.1064	-44.7267	-6.1443
3	Cilengitide (3)	4G1M	1	3	-6.4233	1.9438	-436.9060	-57.1871	-11.9768	-32.3184	-6.4233
4	Cilengitide (4)	4G1M	1	4	-6.0647	1.3674	-346.9453	-69.1853	-13.2787	-44.0310	-6.0647
5	Cilengitide (5)	4G1M	1	5	-6.1925	2.0273	-432.7704	-94.0489	-13.7173	-42.2189	-6.1925
6	Cilengitide (6)	4G1M	1	6	-7.0101	2.5965	-426.8966	-64.6985	-14.1673	-32.0753	-7.0101
7	OC(=O)CC1NC(=O)	4G1M	1	7	-6.5935	2.3134	-493.2240	-54.9155	-16.3221	-52.9942	-6.5935
8	4WK2_1.C	4G1M	1	8	-6.4722	1.9462	-341.4252	-54.4979	-18.6630	-67.4531	-6.4722
9	4WK2_1.C	4G1M	1	9	-6.0364	1.3101	-331.2284	-87.7134	-13.7376	-49.8797	-6.0364
10	4WK2_1.C	4G1M	1	10	-6.3045	2.2154	-369.4706	-75.5444	-18.5010	-44.0580	-6.3045
11		4G1M	1	11	-6.4940	1.7581	-327.8553	-89.5012	-15.2626	-58.7710	-6.4940
12	3VI4.I	4G1M	1	12	-4.8970	1.1850	-325.9924	-65.6738	-12.5147	-11.3907	-4.8970

Figure 88: docking results of the different compounds with 4G1M in 1L5G.C binding site: 1) cilengitide_mod_1 ; 2) cilengitide_mod_2; 3) cilengitide_mod_3; 4) cilengitide_mod_4; 5) cilengitide_mod_5; 6) cilengitide_mod_6) ; 7) cilengitide_mod_7 ; 8) cilengitide_mod_a ; 9) cilengitide_mod_b ; 10) cilengitide_mod_c ; 11) cilengitide_mod_d ; 12) cilengitide_mod_e.

By searching on PubChem, GooglePatents and WIPO, the new compounds are not been patented yet. Of course, the original Cilengitide, and similar RGD-based antagonist like Tirofiban and Eptifibatide are already patented.

ADMET analysis on the compound built from the cilengitide used are listed in the image below (**Figure 86a-b**)

AP_FWeight	ADMET_Risk	ADMET_Code	S+Acldic_pKa	S+Basic_pKa	DiffCoef	MlogP	S-logP	S-logD	logHLC	S+Peff
Formula we	Full ADMET f	Full ADMET Risk rule codes: Size, RotB=rotatable bonds, HBD=H-bond donors, HBA=H-bond accep	Predicted ma	Predicted m	Hayduk-L	Moriguct	Simulatic log D	at	Logarithm	Effective
622,769	11.5	Size; RotB; HBD; HBA; ch; Kow; Peff; Sw-; Vd-; rat; Xr+; Xm-; 2C9-; 2C19-; 2D6-; CL-	11.59; 4.58	17.12; 0.74	0.489	-1,249	-1,604	-1,603	-16,556	0,168
576,593	11	Size; RotB; HBD; HBA; ch; Kow; Peff; Sw-; Vd-; hERG-; Xr+; Xm-; 2C9-; 2C19-; 2D6-; CL-	4.53	16.91; 7.36	0,516	-2,937	-3,297	-3,234	-17,423	0,099
637,696	12	Size; RotB; HBD; HBA; ch; Kow; Peff; Sw-; Vd-; hERG-; Xr+; Xm-; MUT; 2C9-; 2C19-; 2D6-; CL-	9.90; 8.37	16.21; 0.31	0,486	-1,786	-3,917	-4,052	-16,761	0,115
561,601	11.5	Size; RotB; HBD; HBA; ch; Kow; Peff; Sw-; Vd-; hERG-; rat; Xr+; Xm-; 2C9-; 2C19-; 2D6-; CL-	4.49	16.92; 6.18	0,52	-2,245	-2,994	-2,973	-17,135	0,12
631,757	11	Size; RotB; HBD; HBA; ch; Kow; Peff; Sw-; Vd-; hERG-; Xr+; Xm-; 2C9-; 2C19-; 2D6-; CL-	10.45; 4.47	17.19; 8.68	0,469	-1,3	-1,46	-2,097	-17,027	0,13
688,745	11	Size; RotB; HBD; HBA; ch; Kow; Peff; Sw-; Vd-; hERG-; Xr+; Xm-; 2C9-; 2C19-; 2D6-; CL-	4.78; 4.01	17.06; 14.7	0,458	-3,252	-2,508	-2,7	-17,323	0,097
600,569	11.5	Size; RotB; HBD; HBA; ch; Kow; Peff; Sw-; Vd-; hERG-; Xr+; Xm-; HEPX-; 2C9-; 2C19-; 2D6-; CL-	11.18; 10.45	17.12; 0.37	0,516	-4,423	-3,561	-3,744	-17,378	0,098
598,553	11.5	Size; RotB; HBD; HBA; ch; Kow; Peff; Sw-; Vd-; hERG-; Xr+; Xm-; HEPX-; 2C9-; 2C19-; 2D6-; CL-	10.90; 4.84	16.81; 0.31	0,523	-3,611	-3,502	-3,938	-17,154	0,123
584,569	12	Size; RotB; HBD; HBA; ch; Kow; Peff; Sw-; Vd-; hERG-; Xr+; Xm-; MUT; 2C9-; 2C19-; 2D6-; CL-	3.95; 3.36	16.91; 5.10	0,523	-3,773	-3,698	-3,896	-17,158	0,09
589,632	11	Size; RotB; HBD; HBA; ch; Kow; Peff; Sw-; Vd-; hERG-; Xr+; Xm-; 2C9-; 2C19-; 2D6-; CL-	4.59; 3.59	16.54; 8.39	0,502	-2,239	-2,841	-2,873	-17,404	0,129
342,357	10	HBD; HBA; ch; Kow; Peff; Sw-; Vd-; hERG-; Xr; MUT; 2C9-; 2C19-; 2D6-; CL-	4.15	16.97; 6.20	0,697	-2,528	-3,837	-3,825	-15,731	0,14
588,668	9,059	Size; RotB; HBD; HBA; ch; Peff; hERG-; rat; Xr+; Xm-; 3A4; CL-	4.60	12.38; -0.02	0,496	-1,342	-2,056	-2,056	-16,358	0,161

Risks	*PCB*	*TRNS Substr*	*TRNS Inh*	*CYP Substr*	*CYP Clint*	*UGT Substr*	*Ames Mut*	*Tox*
ADMET_Risk [9.059,12.0]	MWT [342.357,688.745]	Pgp_Substr [1.000,1.000]	Pgp_Inh [0.000,0.000]	B_CYP1A2_Substr [0.000,0.000]	log(CYP1A2_Clint) [0.00,0.00]	UGT1A1 [0.000,0.000]	UMUT_97+1537 [0.000,0.000]	hERG_pIC50 [3.645,4.263]

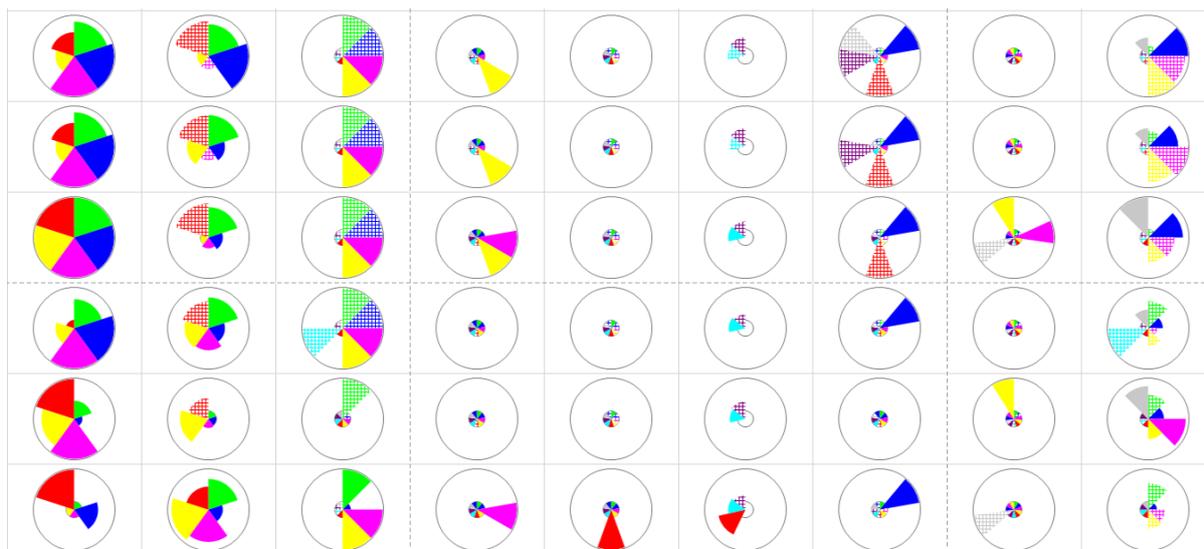


Figure 87: ADMET result of the 11 compounds built. From top to bottom are listed ADMET analysis of cilengitide_mod_1, cilengitide_mod_2, cilengitide_mod_3, cilengitide_mod_4, cilengitide_mod_5, cilengitide_mod_6, cilengitide_mod_7 and compound_mod_a, compound_mod_b, compound_mod_c, compound_mod_d, compound_mod_e.

The comparison between the docking of RGD ligand and cilengitide on PDB entry 1L5G infers that the affinity for the receptor is of similar strength and cilengitide could inhibit binding of other RGD-based interactions. This is also confirmed by literature, that also points out how probably allosteric inhibitors have better affinity and could better inhibit binding between the integrin receptors and RGD ligands.

From docking results on different conformational states of integrin, it is confirmed that allosteric inhibition between monoclonal antibodies and RGD-based peptide or protein could serve as promising therapeutic method. RGD-based peptide affinity for integrin is pretty low compared to the monoclonal antibodies, and there are still reserve on whether these compounds (e.g cilengitide) are capable of disrupting the integrin-ligand complexes after their binding.

Allosteric integrin antagonists that simulate the effect of Ca^{2+} in promoting the closed form of the headpiece have the potential in not only preventing the formation of integrin-ligand complexes, but could also disrupt pre-existing complexes and dissociate pre-formed adhesions.

Also, the oral bioavailability of RGD-based antagonists has been restricted by the essential requirement to include a carboxylic acid moiety, while allosteric inhibitors don't have such limitations. Allosteric inhibitors are also much less prone to risk of thrombosis and thrombocytopenia.

mAb LM609 is a promising candidate in designing a allosteric antagonist for the $\alpha V\beta 3$ type integrins and presents much more specificity against these integrin types that RGD-based antagonists, that target in practice all the integrins recognizing the RGD motif. From superposition of PDB entries 1FNF and 4MMX we could also observe that LM609 doesn't directly occlude the RGD-binding pocket (it doesn't directly bind in the RGD-recognizing motif) and small RGD-peptide could still bind with low affinity to the pocket, but probably prevent binding of the bigger physiological ligands, like fibrinogen or fibronectin, through steric hindrance. In fact, Fibronectin also needs sites in 8FnIII and 9FnIII to achieve an efficient binding to $\alpha V\beta 3$ integrins, and the presence of mAb LM609 precludes binding due to expected clashes with fibronectin domain 8FnIII and 9FnIII.

Regarding type $\alpha 5\beta 1$ integrins, mAb13 presents effects much like LM609 for $\alpha V\beta 3$. Similarly, binding of mAb13 to $\alpha 5\beta 1$ integrins is inversely proportional to the quantity of ligands binding, suggesting the presence of an allosteric inhibition, likely caused either by conformational changes in the structure of

integrins that cause the expulsion of ligands from their binding pocket or by conformational changes that push integrins in their inactive state, incapable of ligand recognition, but the latter hypothesis is contested by the fact that in experiments in which integrin was preincubated with mAb13, antibody binding could be reversed by the CCBD fragment or GRGDS peptide, demonstrating that antibody-occupied integrin was still capable of binding ligand.

These data are validated by the fact that crystallographic structures presenting monoclonal antibodies that should hinder binding of ligand, although not completely forbidding it, presents S score of docking simulation with both RGD based peptide and cilengitide way lower than the crystallographic structure that don't present these antibodies and are in the open and active form of the integrin.

Binding of bigger RGD ligand to monoclonal bound integrins presents S score gradually decreasing, confirming the fact that small RGD peptides could still bind with low affinity, but bigger RGD domain (such as the one present in fibronectin, that span through 3 domains) have much more difficulty to bind, probably due to allosteric inhibition and/or steric hindrance.

Future directions of this work should be primarily to create a pharmacophore query for the RGD interactions and for known inhibitor that bind to the RGD binding site, to improve and simplify the research and computation. The Lipinski rule of 5 was not used in this work to filter the new compounds, because the focus was on the binding affinity and ADMET analysis. Focus on finding novel therapeutic RGD-based that present higher specificity and less adverse effects is necessary. RGD compounds tend to be more difficult to deliver orally and tend to be generally more toxic. This, united with lower specificity and high chance of collateral effects, make ADMET and in vitro analysis essential to complete computational studies. Checking if these compounds are synthesizable (e.g. through RECXYS software) is a necessary step following ADMET analysis. Use in synchrony of mAbs, allosteric inhibitor and RGD-based peptides could be a promising combination to achieve high specificity and high efficacy.

4. Bibliography

- Adair, B. D., Alonso, J. L., van Agthoven, J., Hayes, V., Ahn, H. S., Yu, I. S., Lin, S. W., Xiong, J. P., Poncz, M., & Arnaout, M. A. (2020). Structure-guided design of pure orthosteric inhibitors of α IIb β 3 that prevent thrombosis but preserve hemostasis. *Nature Communications*, *11*(1). <https://doi.org/10.1038/s41467-019-13928-2>
- Aota -i., S., Nagai, T., & Yamada, K. M. (1991). Characterization of regions of fibronectin besides the arginine- glycine-aspartic acid sequence required for adhesive function of the cell-binding domain using site-directed mutagenesis. *Journal of Biological Chemistry*, *266*(24), 15938–15943. [https://doi.org/10.1016/s0021-9258\(18\)98498-x](https://doi.org/10.1016/s0021-9258(18)98498-x)
- Bachman, H., Nicosia, J., Dysart, M., & Barker, T. H. (2015). Utilizing Fibronectin Integrin-Binding Specificity to Control Cellular Responses. *Advances in Wound Care*, *4*(8), 501–511. <https://doi.org/10.1089/wound.2014.0621>
- Bae, Y. K., Kim, A., Kim, M. K., Choi, J. E., Kang, S. H., & Lee, S. J. (2013). Fibronectin expression in carcinoma cells correlates with tumor aggressiveness and poor clinical outcome in patients with invasive breast cancer. *Human Pathology*, *44*(10), 2028–2037. <https://doi.org/10.1016/J.HUMPATH.2013.03.006>
- Baldwin, J., & Chothia, C. (1979). Haemoglobin: The structural changes related to ligand binding and its allosteric mechanism. *Journal of Molecular Biology*, *129*(2), 175–220. [https://doi.org/10.1016/0022-2836\(79\)90277-8](https://doi.org/10.1016/0022-2836(79)90277-8)

- Barton, S. J., Travis, M. A., Askari, J. A., Buckley, P. A., Craig, S. E., Humphries, M. J., & Mould, A. P. (2004). Novel activating and inactivating mutations in the integrin $\beta 1$ subunit A domain. *Biochemical Journal*, *380*(2), 401–407. <https://doi.org/10.1042/BJ20031973>
- BENESCH, R., & BENESCH, R. E. (1969). Intracellular Organic Phosphates as Regulators of Oxygen Release by Haemoglobin. *Nature*, *221*(5181), 618–622. <https://doi.org/10.1038/221618a0>
- Bingham, R. J., Rudiño-Piñera, E., Meenan, N. A. G., Schwarz-Linek, U., Turkenburg, J. P., Höök, M., Garman, E. F., & Potts, J. R. (2008). Crystal structures of fibronectin-binding sites from *Staphylococcus aureus* FnBPA in complex with fibronectin domains. *Proceedings of the National Academy of Sciences*, *105*(34), 12254–12258. <https://doi.org/10.1073/PNAS.0803556105>
- Borst, A. J., James, Z. M., Zagotta, W. N., Ginsberg, M., Rey, F. A., DiMaio, F., Backovic, M., & Veesler, D. (2017). The Therapeutic Antibody LM609 Selectively Inhibits Ligand Binding to Human $\alpha V\beta 3$ Integrin via Steric Hindrance. *Structure*, *25*(11), 1732–1739.e5. <https://doi.org/10.1016/j.str.2017.09.007>
- Bowditch, R. D., Hariharan, M., Tominna, E. F., Smith, J. W., Yamada, K. M., Getzoff, E. D., & Ginsberg, M. H. (1994). Identification of a novel integrin binding site in fibronectin. Differential utilization by $\beta 3$ integrins. *Journal of Biological Chemistry*, *269*(14), 10856–10863. [https://doi.org/10.1016/s0021-9258\(17\)34137-6](https://doi.org/10.1016/s0021-9258(17)34137-6)
- Catterall, W. A., Rev, A., Dev, C., Xiong, J., Stehle, T., Diefenbach, B., Zhang, R., Dunker, R., & Arnaout, M. A. (2001). Crystal Structure of the Extracellular Segment of Integrin $\alpha V\beta 3$. *October*, *294*(October), 339–345.
- Chen, J. F., Salas, A., & Springer, T. A. (2003). Bistable regulation of integrin adhesiveness by a bipolar metal ion cluster. *Nature Structural Biology*, *10*(12), 995–1001. <https://doi.org/10.1038/nsb1011>
- Craig, D., Gao, M., Schulten, K., & Vogel, V. (2004). Structural insights into how the MIDAS ion stabilizes integrin binding to an RGD peptide under force. *Structure*, *12*(11), 2049–2058. <https://doi.org/10.1016/j.str.2004.09.009>
- Dalton, C. J., & Lemmon, C. A. (2021). Fibronectin: Molecular structure, fibrillar structure and mechanochemical signaling. In *Cells* (Vol. 10, Issue 9, p. 2443). Multidisciplinary Digital Publishing Institute. <https://doi.org/10.3390/cells10092443>
- de Vos, A. M. (1993). Crystal structure of the complex between human growth hormone and the extracellular domain of its reception. *Biomedicine and Pharmacotherapy*, *47*(1), 46–47. [https://doi.org/10.1016/0753-3322\(93\)90042-J](https://doi.org/10.1016/0753-3322(93)90042-J)
- Dong, X., Mi, L. Z., Zhu, J., Wang, W., Hu, P., Luo, B. H., & Springer, T. A. (2012). $\alpha V\beta 3$ Integrin Crystal Structures and Their Functional Implications. *Biochemistry*, *51*(44), 8814–8828. <https://doi.org/10.1021/bi300734n>
- Dransfield, I., Cabañas, C., Craig, A., & Hogg, N. (1992). Divalent cation regulation of the function of the leukocyte integrin LFA-1. *Journal of Cell Biology*, *116*(1), 219–226. <https://doi.org/10.1083/jcb.116.1.219>
- Faull, R. J., Kovach, N. L., Harlan, J. M., & Ginsberg, M. H. (1993). Affinity modulation of integrin $\alpha 5\beta 1$: Regulation of the functional response by soluble fibronectin. *Journal of Cell Biology*, *121*(1), 155–162. <https://doi.org/10.1083/jcb.121.1.155>
- Ferrari, E., Menabue, L., & Saladini, M. (2004). Characterization and metal affinity of Tirofiban, a pharmaceutical compound used in acute coronary syndromes. *BioMetals*, *17*(2), 145–155.

<https://doi.org/10.1023/B:BIOM.0000018376.52169.40>

- Frelinger, A. L., Cohen, I., Plow, E. F., Smith, M. A., Roberts, J., Lam, S. C. T., & Ginsberg, M. H. (1990). Selective inhibition of integrin function by antibodies specific for ligand-occupied receptor conformers. *Journal of Biological Chemistry*, *265*(11), 6346–6352. [https://doi.org/10.1016/s0021-9258\(19\)39332-9](https://doi.org/10.1016/s0021-9258(19)39332-9)
- Frelinger, A. L., Du, X., Plow, E. F., & Ginsberg, M. H. (1991). Monoclonal antibodies to ligand-occupied conformers of integrin α (IIb) β 3 (glycoprotein IIb-IIIa) alter receptor affinity, specificity, and function. *Journal of Biological Chemistry*, *266*(26), 17106–17111. [https://doi.org/10.1016/s0021-9258\(19\)47346-8](https://doi.org/10.1016/s0021-9258(19)47346-8)
- Frelinger, A. L., Lam, S. C. T., Plow, E. F., Smith, M. A., Loftus, J. C., & Ginsberg, M. H. (1988). Occupancy of an adhesive glycoprotein receptor modulates expression of an antigenic site involved in cell adhesion. *Journal of Biological Chemistry*, *263*(25), 12397–12402. [https://doi.org/10.1016/s0021-9258\(18\)37769-x](https://doi.org/10.1016/s0021-9258(18)37769-x)
- Gailit, J., & Ruoslahti, E. (1988). Regulation of the fibronectin receptor affinity by divalent cations. *Journal of Biological Chemistry*, *263*(26), 12927–12932. [https://doi.org/10.1016/s0021-9258\(18\)37650-6](https://doi.org/10.1016/s0021-9258(18)37650-6)
- Garrigues, H. J., Rubinchikova, Y. E., DiPersio, C. M., & Rose, T. M. (2008). Integrin α V β 3 Binds to the RGD Motif of Glycoprotein B of Kaposi's Sarcoma-Associated Herpesvirus and Functions as an RGD-Dependent Entry Receptor. *Journal of Virology*, *82*(3), 1570–1580. <https://doi.org/10.1128/jvi.01673-07>
- Ginsberg, M. H., Partridge, A., & Shattil, S. J. (2005). Integrin regulation. *Current Opinion in Cell Biology*, *17*(5 SPEC. ISS.), 509–516. <https://doi.org/10.1016/j.ceb.2005.08.010>
- Griggs, L. A., Hassan, N. T., Malik, R. S., Griffin, B. P., Martinez, B. A., Elmore, L. W., & Lemmon, C. A. (2017). Fibronectin fibrils regulate TGF- β 1-induced Epithelial-Mesenchymal Transition. *Matrix Biology*, *60–61*, 157–175. <https://doi.org/10.1016/J.MATBIO.2017.01.001>
- Harding, M. M. (2001). *D57, 401±411 Harding Geometry of metal±ligand interactions*. 401–411. <http://www.ebi.ac.uk/pdb/>
- Henzler-Wildman, K., & Kern, D. (2007). Dynamic personalities of proteins. *Nature*, *450*(7172), 964–972. <https://doi.org/10.1038/nature06522>
- Hynes, R. O. (1992). Integrins: Versatility, modulation, and signaling in cell adhesion. *Cell*, *69*(1), 11–25. [https://doi.org/10.1016/0092-8674\(92\)90115-5](https://doi.org/10.1016/0092-8674(92)90115-5)
- Hynes, R. O., & Yamada, K. M. (1982). Fibronectins: Multifunctional modular glycoproteins. *Journal of Cell Biology*, *95*(2), 369–377. <https://doi.org/10.1083/jcb.95.2.369>
- Jieqing Zhu Bing-Hao Luo, Patrick Barth, Jack Schonbrun, David Baker, and T. A. S. (2009). *The structure of a receptor with two associating transmembrane domains on the cell surface: integrin α IIb β 3*. <https://doi.org/10.1016/j.molcel.2009.02.022>.The
- Johansson, S., Svineng, G., Wennerberg, K., Armulik, A., & Lohikangas, L. (1997). FIBRONECTIN-INTEGRIN INTERACTIONS. In *Frontiers in Bioscience* (Vol. 2).
- Kamata, T., Handa, M., Ito, S., Sato, Y., Ohtani, T., Kawai, Y., Ikeda, Y., & Aiso, S. (2010). Structural requirements for activation in α IIb β 3 integrin. *Journal of Biological Chemistry*, *285*(49), 38428–38437. <https://doi.org/10.1074/jbc.M110.139667>
- Kamata, T., Tieu, K. K., Tarui, T., Puzon-McLaughlin, W., Hogg, N., & Takada, Y. (2002). The Role of the CPNKEKEC Sequence in the β 2 Subunit I Domain in Regulation of Integrin α L β 2 (LFA-1). *The*

Journal of Immunology, 168(5), 2296–2301. <https://doi.org/10.4049/jimmunol.168.5.2296>

- Kornblihtt, A. R., Umezawa, K., Vibe-Pedersen, K., & Baralle, F. E. (1985). Primary structure of human fibronectin: differential splicing may generate at least 10 polypeptides from a single gene. *The EMBO Journal*, 4(7), 1755–1759. <https://doi.org/10.1002/j.1460-2075.1985.tb03847.x>
- Kovach, N. L., Carlos, T. M., Yee, E., & Harlan, J. M. (1992). A monoclonal antibody to $\beta 1$ integrin (CD29) stimulates VLA-dependent adherence of leukocytes to human umbilical vein endothelial cells and matrix components. *Journal of Cell Biology*, 116(2), 499–509. <https://doi.org/10.1083/jcb.116.2.499>
- Leahy, D. J., Aukhil, I., & Erickson, H. P. (1996). 2.0 Å Crystal Structure of a Four-Domain Segment of Human Fibronectin Encompassing the RGD Loop and Synergy Region. *Cell*, 84(1), 155–164. [https://doi.org/10.1016/S0092-8674\(00\)81002-8](https://doi.org/10.1016/S0092-8674(00)81002-8)
- Lin, F. Y., Zhu, J., Eng, E. T., Hudson, N. E., & Springer, T. A. (2016). β -Subunit binding is sufficient for ligands to open the integrin α IIb β 3 headpiece. *Journal of Biological Chemistry*, 291(9), 4537–4546. <https://doi.org/10.1074/jbc.M115.705624>
- Luo, B., Carman, C. V., & Springer, T. A. (2007). Structural Basis of Integrin Regulation and Signaling. *Annu Rev Immunol*, 619–647.
- Luo, B. H., Carman, C. V., & Springer, T. A. (2007). Structural basis of integrin regulation and signaling. *Annual Review of Immunology*, 25, 619–647. <https://doi.org/10.1146/annurev.immunol.25.022106.141618>
- Luo, B. H., Karanicolas, J., Harmacek, L. D., Baker, D., & Springer, T. A. (2009). Rationally designed integrin $\beta 3$ mutants stabilized in the high affinity conformation. *Journal of Biological Chemistry*, 284(6), 3917–3924. <https://doi.org/10.1074/jbc.M806312200>
- Luo, B. H., & Springer, T. A. (2006). Integrin structures and conformational signaling. *Current Opinion in Cell Biology*, 18(5), 579–586. <https://doi.org/10.1016/j.ceb.2006.08.005>
- Luo, B. H., Springer, T. A., & Takagi, J. (2003). Stabilizing the open conformation of the integrin headpiece with a glycan wedge increases affinity for ligand. *Proceedings of the National Academy of Sciences of the United States of America*, 100(5), 2403–2408. <https://doi.org/10.1073/pnas.0438060100>
- Luo, B. H., Springer, T. A., & Takagi, J. (2004). A specific interface between integrin transmembrane helices and affinity for ligand. *PLoS Biology*, 2(6). <https://doi.org/10.1371/journal.pbio.0020153>
- Luo, B. H., Takagi, J., & Springer, T. A. (2004). Locking the $\beta 3$ Integrin I-like Domain into High and Low Affinity Conformations with Disulfides. *Journal of Biological Chemistry*, 279(11), 10215–10221. <https://doi.org/10.1074/jbc.M312732200>
- Luque, A., Sánchez-Madrid, F., & Cabañas, C. (1994). Functional regulation of the human integrin VLA-1 (CD49a/CD29) by divalent cations and stimulatory $\beta 1$ antibodies. *FEBS Letters*, 346(2–3), 278–284. [https://doi.org/10.1016/0014-5793\(94\)00490-0](https://doi.org/10.1016/0014-5793(94)00490-0)
- Mahalingam, B., Van Agthoven, J. F., Xiong, J. P., Alonso, J. L., Adair, B. D., Rui, X., Anand, S., Mehrbod, M., Mofrad, M. R. K., Burger, C., Goodman, S. L., & Arnaout, M. A. (2014). Atomic basis for the species-specific inhibition of αV integrins by monoclonal antibody 17E6 is revealed by the crystal structure of $\alpha V\beta 3$ ectodomain-17E6 Fab complex. *Journal of Biological Chemistry*, 289(20), 13801–13809. <https://doi.org/10.1074/jbc.M113.546929>
- Main, A. L., Harvey, T. S., Baron, M., Boyd, J., & Campbell, I. D. (1992). The three-dimensional structure of the tenth type III module of fibronectin: An insight into RGD-mediated interactions.

Cell, 71(4), 671–678. [https://doi.org/10.1016/0092-8674\(92\)90600-H](https://doi.org/10.1016/0092-8674(92)90600-H)

- Mezzenga, R., & Mitsi, M. (2019). The Molecular Dance of Fibronectin: Conformational Flexibility Leads to Functional Versatility. In *Biomacromolecules* (Vol. 20, Issue 1, pp. 55–72). American Chemical Society. <https://doi.org/10.1021/acs.biomac.8b01258>
- Moore, H. D. (2001). Molecular biology of fibronectin. In *Journal of reproduction and fertility. Supplement* (Vol. 57, pp. 105–110).
- Mosesson, M. W. (2005). Structure and Functions of Fibrinogen and Fibrin. *Recent Advances in Thrombosis and Hemostasis 2008*, 3–26. https://doi.org/10.1007/978-4-431-78847-8_1
- Mosesson, M. W., Chen, A. B., & Huseby, R. M. (1975). The cold-insoluble globulin of human plasma: Studies of its essential structural features. *BBA - Protein Structure*, 386(2), 509–524. [https://doi.org/10.1016/0005-2795\(75\)90294-9](https://doi.org/10.1016/0005-2795(75)90294-9)
- Mosher, D. F. (1984). Physiology of fibronectin. *Annual Review of Medicine*, 35, 561–575. <https://doi.org/10.1146/annurev.me.35.020184.003021>
- Mould, A. P., Askari, J. A., Barton, S., Kline, A. D., McEwan, P. A., Craig, S. E., & Humphries, M. J. (2002). Integrin activation involves a conformational change in the $\alpha 1$ helix of the β subunit A-domain. *Journal of Biological Chemistry*, 277(22), 19800–19805. <https://doi.org/10.1074/jbc.M201571200>
- Mould, A. P., Barton, S. J., Askari, J. A., Craig, S. E., & Humphries, M. J. (2003). Role of ADMIDAS Cation-binding Site in Ligand Recognition by Integrin $\alpha 5\beta 1$. *Journal of Biological Chemistry*, 278(51), 51622–51629. <https://doi.org/10.1074/jbc.M306655200>
- Mould, A. P., Barton, S. J., Askari, J. A., McEwan, P. A., Buckley, P. A., Craig, S. E., & Humphries, M. J. (2003). Conformational changes in the integrin βA domain provide a mechanism for signal transduction via hybrid domain movement. *Journal of Biological Chemistry*, 278(19), 17028–17035. <https://doi.org/10.1074/jbc.M213139200>
- Mould, A. P., Craig, S. E., Byron, S. K., Humphries, M. J., & Jowitt, T. A. (2014). Disruption of integrin-fibronectin complexes by allosteric but not ligand-mimetic inhibitors. *Biochemical Journal*, 464(3), 301–313. <https://doi.org/10.1042/BJ20141047>
- Mould, A. P., Garratt, A. N., Askari, J. A., Akiyama, S. K., & Humphries, M. J. (1995). Identification of a novel anti-integrin monoclonal antibody that recognises a ligand-induced binding site epitope on the $\beta 1$ subunit. *FEBS Letters*, 363(1–2), 118–122. [https://doi.org/10.1016/0014-5793\(95\)00301-0](https://doi.org/10.1016/0014-5793(95)00301-0)
- Muller, B., Zerwes, H. G., Tangemann, K., Peter, J., & Engel, J. (1993). Two-step binding mechanism of fibrinogen to $\alpha IIb\beta 3$ integrin reconstituted into planar lipid bilayers. *Journal of Biological Chemistry*, 268(9), 6800–6808. [https://doi.org/10.1016/S0021-9258\(18\)53320-2](https://doi.org/10.1016/S0021-9258(18)53320-2)
- Nagae, M., Re, S., Mihara, E., Nogi, T., Sugita, Y., & Takagi, J. (2012). Crystal structure of $\alpha 5\beta 1$ integrin ectodomain: Atomic details of the fibronectin receptor. *Journal of Cell Biology*, 197(1), 131–140. <https://doi.org/10.1083/jcb.201111077>
- Obara, M., Kang, M. S., & Yamada, K. M. (1988). Site-directed mutagenesis of the cell-binding domain of human fibronectin: Separable, synergistic sites mediate adhesive function. *Cell*, 53(4), 649–657. [https://doi.org/10.1016/0092-8674\(88\)90580-6](https://doi.org/10.1016/0092-8674(88)90580-6)
- Orlando, R. A., & Cheresh, D. A. (1991). Arginine-glycine-aspartic acid binding leading to molecular stabilization between integrin $\alpha(v)\beta 3$ and its ligand. *Journal of Biological Chemistry*, 266(29), 19543–19550. [https://doi.org/10.1016/s0021-9258\(18\)55029-8](https://doi.org/10.1016/s0021-9258(18)55029-8)

- Pankov, R., & Yamada, K. M. (2002). Fibronectin at a glance. *Journal of Cell Science*, *115*(20), 3861–3863. <https://doi.org/10.1242/jcs.00059>
- Parise, L. V., Helgerson, S. L., Steiner, B., Nannizzi, L., & Phillips, D. R. (1987). Synthetic peptides derived from fibrinogen and fibronectin change the conformation of purified platelet glycoprotein IIb-IIIa. *Journal of Biological Chemistry*, *262*(26), 12597–12602. [https://doi.org/10.1016/s0021-9258\(18\)45247-7](https://doi.org/10.1016/s0021-9258(18)45247-7)
- Paul Mould, A., Akiyama, S. K., & Humphries, M. J. (1996). The inhibitory anti- β 1 integrin monoclonal antibody 13 recognizes an epitope that is attenuated by ligand occupancy. Evidence for allosteric inhibition of integrin function. *Journal of Biological Chemistry*, *271*(34), 20365–20374. <https://doi.org/10.1074/jbc.271.34.20365>
- Petersen, T. E., Thøgersen, H. C., Skorstengaard, K., Vibe-Pedersen, K., Sahl, P., Sottrup-Jensen, L., & Magnusson, S. (1983). Partial primary structure of bovine plasma fibronectin: Three types of internal homology. *Proceedings of the National Academy of Sciences*, *80*(1), 137–141. <https://doi.org/10.1073/PNAS.80.1.137>
- Pickford, A. R., Smith, S. P., Staunton, D., & Boyd, J. (2001). The hairpin structure of the 6F11F22F2 fragment from human fibronectin enhances gelatin binding. *EMBO Journal*, *20*(7), 1519–1529. <https://doi.org/10.1093/emboj/20.7.1519>
- Plow, E. F., Pierschbacher, M. D., Ruoslahti, E., Marguerie, G. A., & Ginsberg, M. H. (1985). The effect of Arg-Gly-Asp-containing peptides on fibrinogen and von Willebrand factor binding to platelets. *Proceedings of the National Academy of Sciences of the United States of America*, *82*(23), 8057–8061. <https://doi.org/10.1073/pnas.82.23.8057>
- Potts, J. R., Bright, J. R., Bolton, D., Pickford, A. R., & Campbell, I. D. (1999). Solution structure of the N-terminal F1 module pair from human fibronectin. *Biochemistry*, *38*(26), 8304–8312. <https://doi.org/10.1021/bi990202b>
- Potts, J. R., & Campbell, I. D. (1994). Fibronectin structure and assembly. *Current Opinion in Cell Biology*, *6*(5), 648–655. [https://doi.org/10.1016/0955-0674\(94\)90090-6](https://doi.org/10.1016/0955-0674(94)90090-6)
- Puklin-Faucher, E., Gao, M., Schulten, K., & Vogel, V. (2006). How the headpiece hinge angle is opened: New insights into the dynamics of integrin activation. *Journal of Cell Biology*, *175*(2), 349–360. <https://doi.org/10.1083/jcb.200602071>
- Puklin-Faucher, E., & Vogel, V. (2009). Integrin activation dynamics between the RGD-binding site and the headpiece hinge. *Journal of Biological Chemistry*, *284*(52), 36557–36568. <https://doi.org/10.1074/jbc.M109.041194>
- Reardon, D. A., & Cheresch, D. (2011). Cilengitide: A Prototypic Integrin Inhibitor for the Treatment of Glioblastoma and Other Malignancies. *Genes and Cancer*, *2*(12), 1159–1165. <https://doi.org/10.1177/1947601912450586>
- Richardson, J. S., Getzoff, E. D., & Richardson, D. C. (1978). The β bulge: A common small unit of nonrepetitive protein structure. *Proceedings of the National Academy of Sciences of the United States of America*, *75*(6), 2574–2578. <https://doi.org/10.1073/pnas.75.6.2574>
- Roca-Cusachs, P., Gauthier, N. C., Del Rio, A., & Sheetz, M. P. (2009). Clustering of α 5 β 1 integrins determines adhesion strength whereas α v β 3 and talin enable mechanotransduction. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(38), 16245–16250. <https://doi.org/10.1073/pnas.0902818106>
- Rudiño-Piñera, E., Ravelli, R. B. G., Sheldrick, G. M., Nanao, M. H., Korostelev, V. V., Werner, J. M., Schwarz-Linek, U., Potts, J. R., & Garman, E. F. (2007). The Solution and Crystal Structures of a

Module Pair from the Staphylococcus aureus-Binding Site of Human Fibronectin-A Tale with a Twist. *Journal of Molecular Biology*, 368(3), 833–844.
<https://doi.org/10.1016/j.jmb.2007.02.061>

- Sachs, D. H., Schechter, A. N., Eastlake, A., & Anfinsen, C. B. (1972). An immunologic approach to the conformational equilibria of polypeptides. *Proceedings of the National Academy of Sciences of the United States of America*, 69(12), 3790–3794. <https://doi.org/10.1073/pnas.69.12.3790>
- Scarborough, R. M., & Gretler, D. D. (2000). *Novel Parenteral and Potential Oral Antithrombotic Agents*. 43(19).
- Schiefner, A., Gebauer, M., Richter, A., & Skerra, A. (2018). Anticalins Reveal High Plasticity in the Mode of Complex Formation with a Common Tumor Antigen. *Structure*, 26(4), 649–656.e3.
<https://doi.org/10.1016/j.str.2018.02.003>
- Schiefner, A., Gebauer, M., & Skerra, A. (2012). Extra-domain B in oncofetal fibronectin structurally promotes fibrillar head-to-tail dimerization of extracellular matrix protein. *Journal of Biological Chemistry*, 287(21), 17578–17588. <https://doi.org/10.1074/jbc.M111.303131>
- Schumacher, S., Dedden, D., Nunez, R. V., Matoba, K., Takagi, J., Biertümpfel, C., & Mizuno, N. (2021). Structural insights into integrin $\alpha 5\beta 1$ opening by fibronectin ligand. *Science Advances*, 7(19), 1–18. <https://doi.org/10.1126/sciadv.abe9716>
- Sen, M., Yuki, K., & Springer, T. A. (2013). An internal ligand-bound, metastable state of a leukocyte integrin, $\alpha X\beta 2$. *Journal of Cell Biology*, 203(4), 629–642. <https://doi.org/10.1083/jcb.201308083>
- Sharma, A. (1999). Crystal structure of a heparin-and integrin-binding segment of human fibronectin. *The EMBO Journal*, 18(6), 1468–1479. <https://doi.org/10.1093/emboj/18.6.1468>
- Sheldrake, H. M., & Patterson, L. H. (2014). Strategies to inhibit tumor associated integrin receptors: Rationale for dual and multi-antagonists. *Journal of Medicinal Chemistry*, 57(15), 6301–6315.
<https://doi.org/10.1021/jm5000547>
- Shi, M. L., Shen, Y. F., Tan, S. M., Mitchell, E. P., Law, S. K. A., & Lescar, J. (2007). A structural hypothesis for the transition between bent and extended conformations of the leukocyte $\beta 2$ integrins. *Journal of Biological Chemistry*, 282(41), 30198–30206.
<https://doi.org/10.1074/jbc.M701670200>
- Shi, M. L., Sundramurthy, K., Liu, B., Tan, S. M., Law, S. K. A., & Lescar, J. (2005). The crystal structure of the plexin-semaphorin-integrin domain/hybrid domain/I-EGF1 segment from the human integrin $\beta 2$ subunit at 1.8-Å resolution. *Journal of Biological Chemistry*, 280(34), 30586–30593.
<https://doi.org/10.1074/jbc.M502525200>
- Shi, M., Zhu, J., Wang, R., Chen, X., Mi, L., Walz, T., & Springer, T. A. (2011). Latent TGF- β structure and activation. *Nature*, 474(7351), 343–351. <https://doi.org/10.1038/nature10152>
- Springer, T. A. (2009). Structural basis for selectin mechanochemistry. *Proceedings of the National Academy of Sciences of the United States of America*, 106(1), 91–96.
<https://doi.org/10.1073/pnas.0810784105>
- Springer, T. A., & Dustin, M. L. (2012). Integrin inside-out signaling and the immunological synapse. *Current Opinion in Cell Biology*, 24(1), 107–115. <https://doi.org/10.1016/j.ceb.2011.10.004>
- Springer, T. A., Zhu, J., & Xiao, T. (2008). Structural basis for distinctive recognition of fibrinogen γC peptide by the platelet integrin $\alpha IIb\beta 3$. *Journal of Cell Biology*, 182(4), 791–800.
<https://doi.org/10.1083/jcb.200801146>
- Stewart, M., Thiel, M., & Hogg, N. (1995). Leukocyte integrins. *Current Opinion in Cell Biology*, 7(5),

690–696. [https://doi.org/10.1016/0955-0674\(95\)80111-1](https://doi.org/10.1016/0955-0674(95)80111-1)

- Takada, Y., & Puzon, W. (1993). Identification of a regulatory region of integrin $\beta 1$ subunit using activating and inhibiting antibodies. *Journal of Biological Chemistry*, *268*(23), 17597–17601. [https://doi.org/10.1016/s0021-9258\(19\)85374-7](https://doi.org/10.1016/s0021-9258(19)85374-7)
- Takada, Y., Ye, X., & Simon, S. (2007). The integrins. *Genome Biology*, *8*(5). <https://doi.org/10.1186/gb-2007-8-5-215>
- Van Agthoven, J. F., Xiong, J. P., Alonso, J. L., Rui, X., Adair, B. D., Goodman, S. L., & Arnaout, M. A. (2014a). Structural basis for pure antagonism of integrin $\alpha V\beta 3$ by a high-affinity form of fibronectin. *Nature Structural and Molecular Biology*, *21*(4), 383–388. <https://doi.org/10.1038/nsmb.2797>
- Van Agthoven, J. F., Xiong, J. P., Alonso, J. L., Rui, X., Adair, B. D., Goodman, S. L., & Arnaout, M. A. (2014b). Structural basis for pure antagonism of integrin $\alpha V\beta 3$ by a high-affinity form of fibronectin. *Nature Structural and Molecular Biology*, *21*(4), 383–388. <https://doi.org/10.1038/nsmb.2797>
- Wang, H. Y., Chen, Z., Wang, Z. H., Wang, H., & Huang, L. M. (2013). Prognostic significance of $\alpha 5\beta 1$ -integrin expression in cervical cancer. In *Asian Pacific Journal of Cancer Prevention* (Vol. 14, Issue 6, pp. 3891–3895). <https://doi.org/10.7314/APJCP.2013.14.6.3891>
- Wilmot, C. M., & Thornton, J. M. (1990). β -turns and their distortions: A proposed new nomenclature. *Protein Engineering, Design and Selection*, *3*(6), 479–493. <https://doi.org/10.1093/protein/3.6.479>
- Wojcik, J., Lamontanara, A. J., Grabe, G., Koide, A., Akin, L., Gerig, B., Hantschel, O., & Koide, S. (2016). Allosteric inhibition of Bcr-Abl kinase by high affinity monobody inhibitors directed to the Src homology 2 (SH2)-kinase interface. *Journal of Biological Chemistry*, *291*(16), 8836–8847. <https://doi.org/10.1074/jbc.M115.707901>
- Wu, H., Beuerlein, G., Nie, Y., Smith, H., Lee, B. A., Hensler, M., Huse, W. D., & Watkins, J. D. (1998). Stepwise in vitro affinity maturation of Vitaxin, an $\alpha v\beta 3$ -specific humanized mAb. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(11), 6037–6042. <https://doi.org/10.1073/pnas.95.11.6037>
- Xia, W., & Springer, T. A. (2014). Metal ion and ligand binding of Integrin $\alpha 5\beta 1$. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(50), 17863–17868. <https://doi.org/10.1073/pnas.1420645111>
- Xiao, T., Takagi, J., Collier, B. S., Wang, J.-H., & Springer, T. A. (2004). Structural basis for allostery in integrins and binding to fibrinogen-mimetic therapeutics. *Nature*, *432*(7013), 59–67. <https://doi.org/10.1038/nature02976>
- Xie, C., Zhu, J., Chen, X., Mi, L., Nishida, N., & Springer, T. A. (2010). Structure of an integrin with an αi domain, complement receptor type 4. *EMBO Journal*, *29*(3), 666–679. <https://doi.org/10.1038/emboj.2009.367>
- Xiong, J. P., Stehle, T., Zhang, R., Joachimiak, A., Frech, M., Goodman, S. L., & Arnaout, M. A. (2002). Crystal structure of the extracellular segment of integrin $\alpha v\beta 3$ in complex with an Arg-Gly-Asp ligand. *Science*, *296*(5565), 151–155. <https://doi.org/10.1126/science.1069040>
- Xiong, J., Stehle, T., Diefenbach, B., Zhang, R., Dunker, R., Scott, D. L., Joachimiak, A., Goodman, S., & Arnaout, M. A. (2001). Crystal structure of the extracellular segment of integrin alpha V beta 3. *Science*, *294*(5541), 339–345. <http://www.sciencemag.org/cgi/content/abstract/294/5541/339%5Cnpapers://6cd7fee5->

76aa-4f6d-875a-48c3d3370b0d/Paper/p1176

- Yan, C., Wu, F., Jernigan, R. L., Dobbs, D., & Honavar, V. (2008). Characterization of protein-protein interfaces. *Protein Journal*, 27(1), 59–70. <https://doi.org/10.1007/s10930-007-9108-x>
- Yang, W., Shimaoka, M., Chen, J. F., & Springer, T. A. (2004). Activation of integrin β -subunit I-like domains by one-turn C-terminal α -helix deletions. *Proceedings of the National Academy of Sciences of the United States of America*, 101(8), 2333–2338. <https://doi.org/10.1073/pnas.0307291101>
- Yu, Y., Schürpf, T., & Springer, T. A. (2013). How natalizumab binds and antagonizes α 4 integrins. *Journal of Biological Chemistry*, 288(45), 32314–32325. <https://doi.org/10.1074/jbc.M113.501668>
- Yu, Y., Zhu, J., Mi, L. Z., Walz, T., Sun, H., Chen, J. F., & Springer, T. A. (2012). Structural specializations of α 4 β 7, an integrin that mediates rolling adhesion. *Journal of Cell Biology*, 196(1), 131–146. <https://doi.org/10.1083/jcb.201110023>
- Zhu, J., Choi, W. S., McCoy, J. G., Negri, A., Zhu, J., Naini, S., Li, J., Shen, M., Huang, W., Bougie, D., Rasmussen, M., Aster, R., Thomas, C. J., Filizola, M., Springer, T. A., & Collier, B. S. (2012). Structure-guided design of a high-affinity platelet integrin α IIb β 3 receptor antagonist that disrupts Mg²⁺ binding to the MIDAS. *Science Translational Medicine*, 4(125). <https://doi.org/10.1126/scitranslmed.3003576>
- Zhu, J., Luo, B. H., Xiao, T., Zhang, C., Nishida, N., & Springer, T. A. (2008). Structure of a Complete Integrin Ectodomain in a Physiologic Resting State and Activation and Deactivation by Applied Forces. *Molecular Cell*, 32(6), 849–861. <https://doi.org/10.1016/j.molcel.2008.11.018>
- Zhu, J., Zhu, J., Negri, A., Provasi, D., Filizola, M., Collier, B. S., & Springer, T. A. (2010). Closed headpiece of integrin α IIb β 3 and its complex with an α IIb β 3-specific antagonist that does not induce opening. *Blood*, 116(23), 5050–5059. <https://doi.org/10.1182/blood-2010-04-281154>
- Zhu, J., Zhu, J., & Springer, T. A. (2013). Complete integrin headpiece opening in eight steps. *Journal of Cell Biology*, 201(7), 1053–1068. <https://doi.org/10.1083/jcb.201212037>
- Zollinger, A. J., & Smith, M. L. (2017). Fibronectin, the extracellular glue. In *Matrix Biology* (Vols. 60–61, pp. 27–37). Elsevier B.V. <https://doi.org/10.1016/j.matbio.2016.07.011>
- Winklbauer R, Stoltz C. Fibronectin fibril growth in the extracellular matrix of the Xenopus embryo. *J Cell Sci*. 1995 Apr;108 (Pt 4):1575-86. doi: 10.1242/jcs.108.4.1575. PMID: 7615676.