Structure-Function Relationship of the Disintegrin Family: Sequence Signature and Integrin Interaction

Ariana A. Vasconcelos1,2†, Jorge C. Estrada3†, Victor David3, Luciana S. Wermelinger4, Fabio C. L. Almeida1,2* and Russolina B. Zingali3*

1Instituto de Bioquímica Médica (IBqM) Leopoldo de Meis, Centro Nacional de Ressonância Magnética Nuclear, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil, 2Centro Nacional de Ressonância Magnética Nuclear (CENARBIO), Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil, 3Laboratório de Hemostase e Venenos, Instituto de Bioquímica Médica (IBqM) Leopoldo de Meis, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil, 4Faculdade de Farmácia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil

Disintegrins are small cysteine-rich proteins found in a variety of snake venom. These proteins selectively modulate integrin function, heterodimeric receptors involved in cell-cell and cell-matrix interaction that are widely studied as therapeutic targets. Snake venom disintegrins emerged from the snake venom metalloproteinase and are classified according to the sequence size and number of disulfide bonds. Evolutive structure and function diversification of disintegrin family involves a stepwise decrease in the polypeptide chain, loss of cysteine residues, and selectivity. Since the structure elucidation of echistatin, the description of the structural properties of disintegrins has allowed the investigation of the mechanisms involved in integrin-cell-extracellular matrix interaction. This review provides an analysis of the structures of all family groups enabling the description of an expanded classification of the disintegrin family in seven groups. Each group presents a particular disulfide pattern and sequence signatures, facilitating the identification of new disintegrins. The classification was based on the disintegrin-like domain of the human metalloproteinase (ADAM-10). We also present the sequence and structural signatures important for disintegrin-integrin interaction, unveiling the relationship between the structure and function of these proteins.

Keywords: snake venom disintegrin, structure, integrin, NMR, crystallography

INTRODUCTION

Integrin antagonists comprise molecules that can bind and interfere with the activity of the cellular receptors integrins. Some proteins or peptides are found in snake venoms such as the C-type lectins EMS6 (Marcinkiewicz et al., 2000) and Vixapatin (Momic et al., 2012) or disintegrins (Calvete, 2005a), or found in other animals venoms such as arthropods (Khamessi et al., 2018). The snake venom disintegrin family comprises a group of cysteine-rich proteins (40–100 amino acids) found in the venom from snakes from Elapidae, Viperidae, Atractaspididae, and Colubridae families (Kini and Evans, 1992; Calvete et al., 2005b; Arruda Macedo et al., 2015). These proteins are released in the venom as a result of the proteolytic process of the so-called PII snake venom metalloproteinases (SVMP) (Kini and Evans, 1992; Calvete et al., 2005b). It is worth emphasising that venoms from other venomous animals also contain metalloproteinases similar to that from snake venoms (Xia et al., 2013; Borges et al., 2016). Whether these molecules are able to generate disintegrins is yet to be
elucidated. Snake venoms disintegrins were firstly described by the capacity to inhibit the platelet fibrinogen receptor integrin αIIbβ3 (Huang et al., 1987). Disintegrins are capable of modulating the function of a broad range of integrins (Gould et al., 1990; McLane et al., 2004), a family of heterodimeric receptors that play a fundamental role in mediating physiological and pathological processes, such as hemostasis and cancer (Phillips et al., 1980; Hynes, 1987). Many aspects of the disintegrin protein family were reviewed recently: as tools for antitumor activity (Schönhal et al., 2020), for antithrombotic agents (Kuo et al., 2019), recombinant and chimeric proteins on research (Huang et al., 2016; David et al., 2018; Uzair et al., 2018; Cesar et al., 2019; Lazarovici et al., 2019). This review focuses on structural studies of snake venom disintegrins and the features of integrin-disintegrin interaction.

**STRUCTURAL AND EVOLUTIVE BASES OF DISINTEGRIN MOLECULES**

**Evolutive Diversification of Metalloproteinases Into Disintegrins**

The emergence of disintegrins results from the divergence of snake venom Zn$^{2+}$-Metalloproteinases (SVMPs). These proteins are present in the venom of the majority of venomous snakes and are capable of degrading the extracellular matrix and/or proteins belonging to the hemostasis system (de Lima et al., 2009; Fox and Serrano, 2009; Mackessy, 2009). SVMPs share a common ancestor with matrix-degrading metalloproteinases, the A Disintegrin And Metalloproteinase (ADAM) family (Moura-da-Silva et al., 1996). SVMPs are classified according to their domain organization: the PIII class contains the metalloproteinase domain followed by the C-terminal disintegrin-like and cysteine-rich domain; the PII class, contains the disintegrin domain at the C-terminal of the metalloproteinase domain and the P class contains only the metalloproteinase domain (Fox and Serrano, 2009). The subfamily of ADAM containing the thrombospondin type-1 motif (ADAMTS) are also related to SVMPs, but the crystallographic structures showed that the disintegrin-like domain of ADAMTS adopts a different fold and are not structurally homologous to disintegrins (Takeda et al., 2012). For this reason, the analysis presented in this review will consider only the ADAM family.

The evolution of disintegrin family was extensively studied by many authors (for review see: Juárez et al., 2008; Calvete, 2013; Fox and Serrano, 2009). The accepted hypothesis is that SVMPs have evolved relatively late from a common ancestor by speciation and positive Darwinian selection. The evolution started by gene duplication of the ancestral PIII disintegrin-like domain followed by neofunctionalization in the snake venoms generating the disintegrin domains of PII (Fox and Serrano, 2009). The evolution from disintegrin-like (PIII) to disintegrin domain (PII) occurred through the successive loss of disulfide bonds and reduction in size to the different snake venom disintegrin subfamilies (long, medium-sized, dimeric, and short (Kini and Evans, 1992; Calvete, 2005a). The emergence of PII SVMP precursors follows another key event that includes deletion of the C-terminal cysteine-rich domain, due to a mutation causing the appearance of a stop codon, and the removal disulfide bonds (Calvete et al., 2003; Júarez et al., 2006a). PII SVMP undergoes limited proteolysis, releasing disintegrins domains (short, medium, and long) into the venom (Kini and Evans, 1992; Jia et al., 1996). Most proteins from dimeric or short groups are synthesized from short-coding mRNAs, lacking the metalloproteinase domain (Juárez et al., 2006b; Okuda et al., 2002; Bazaa et al., 2007). In fact, analysis of the pre-sequence of the cDNA disintegrin jerdostatin suggests that this disintegrin originates from a short-coding gene, instead of a proteolytic process such as the majority of disintegrin from PII SVMP (Sanz et al., 2005). Moreover, Bazaa et al., 2007 suggested that the shortening of the gene is due to the loss of introns and coding regions that contribute to the formation of the short-coding disintegrins. In summary, disintegrins originate from a multigene family from metalloproteinases that undergoes an accelerated evolutionary pathway resulting in great diversification (Juárez et al., 2008).

**Structural Diversification of Disintegrins**

Disintegrins have been classified according to the number of residues and disulfide bridges, into four subfamilies: smaller ones are the short disintegrins composed of 49–51 residues and four disulfide bridges; the medium size disintegrins with about 70 residues and six bridges; the largest consist in the long disintegrin with about 84 residues cross-linked by seven disulfide bridges; and homo- and heterodimers disintegrins (McLane et al., 1998; Calvete et al., 2003). Dimeric disintegrins contain subunits about 67 residues cross-linked by 2 interchain cysteine linkages and 4 intra-chain disulfide bonds (Calvete et al., 2000; Bilgrami et al., 2005; Figure 1). The evolution pathway of disintegrin structure diversification involved the reduction of the polypeptide chain and selective loss of pairs of cysteine residues that form disulfide bonds (Calvete et al., 2003; Calvete, 2005a, 2010; Bazaa et al., 2007). Phylogenetic analysis suggests that PII-dimeric and short disintegrins represent the more recent diverging lineages of disintegrins (Figure 1A) (Juárez et al., 2008).

Additionally, disintegrins present much more complex structural diversity. To illustrate, disintegrins jarastatin (Bothrops jararaca)1, trilavín (Trimeresurus flavoviridis), and obtustatin exhibit different disulfide bridge patterns and structural features. Jarastatin and trilavín are disintegrins that contain six disulfide bonds with the same cysteine pattern, however, these disintegrins have different cystine connectives (Huang T. F. et al., 1991; Coelho et al., 1999; Cidade et al., 2006). Likewise, trilavín has an elongated and rigid structure composed of turns and antiparallel

---

1The snake species names were taken from the original papers that describe the disintegrins, some scientifics names of these species could have changed.
β-strands, and obtustatin has a compact globular structure composed of turns and without regular secondary structure (Huang T. F. et al., 1991; Coelho et al., 1999; Fujii et al., 2003; McLane et al., 2004; Cidade et al., 2006; Wermelinger et al., 2009). These cysteine pairing characteristics motivated us to perform a re-analysis of the classification of disintegrins based on the 3D structure to seek a better structural/function comprehension of disintegrins (see point section 4.1).

Moreover, despite the majority of disintegrins following canonical structure features, some of them follow a different pathway (Calvete et al., 2003). The disintegrin graminelysin belongs to medium-size disintegrins but has some features of PIII derived disintegrins, such as the Cys13-Cys16 disulfide bond, which represent an intermediate step in the evolution pathway of medium-sized disintegrins (Wu et al., 2001). Bilitoxin-1 is a long disintegrin containing an additional cysteine residue involved in a disulfide bond homodimer (Nikai et al., 2000). Another example includes the identification of different PII SVMP transcripts found in the venom of Bothrops neuwiedi, suggesting recombination of PII SVMP with the catalytic domain of P1 or PIII SVMP (Moura-da-Silva et al., 2011). This fact is well described below, where the groups are presented based on structural and cysteine pairing features. We also show a cladogram that illustrates the evolutionary path of the disintegrins.

**FIGURE 1** | Evolutionary relationships between disintegrins and current classification. (A) Dendrogram shows the evolutionary relationships between the different disintegrin subfamilies. The dendrogram also includes the disintegrin-like domain of ADAMs and PIII-SVMPs from which the different snake venom disintegrin subfamilies (long, medium, dimeric and short) evolved through the successive loss of disulfide bonds and size reduction. (B) Current classification of disintegrins: long (~84 amino acids and 7 disulfide bonds) (connections between cysteines in lavender), medium (~70 amino acids and 6 disulfide bonds) (connections between cysteines in red), dimeric (~67 amino acids and 4 intrachain disulfide bonds for each subunit and 2 interchain) (connections in green), and short (41–51 amino acids and 4 disulfide bonds) (connections in blue). The integrin-binding RGD, KGD, MGD, MDL, KTS, and RTS tripeptide motifs localization is indicated by the magenta rectangle. (Adapted from Juárez et al., 2008).

**DISINTEGRIN STRUCTURE: NMR AND CRYSTALLIZATION STUDIES**

The number of disintegrins purified from the venom or identified by proteome or transcriptome analysis grows each year. The NMR studies expanded the panorama of the structure of disintegrins. Nevertheless, the structure of only 16 disintegrins has been solved by NMR or...
crystallography until now. Figure 2 presents the timeline of experimental information on disintegrin structure.

**An Expanded Classification of Disintegrins Based on the Structural Features, Disulfide Pattern, and Comparison With Ancestral Disintegrin Fold**

In this section, we propose an expanded classification based on the structural features of disintegrins and use the disintegrin-like domain of a PIII ADAM as a reference for folding and disulfide pattern. This reference is suitable since it is well established by several in-death phylogenetic studies (Juárez et al., 2008; Fox and Serrano, 2009; Calvete, 2013) that the PIII disintegrin-like domain is ancestral of all disintegrins. As previously mentioned in this work, disintegrins have a vast structural diversity. Within the disintegrin family there is a wide variety of disulfide bridge patterns and protein sizes and, at the same time, conserved structural features (Huang T. F. et al., 1991; Coelho et al., 1999; Cidade et al., 2006; Cheng et al., 2012). For instance, jarastatin and triflavin disintegrins contain six disulfide bonds with the same cysteine pattern while the disintegrin domain of ADAM metalloproteinases present the same size, structural homology, and different disulfide pattern (Huang T.-F. et al., 1991; Coelho et al., 1999; Cidade et al., 2006).

Pairing cysteine residues in disintegrins is already known to play an important role in exposing the RGD binding motif that mediates inhibition of platelet aggregation, neutrophils, or endothelial cell function (Blodel and White, 1992; Takagi et al., 2002; Calvete et al., 2003; Calvete, 2005a). In addition, the modulatory activity of the disintegrins depends on the proper pairing of the cysteine residues, contributing to the conformation of the disintegrin structure (Niewiarowski et al., 1994). The conservative aspect of cysteine residues and
FIGURE 3] Sequence alignments of representative disintegrins. The sequences were obtained through research in the Blast software (Altschul et al., 1990; Gish and States, 1993). The pink color highlights the sequence signatures, green the residues that are not conserved within the signatures, orange the cysteines. Dotted blue squares show the conserved L/I that compose the hydrophobic core. RGD/KGD or related three-peptides are underlined. The codes representing the sequences were obtained from PDB and UniProt. Abbreviated: VAP-1 (Vascular apoptosis-inducing protein 1); RRV-X (Russel’s viper venom factor X activator); EcoHIT (Viper PIII hemorrhagic SVMP Echis coloratus); EcH (Metalloproteinase Echis pyramidum leakeyi); G. shedaoensis (G. shedaoensis); SSV (saw-scaled Viper). The multiple sequence alignment was performed using the ClustalW2 server (https://www.genome.jp/tools-bin/clustalw).
the disulfide bond pattern between the disintegrins subfamily contribute to the hypothesis of strong selection for maintaining the active conformation of these proteins (Juárez et al., 2008). Therefore, we decided to investigate in more detail the patterns of protein disulfide bridges within the structure of some members of the disintegrin family, which show disintegrin fold either as an isolated protein or as a domain of a larger protein.

To select these proteins, we used the PFAM family (PF0020) classification, which describes all the subfamilies and their correspondent structures, contains annotations and multiple sequence alignments. We analyzed the sequence and structure...
alignment and proposed a classification based on the pattern of disulfide bonds using the ancestral subfamily ADAM as a reference. We assigned letters from A to N for each possible position occupied by these cysteines. The analysis pointed out seven possible disulfide patterns, named group 1, for proteins of ADAM subfamily; group 2, for disintegrins similar to bitistatin A; group 3, bitistatin B; group 4, for kistrin; group 5 for salmosin; group 6, for dimers; group 7, for obtustatin. Figure 3 shows the sequence alignment of each group proposed in this classification.

The classification of the disulfide bond patterns followed the tertiary structure alignments with the reference protein ADAM10 Extracellular Domain (ADAM10, PDB id 6BE6). Only the alignments of primary structure did not provide sufficient information to classify the pattern within the disintegrin family, whereas the comparison with ADAM10 provided the exact location and classification for each cysteine. Figure 4 shows the disulfide pattern for each group and the superposition of one component of each group with ADAM10. It illustrates the rationale of the classification proposed here.

For group 1 (Figure 4A), we analyzed 11 representatives, which have a disintegrin domain with up to 14 cysteines. They are organized into 6 disulfide bonds and two cysteines that may form disulfide bonds with cysteines in the other domain of the protein (Jia et al., 1996). The letter-based pattern is defined as AE BC DJ F GI HM L KN, where F and L, which appear here unmatched, are cysteines involved in connections with the other domains. This group is formed by mammalian (ADAM) and snake venom metalloproteinases (SVMP), both containing the disintegrin domain with the same fold.

Group 2 and 3 represent two disulfide patterns (alternate folding) for the same sequence. They were first described for the disintegrin bitistatin, group 2 for bitistatin A (Figure 4B), and 3 for bitistatin B (Figure 4C). The disulfide pattern for group 2 is AD BG CF EK HJ IM LN and AG BF CD EK HJ IM LN for group 3 differing only at disulfide bonds at the N-terminal region.

Group 4 (Figure 4D) is the group of kistrin, which is also presented as only 1 disintegrin monomeric domain. It consists of 15 proteins that have 12 cysteines, all involved in disulfide bonds. The pattern is AF BE DJ GI HM KN with the cysteines of positions C and L were evolutionarily replaced by other amino acids. We noted that the disulfide pattern is the same as that found for the ancestral subfamily ADAM, differing only at the N-terminal region. Despite this, the detailed analysis of the sequence alignment (Figure 3) of the studied groups showed that the ADAM group (group 1) and the Kistrin group (group 4), have many conserved amino acid sequences in the N-terminal portion.

Group 5 (Figure 4E) are also monomeric proteins with 1 disintegrin domain. It consists of 11 proteins, but the only member with a high-resolution structure available is the salmosin. This disintegrin has a different disulfide pattern, represented by AD BE FI HJ GM KN. Interestingly, the cysteines that occupied the C and L positions in the ADAM ancestor were replaced by other amino acids throughout the evolutionary process and for this reason, they are not included in this pairing. Therefore, members of this group include 12 cysteines involved in 6 disulfide bonds. Salmosin is the only available disintegrin structure solved so far that cannot be completely superposed, differing at the N-terminal region, with the reference group (group 1) and other disintegrins.

Group 6 (Figure 4F) are the dimeric disintegrins. They include 4 representatives with resolved structures. The dimeric disintegrins are small and present the N-terminal region quite conserved among themselves, but distinct when compared to the other disintegrins. They have 10 cysteines, of which 8 are involved in intra-chain disulfide bonds and two are involved in inter-chain cystine linkages (cysteines E and F). Also, they present evolutionary replacement at the L position for other amino acids. The pattern observed for this group is DJ EF GI HM KN, containing the sequence signature NPCC at the N-terminal (Figure 3 and Table 1).

Group 7 (Figure 4G) comprises the small disintegrins, which have members that have, in addition to the RGD motif, KTS, and RTS sequences, with a GJ HM IL KN binding pattern, these disintegrins have 8 cysteines that form 4 disulfide bonds. In addition, group 7 displays few sequence signatures in common with other groups (Table 1).

The new classification of the disintegrin family provided the description of the sequence signatures that will enable localization of an unknown sequence within a group and to be able to predict properties such as structural features and functional capabilities. We looked for sequence signatures within each group that may help to classify and model the structure of an unknown sequence. The signature GxECDC sequence is common for the groups G1, G2, G3, G4, and G5, absent in groups G6 and G7 (short and dimeric disintegrins, Table 1). Here, we emphasize that group 7 was subdivided into 7A and 7B.

The sequences CRxARGD and CCxQCxF are common to groups G2, G3, G4, and G5, which relate to medium and long proteins. The CRxARGD signature is shared with the salmosin group, a more distinct group due to its more differentiated disulfide bond pattern.

We also observed signatures present in some related groups, such as CCDAATCKLxxGAQC and DDxCxG, which were common to groups G4 and G5. Both groups harbor medium-sized disintegrins, but with different folds. We also observed that short disintegrins exhibit the highest variability in their integrin recognition motifs, including, in addition to RGD, the KTS and RTS motifs. This variability in the motifs for integrin recognition was also observed for G6, a group of dimeric disintegrins, which

| Table 1 | Signatures of the sequences of the study groups based on the classification of seven groups.

| Sequence | G1 | G2 | G3 | G4 | G5 | G6 | G7A | G7B |
|----------|----|----|----|----|----|----|-----|-----|
| GxECDCC | ✔  | ✔  | ✔  | ✔  | ✔  | ✔  | –   | –   |
| CCCDAATCKLxxGAQC | –  | –  | –  | ✔  | ✔  | ✔  | ✔   | –   |
| GPCCR    | –  | –  | ✔  | ✔  | ✔  | ✔  | ✔   | ✔   |
| CRxARGD  | –  | ✔  | ✔  | ✔  | ✔  | ✔  | –   | –   |
| CxQxCxF  | –  | ✔  | ✔  | ✔  | ✔  | ✔  | ✔   | ✔   |
| DDxCxG   | –  | –  | –  | –  | ✔  | ✔  | –   | ✔   |
| DCPFR    | –  | –  | ✔  | ✔  | ✔  | ✔  | –   | –   |
| NxCC     | –  | –  | ✔  | ✔  | ✔  | ✔  | ✔   | ✔   |

x variation of amino acid residue.

Frontiers in Molecular Biosciences | www.frontiersin.org 7 December 2021 | Volume 8 | Article 783301

Vasconcelos et al. Structure and Function of Disintegrins
Disintegrin Structural Properties

Since the purification of Echistatin, isolated for the first time in 1988 from the venom of *Echis carinatus* snake, some disintegrins’ structures were solved leading to a better understanding of the structure-activity relationship. As mentioned before, the disintegrins vary in size and fold, but they display many conserved structural features. Table 2 summarizes the information about all disintegrins that had their structure solved so far. They are stabilized by multiple disulfide bonds (from 4 to 7), they lack a canonical hydrophobic core and well-defined secondary structures.

The disintegrin structure consists of a series of loops tightly held together by disulfide bonds with almost no regular secondary structure. All hydrophobic residues are at least partially exposed to the protein surface, except for one leucine/isoleucine residue that is buried in a protein core. This I/L residue is present in all groups and should be important for protein stabilization. More studies are necessary to understand the role of the hydrophobic residues in the overall folding. The exposure of hydrophobic residues and folding without a canonical hydrophobic core is not exclusive for disintegrins and it is also present in defensins (Machado et al., 2018; Pinheiro-Aguiar et al., 2020). It was recently proposed that these proteins are stabilized by hydrophobic surface clusters acting as an independent folding unit. In the absence of a canonical hydrophobic core, the surface clusters promote the folding by the interaction of exposed hydrophobic residues with the adjacent side chains regulated by solvation forces (Almeida et al., 2021).

There are exceptions for the disintegrin general structure. One exception is bitistatin, which is mainly found in two different forms, as bitistatin A and B. Both have an identical primary structure, their molecular architecture can be defined, generally, as a fold with an elongated shape, including a “ladder” of seven disulfide bonds, representing the dominant organizational characteristic of polypeptide folds. Some differences are found in the intra-domain interactions between the two bitistanins. In bitistatin A, some hydrophobic interactions are present between Val4-Pro21 and between Ile9 and the fragment Glu11-Cys18. In contrast, the Bitistatin B has the Pro3 positioned adjacent to the side chain.
### TABLE 2 | Summary of the sequence, ligands, structural information and references of the disintegrins structures reported in PDB database.

| Disintegrin name                        | Length (aa) | Number of cysteines | Ligand | Specie | Loop | Method | RMSD resolution (Å) | PDB ID | Group | References                          |
|-----------------------------------------|-------------|---------------------|--------|--------|------|--------|---------------------|--------|-------|-------------------------------------|
| Bitistatin                              | 89          | 14                 | αβ3    | Bitis arietans | RGD | NMR   | 0.50                | 2MOP/   | 2MPS  | 2–3 Carbaj et al. (2015)            |
| Kistrin                                 | 68          | 12                 | αβ3    | Agkistrodon rhodostoma | RGD | NMR   | 0.47                | 1N4Y   | 4      | Cooke et al. (1992)                 |
| Rhodostomin                             | 68          | 12                 | αβ3    | Calloselasma rhodostoma | RGD | NMR   | 0.61                | 2PJF   | 4      | Guo et al. (2001)                  |
| Rhodostomin R46E mutant                 | 68          | 12                 | αβ3    | Calloselasma rhodostoma | RGD | CRYSTAL | NA                | 4R5U   | 4      | Huang et al. (2015a)               |
| Rhodostomin DS1E mutant                 | 68          | 12                 | αβ3    | Calloselasma rhodostoma | RGD | NMR   | 0.59                | 2PJG/   | 2PJF  | 4 Chen et al. (2009)               |
| Rhodostomin P48A mutant                 | 68          | 12                 | αβ1    | Calloselasma rhodostoma | RGD | NMR   | 0.57                | 2PJ1    | 4      | Chuang et al. (2007)               |
| Rhodostomin G50L mutant                 | 68          | 12                 | αβ3    | Calloselasma rhodostoma | RGD | NMR   | 0.45                | 2LJV    | 4      | Chuang et al. (2012)               |
| Rhodostomin ARGDP mutant                | 68          | 12                 | αβ3    | Calloselasma rhodostoma | RGD | CRYSTAL | NA                | 4M4C    | 4      | Chang et al. (2014)                |
| Rhodostomin KKKKT mutant                | 68          | 12                 | αβ3    | Calloselasma rhodostoma | RGD | CRYSTAL | NA                | 4R5R    | b 4 Huang et al. (2015b)           |
| Rhodostomin ARLDDL mutant               | 72          | 12                 | αβ3    | Calloselasma rhodostoma | RGD | CRYSTAL | NA                | 3UCI    | 4      | Shiu et al. (2012)                 |
| Disintegrin with the AKGDW motif        | 68          | 12                 | αβ3    | Calloselasma rhodostoma | RGD | NMR   | 0.75                | 1Q7I    | 4      | Chuang et al. (2004)               |
| Rhodostomin mutante Protein RGD Motif C-terminal mutation | 79      | 12                 | αβ3    | Calloselasma rhodostoma | RGD | NMR,   | 0.54                | 2M7F    | 4      | Chang et al. (2014)                |
| Rhodostomin mutante Protein RGD Motif C-terminal mutation | 79      | 12                 | αβ3    | Calloselasma rhodostoma | RGD | NMR   | 0.51                | 2M7H    | 4      | Chuang et al. (2013)               |
| Rhodostomin mutante Protein RGD Motif C-terminal mutation | 76      | 12                 | αβ3    | Calloselasma rhodostoma | RGD | NMR   | 0.60                | 2M7S    | 4      | Chuang et al. (2013)               |
| Flavoridin                              | 70          | 12                 | αβ3    | Trimeresurus flavoviris | RGD | NMR   | 0.85                | 1FVL    | 4      | Senn and Werner, (1993)            |
| Trimestatin                             | 70          | 12                 | αβ3    | Trimeresurus flavoviris | RGD | CRYSTAL | NA                | 1J2L    | 4      | Fuji et al. (2003)                 |
| Acostatin                               | 64          | 10                 | αβ3    | Agkistrodon contortrix contortrix | RGD | CRYSTAL | NA                | 3CO5    | 4–5 Moiseeva et al. (2008)         |
| Abolabrin                               | 84          | 12                 | αβ3    | Trimeresurus abalabris  | RGD | NMR   | NA                  | NA      | 5      | Smith et al. (1996)                |
| Salmosin                                | 70          | 12                 | αβ3    | Agkistrodon haley      | RGD | NMR   | 0.49                | 1L3X    | 5      | Shin et al. (2003)                 |
| Schistatin                              | 64          | 10                 | αβ3    | Echis carinatus       | RGD | CRYSTAL | NA                | 1RMR    | 6      | Bilgrami et al. (2004)             |
| Heterodimer from Saw-scaled viper       | 64          | 10                 | αβ3    | Echis carinatus       | RGD | CRYSTAL | NA                | 1TEJ    | 6      | Bilgrami et al. (2005)             |

(Continued on following page)
The structures present the same disulfide pattern of group 6, a heterodimeric disintegrin isolated from the venom of *Agkistrodon contortrix contortrix*. The crystal structure shows a tetramer (dimer of dimer), with each dimer with a similar fold. The structures present the same disulfide pattern of group 6, with Cys residues DJ, GI, HM, and KN forming intrachain disulfide bridges and E and F interchain disulfide bridges. Both dimers form an identical disulfide pattern. (Figure 4) (Moiseeva et al., 2008). Interestingly, only the sequence were to be taken into account, as in the cladogram construction, chain α would be wrongly classified as group 4 and chain β as group 5 (Figure 5).

One of the most different structures reported so far is the Salmosin (group 5). This disintegrin isolated from the *Agkistrodon halys* venom has the RGD motif conserved with an unusual finger shape and is distal from the rigid nucleus of the C terminal domain. In addition, although the RGD motif does not interact with the hydrophobic nucleus of the protein, it has been stabilized by a network of molecular contacts through a small beta propeller-like domain and one ligand-binding I domain (interacting domain) named α1 (Shattil and Newman, 2004; Barczyk et al., 2010; Dermont et al., 2010). β-subunit varies from 95 to 117 Kda. Each β-subunit contains a divalent cation binding site located at 100 residues from the amino terminus (Berman et al., 2003). The extracellular portion of the β subunit is composed of two calcium ions, one ligand, one β-propeller domain (β-propeller), and one ligand-binding I domain (interacting domain) named α and β domains possess specific binding regions for metallic ions. The presence of magnesium ion in the domain I of α, known as metal ion dependent-adhesion site (MIDAS), modulates the binding of integrin to the specific ligand. Some integrins, such as αIIbβ3, lacks αI domain, in that case, the ligand-binding lies between βI domain and β-propeller of the α subunit (Dermont et al., 2010; Coller, 2015). αI and βI domains promote a conformational change in the extracellular integrin portion, promoting separation of transmembrane and cytoplasmic

### INTEGRINS

Integrins are a large family of surface receptors, involved in cell-cell and cell interactions with extracellular matrix components, present in biological processes such as angiogenesis and hemostasis (Hynes, 1992; Ata and Antonescu, 2017). Integrins are heterodimers formed by one α and one β-subunit, stabilized by noncovalent bonds (Barczyk et al., 2010) (Figure 6A). Each integrin subunit combines to form 24 heterodimers, composed of 18 α subunits and 8 β (Barczyk et al., 2010). Each subunit has a transmembrane polypeptide type I; containing three domains, a glycosylated domain, a hydrophobic domain, and an endo-domain. The α subunit can vary in the range of 120–180 Kda. At the N-terminus, the α subunit possesses seven homologous domains of 50 amino acids each. The extracellular portion of α subunit is composed of two calcium ions, one ligand, one β-propeller domain (β-propeller), and one ligand-binding I domain (interacting domain) named α and β domains.
portion, a process known as inside-out activation. Intracellular signaling promotes conformational changes in the cytoplasmatic integrin tail until the ligand-binding domain, known as inside-out activation (Barczyk et al., 2010; Dermont et al., 2010). Three conformational states are reported for integrin receptors: a low-affinity state, where the integrin structure is bent (Figure 6A); an intermediate affinity state triggered by the binding to the ligand, where the integrin structure is elongated, however partially activated; and a high affinity conformational, also triggered by the ligand-binding, where the integrin structure is extended and opened (Figure 6A) (Barczyk et al., 2010; Dermont et al., 2010). Structural studies on integrins have been solved mainly by X-ray crystallography of the extracellular structure fragments, αI domain, transmembrane and cytosolic complex, and intracellular protein in complex with the cytosolic tail of integrins (Liddington, 2014). Although few examples of structural experiments of integrin were reported, some integrins described on disintegrin structure studies, such as αIbβ3 and αvβ3, have been characterized.

In addition to structural characteristics of αIbβ3 integrin, crystallography study of integrin αvβ3 reveals that NH2-terminal segments of the α and β subunits consist of an ovoid head with two parallel tails. The head is formed by a seven-bladed β-propeller from αv and a βA domain. Also, four solvent-exposed Ca2+ binding sites are found in the A-B β hairpin loops of blades 4 to seven at the propeller’s bottom. Finally, the αv ends in the thigh and calf-1 and 2 domains. Each calf domain contains two antiparallel β sheets, one with four strands and the other one with five (Xiong et al., 2001). Binding to RGD peptide, reveals that the pentagonal peptide fits into a crevice between the propeller and the βA domains on the integrin head. Also, the binding is associated with tertiary and quaternary changes in the integrin; affecting the α-1 α-2 loops and helices and the α2-C, F-α7 loops. At the same time, the βA and the αv propeller suffer a small change, with the two domains moving closer together at the peptide-binding site (Xiong et al., 2001). A study of the crystal structure of integrin αvβ3 with a cyclic pentapeptide ligand RGD showed that the peptide fits in a crevice between the β-propeller and the βA domains on the integrin head. Also, this study reveals that the Arg side chain gets into a narrow groove at the top of the propeller domain. Furthermore, the Asp made contact with the βA involving the Asp carboxylate group, this made a kind of network of polar hydrophobic interactions with the αv. The fact that the Asp interacts with the βA suggests this residue is responsible for the recognition of the βA (Xiong et al., 2001). Figure 6B shows the interaction of αvβ3 integrin with the RGD domain of fibronectin, representing the fit of the protein into the domains of the integrin. Possibly, disintegrins interact similarly, but a high-resolution structure of the disintegrin/integrin complex is not available.

STRUCTURAL FEATURES OF DISINTEGRIN-INTEGRIN INTERACTION

Only a very few reports show the interaction between disintegrin and integrins. The recognition of Integrins by disintegrins is mediated by the specific binding loop of each disintegrin, some of them containing the RGD sequence. In that matter, it is of interest to look upon the interaction of peptides containing RGD sequence and integrins.
### TABLE 3 | Affinity and interactions with integrins to each group of disintegrins according to the classification proposed.

| Groups | Disintegrin | Integrins | REF |
|--------|-------------|------------|-----|
|        |             | β3 | β1 | β6 | β7 | β5 | β2 |
|        |             | α1b | αv | α2 | α5 | α6 | α3 | α4 | α9 | α1 | α2 | α7 |
| G 1    | VAP1        | No  | Yes | Yes | No  | —  | —  | —  | —  | —  | —  | —  |
|        | SVMP Elapid (stralgin) | Yes | Yes | No  | Yes | —  | —  | —  | —  | —  | —  | —  |
|        | Factor X activator (Russells viper venom factor X activator RVV-X) | Yes | Yes | No  | Yes | No* | —  | —  | —  | —  | —  | —  |
|        | ADAM 22     | Yes | No* | Yes | No* | Yes | —  | —  | —  | —  | —  | —  |
|        | ADAM 10     | Yes | No  | Yes | —  | Yes | —  | —  | —  | —  | —  | —  |
| G 2/3  | Battlesin | Yes | Yes | —  | —  | —  | —  | —  | —  | —  | —  | —  |
|        | Kistrin     | Yes | Yes | No  | —  | —  | —  | —  | —  | —  | —  | —  |
|        | Rhodostomin | Yes | Yes | No  | No | —  | —  | —  | —  | No | —  | No |
|        | Flavondonin | Yes | Yes | No  | No | —  | —  | —  | —  | —  | —  | —  |
|        | Tzabcanin   | Yes | No  | No  | —  | —  | —  | —  | —  | —  | —  | —  |
|        | Barbourin   | Yes | No  | No  | —  | —  | —  | —  | —  | —  | —  | —  |
|        | Cerastin    | Yes | Yes | No  | —  | —  | —  | —  | —  | —  | —  | —  |
|        | Tergemimin  | Yes | Yes | No  | —  | —  | —  | —  | —  | —  | —  | —  |
|        | Jararacin   | Yes | Yes | No  | —  | —  | —  | —  | —  | —  | —  | —  |
|        | Lutosin     | Yes | Yes | No  | —  | —  | —  | —  | —  | —  | —  | —  |
|        | Veldin      | Yes | Yes | No  | —  | —  | —  | —  | —  | —  | —  | —  |
|        | Molosin     | Yes | Yes | No  | —  | —  | —  | —  | —  | —  | —  | —  |
|        | Triffavin   | Yes | Yes | No  | —  | —  | —  | —  | —  | —  | —  | —  |
|        | Cotiarin    | Yes | Yes | No  | —  | —  | —  | —  | —  | —  | —  | —  |
| G 4    | Kistrin     | Yes | Yes | No  | —  | —  | —  | —  | —  | —  | —  | —  |
|        | Rhodostomin | Yes | Yes | No  | No | —  | —  | —  | —  | —  | —  | No |
|        | Flavondonin | Yes | Yes | No  | No | —  | —  | —  | —  | —  | —  | —  |
|        | Tzabcanin   | Yes | No  | No  | —  | —  | —  | —  | —  | —  | —  | —  |
|        | Barbourin   | Yes | No  | No  | —  | —  | —  | —  | —  | —  | —  | —  |
|        | Cerastin    | Yes | Yes | No  | —  | —  | —  | —  | —  | —  | —  | —  |
|        | Tergemimin  | Yes | Yes | No  | —  | —  | —  | —  | —  | —  | —  | —  |
|        | Jararacin   | Yes | Yes | No  | —  | —  | —  | —  | —  | —  | —  | —  |
|        | Lutosin     | Yes | Yes | No  | —  | —  | —  | —  | —  | —  | —  | —  |
|        | Veldin      | Yes | Yes | No  | —  | —  | —  | —  | —  | —  | —  | —  |
|        | Molosin     | Yes | Yes | No  | —  | —  | —  | —  | —  | —  | —  | —  |
|        | Triffavin   | Yes | Yes | No  | —  | —  | —  | —  | —  | —  | —  | —  |
|        | Cotiarin    | Yes | Yes | No  | —  | —  | —  | —  | —  | —  | —  | —  |
| G 5    | Salmosin    | Yes | Yes | No  | —  | —  | —  | —  | —  | —  | —  | —  |
|        | Adinbitor   | Yes | No  | —  | —  | —  | —  | —  | —  | —  | —  | —  |
|        | Jarastatin  | Yes | No  | —  | —  | —  | —  | —  | —  | —  | —  | —  |
|        | Saxatilin   | Yes | No  | —  | —  | —  | —  | —  | —  | —  | —  | —  |
|        | Albolabrin  | Yes | —  | —  | —  | —  | —  | —  | —  | —  | —  | —  |
| G 6    | Disintegrin VLO4 | —  | —  | Yes | Yes | No  | —  | —  | —  | —  | —  | —  |
|        | Labein      | Yes | —  | —  | Yes | Yes | —  | —  | Yes | —  | —  | —  |
|        | Disintegrin VB7A | —  | —  | Yes | Yes | No  | —  | —  | —  | No | —  | —  |
|        | Acostatin   | Yes | —  | —  | —  | —  | —  | —  | —  | —  | —  | —  |
|        | Schistatin  | Yes | Yes | —  | —  | —  | —  | —  | —  | —  | —  | —  |
|        | disintegrin from Saw-scaled viper | Yes | Yes | —  | —  | Yes | —  | —  | —  | —  | —  | —  |
|        | Disintegrin EO4A | —  | —  | Yes | Yes | No  | —  | —  | —  | —  | —  | —  |
| G 7A   | Lubestatin  | —  | —  | —  | —  | —  | —  | —  | —  | —  | —  | —  |
|        | Viperostatin | —  | —  | —  | —  | —  | —  | —  | —  | —  | —  | —  |
|        | Jerdostatin | —  | —  | —  | —  | —  | —  | —  | —  | —  | —  | —  |
|        | Obustatin   | —  | —  | —  | —  | —  | —  | —  | —  | —  | —  | —  |
|        | Obustatin   | —  | —  | —  | —  | —  | —  | —  | —  | —  | —  | —  |

(Continued on following page)
TABLE 3 | (Continued) Affinity and interactions with integrins to each group of disintegrins according to the classification proposed.

| Groups | Disintegrin | Integrins | REF |
|--------|-------------|-----------|-----|
|        |             | α1b | α2 | α5 | α6 | α3 | α4 | α9 | α1 | α2 | α7 | αv | α4b | αv | αM |
| G 7B   | Multisquamatin | Yes | – | – | – | – | – | – | – | – | – | – | – | – | – | Tekha et al. (1994); Okuda et al. (2001) |
|        | Echistatin   | Yes | Yes | – | – | – | – | – | Yes | – | – | – | – | – | – | McLane et al. (1996); Surazytski et al. (2005); Szabo et al. (2012); Jing and Sun (2019) |

Yes – disintegrin interact with integrin, No – disintegrin not interact with integrin, * – not interact with the alpha portion.

**Functional Diversification of Disintegrins Motifs and Bases of Integrin-Modulation**

Functionally, the evolution of the disintegrin family is influenced by positive Darwinian selection, guiding the adaptation of the conformation loop and C-terminal to the target integrin receptor (Juárez et al., 2008). The evolution of integrin-binding motif on disintegrins emerges from an ancestral Arg-Gly-Asp (RGD) sequence, according to phylogenetic and codon substitution studies, to a panel of integrin receptors targets (Juárez et al., 2008). RGD sequence emerged from a subgroup of PIII-SVMP with RDECD sequence. Minimum mutations accomplish the conversion of RDE to RGD sequence, and three mutations (DNA) were the minimum codon changes required to the emergence of inhibitory integrin motifs (Juárez et al., 2008). Dimeric disintegrins present different integrin-binding motifs, suggesting fast evolution and cumulative structure changes (Calvete et al., 2003). Disintegrin inhibitory loop motifs are represented by RGD, capable of modulating integrins such as α1bβ3, αvβ3, and α5β1. Variations of RGD motif include KGD, VGD, MGD, and WGD. KTS and RTS selectively modulate α1β1; MLD motifs, capable of modulating integrins such as α9β1, α4β1 and α4β7 (Table 2); (Sanz et al., 2006). Most RGD-disintegrins are monomeric proteins and this motif is present in all homo-dimeric and some subunits of hetero-dimers disintegrins (Walsh and Marcinkiewicz, 2011).

Different structural features of disintegrins define the interaction with the integrin receptors. Conserved aspartate within the disintegrin motif might be responsible for binding to integrin receptors, while the specificity is ruled by the other two residues within the disintegrin motif. Further, the residues flanking the tripeptide motif, the conformation of the motif loop, and the C-terminal of the disintegrin sequence define the integrin-binding and selectivity (Calvete, 2005a). This characteristic modulates the selectivity of disintegrins to different integrins, as represented in Table 3, which shows the selectivity interactions of disintegrins, divided according to each group of our classification, through different receptors, indicating the affinity of each member.

Moreover, the modulatory activity of disintegrins depends on the appropriate pairing of cysteine residues, contributing to the conformation of the disintegrin structure (Niewiarowski et al., 1994). The conservative aspect of cysteine residues and disulfide bond pattern among disintegrins contribute to the hypothesis of strong selection to maintain the active conformation of these proteins (Juárez et al., 2008).

**Studies of the Structure-activity Relationship**

Integrins recognize many physiological ligands, including soluble and surface proteins. X-ray crystallographic structures of the extracellular domains of αvβ3 have provided insights into the integrin structure-function relationship (Xiong et al., 2002; Van Agthoven et al., 2014). Disintegrins containing RGD or KGD motifs have been reported as unique and potentially useful tools to investigate integrin-ligand interactions. This is because these motifs serve as an integrin ligand-binding site, through which it plays a key role in interacting with integrin receptors. However, these studies are still scarce, for example, there is no report of a structure obtained by crystallography of a disintegrin-integrin complex, most of the studies were conducted by docking analysis.

In general, the RGD motif and the C-terminal of disintegrins have been studied to modulate the Integrin-Disintegrin interaction. To illustrate, the molecular docking of rhodostomin into Integrin α1bβ3, reveal a series of interactions between the Arg residue of the disintegrin with Asp 224 of the Integrin by salt bridges and with Tyr189 and Ser255 by hydrogen bonds. Also, the Asp residue of the RGD loop demonstrates to interact with the Ser123 of the β3 subunit (Chang et al., 2017). Likewise, the C-terminus of echistatin showed to interact with integrin αvβ3. The molecular docking of the echistatin into the integrin αvβ3 showed that the M28 of the side chain may interact with the D126 of the β3 subunit and the H44 of the C-terminus and the K45 of the C-terminus may interact with the side chain of M180 of the β3 subunit (Chen et al., 2020). In this regard, variations in the C-terminal region of disintegrins may be a requirement for the recognition of different integrins. Otherwise, the hydrophobic residues of flavoridin and kistrin dictate the specificity for the α1bβ3 integrin. Also, analyses of the trimestatin reveal that the residues Pro53 and Trp67 may be important for the recognition of the β3 subunit (Bilgrami et al., 2004).
Up to appoint, the interactions between the RGD loop and the C-terminus of the disintegrins play a key role in integrin recognition. Flavoridin shows the contact of the RGD loop and the C terminal domains, involved between the residues Cys27, Ala28, Asp29 and residues Gly7, Asn11, Cys13, Leu21 and the other between Cys13, Ala25, and Cys26. Further, the studies of Senn 1993 found a connection of the C-terminal of the molecule in the residues Cys64 to Trp67 to the loop containing the RGD sequence, which suggests a role of the C-terminal in the recognition and interaction of the disintegrin with the integrin (Senn and Werner, 1993). Many studies have shown that the residues that flank the RGD motif and in the C-terminal region of the disintegrins affect their specificities and binding affinities to integrins (Dennis et al., 1993; Marcinkiewicz et al., 1999; Chen et al., 2009; Cheng et al., 2012).

Finally, the study of the interaction of jararacín and jarastatin with αIIβ3 integrin by docking reveals, in the first place, that jararacín has more interactions with this integrin. The RGD motif of the jararacín interacts with both subunits of the integrin, making hydrogen bonds and ionic interactions, and indicates that the N-terminus and C-terminus region interact with both subunits of the integrin too. In the same way, the docking complex of the jarastatin showed hydrogen and ionic interactions and interactions of the N-terminus and C-terminus with both parts of the integrin (Wermelinger et al., 2009).

### Mutants in the Study of the Structure-Activity Relationship

Several studies have shown that the amino acid residues that flank the RGD motif and the C-terminal region of the disintegrins modulate their specificity of interaction with integrin complexes (Dennis et al., 1993; Scarborough et al., 1993; McLane et al., 1998; Rahman et al., 1998; Cheng et al., 2012; Shiu et al., 2012). For example, disintegrins that have an ARGDW sequence showed a greater affinity for binding to the αIIβ3 integrin, while disintegrins with an ARGDN sequence preferentially bind to αvβ3 and α5β1 integrins (Scarborough et al., 1993). To a better understanding of these interactions, some mutants were produced over the years. To investigate the structural basis of the integrin inhibitory potency, some studies have made mutations in the amino acid residues of the interaction loop of the disintegrins and analyzed the effects produced on the structure and activity of the disintegrins by replacing the native motif (Chen et al., 2009; Carbajo et al., 2011; Calvanese et al., 2015; Chang et al., 2017). These studies have revealed that mutations in the disintegrin that change the aspartate to glutamate, in the RGD loop, decrease their activity, as for Kistrin, which showed a 100-fold decrease in activity with this mutation (Dennis et al., 1993). Likewise, the rhodostomin (Rho) D51E mutant (2PJG/2PJF) was 1,000 times less active than Rho in inhibiting integrins.

A powerful tool to investigate this is NMR spectroscopy and molecular docking has been used as valuable tools to study the relationship between the structure, dynamics, and function of the mutant strains of the rhodostomin protein (2PJF) (Table 2) to understand the important structural requirements for the recognition of integrins. The structural study carried out by Chen et al. (2009), found that Rho and its mutant have the same tertiary fold with three double-stranded antiparallel beta-sheets. There are no structural differences between the RG [D/E] loop. Two small differences between Rho and its mutant D51E were found only in its backbone dynamics and 3D structures. The relaxation parameter R value of E51 is 13% higher than that of residue D51. A difference in charge separation of 1.76 A was found between the positive (R49) and negative (D51 or E51) side chains.

The coupling of Rho to the αvβ3 integrin, by molecular docking, showed that the amide and carbonyl groups of the main structure of the amino acid D51 of Rho formed hydrogen bonds with the amino acids R216 and R214 of the integrin. However, this hydrogen bond does not exist in the structure of the complex formed by the mutant protein D51E and the integrin. Therefore, the study suggests that the hydrogen bonds between the side chain and the backbone of the D51 residue of Rho and integrin are important for their binding to integrin (Chen et al., 2009). In addition, the Rho mutant P48A (2PJH) has its inhibitory capacity to αvβ1 integrin increased by 4.4 times. Docking of P48A showed no difference in the structure of the complex with αvβ1 integrin, pointing out the importance of the dynamics, especially of the RGD loop. The mutant P48A caused differences in the order parameter (S2), conformational exchange contribution (R_e), and local correlation time (τ_r). The authors showed that the thermal flexibility (S2), which are motions in the pico to nanoseconds timescale, are increased for residues R49, G50, and D51 in the mutant P48A (Shiu et al., 2012).

The Rho mutant G50L (2LJV) is a disintegrin that specifically binds to αvβ3 integrin. According to Shiu et al. (2012), the docking models of the mutant G50L and integrin showed that the amino acid L50 mutant G50L can be accommodated by a cavity within the interface between the αv and β3 subunits of the αvβ3 integrin. In contrast, such a pocket is not found in αIIβ3 integrin, due to the formation of hydrogen bonds between αIIβ3 residue Y190 and residue R216 of β3 subunits, resulting in blocking the bonding of residue L50. It was also observed that the G50L mutation increased the rigidity of the RLD motif, and the adjacent residues exhibited a slow conformational exchange. This finding shows that the slow movements of the RLD motif also play a vital role in modulating the integrin recognition link (Chuang et al., 2012; Shiu et al., 2012).

The ARLDDL mutant (3UCI), a potent and selective αvβ3 integrin antagonist, was designed to investigate the function-structure-dynamic relationship. The 3D structure of the ARLDDL mutant was determined by X-ray crystallography, and its tertiary fold is the same as the reported disintegrin structures. The only difference found in the RLD motif of the ARLDDL loop was a compact β-turn structure with a distance of 5.5 Å between R49 (Ca) and D52 (Ca) compared to those of disintegrins ranging from 6.8 to 8.4 Å (Shiu, 2011; Shiu et al., 2012).

In the study by Chang et al. (2017), it was observed that the content of the sequence of the RGD loop and the C terminal of the disintegrins mutually affected their conformations, resulting in
functional and structural differences in the integration of the integrin. Structural analysis by NMR showed that Rho mutants containing a 46ARGDWN-46PRNPWNG sequence exhibited the highest selectivity in inhibiting cell adhesion mediated by αIIbβ3 integrin. The results, of molecular docking, suggested that the content of the sequence and the length of the C-terminal regions in the disintegrins are critical to their ability to bind to the αIIbβ3 integrin (Chang et al., 2017).

Carbajo et al. (2011) made changes to the structure of jerdostatin by replacing the native RTS motif with KTS. These authors demonstrated by NMR that wild-type jerdostatin and its mutant R24K present a common structure, but different dynamic profiles. They found differences in movements on the picosecond to nanosecond time scale and deceleration movements for some residues of the R24K mutant compared to the wild type of jerdostatin (Carbajo et al., 2011). According to the authors, these findings may explain the reduction in the inhibitory potency of the integrin of the mutant jerdostatin R24K (IC₅₀ 703 nM) compared to the wild type (IC₅₀ 180 nM) (Carbajo et al., 2011). A complementary study by Calvanese et al. (2015), used Molecular Dynamics (MD) simulations of the two molecules (wild type jerdostatin and it is mutant) to explore in atomic resolution the structural bases of their different dynamic behaviors, to identify the atomic movements that could differentiate their behaviors dynamic and, therefore, their properties/activities. The analysis confirmed the combined movements between the recognition loop and the C-terminal tail considered relevant to the functional capacity of jerdostatin. It also revealed the residues that dominate such a mechanism. Indicating that both wild-type jerdostatin and the R24K mutant share a common structure but differ in global movements. Studies like this are important to clarify whether disintegrins movements can be functional for integrin binding (Sanz-Soler et al., 2012; Calvanese et al., 2015).

This kind of study helped the development of pharmacological agents that block platelet aggregation by inhibition of integrin αIIbβ3. Some commercialized drugs are the epifibatide, a cyclic heptapeptide that originated from the disintegrin barbourin, and the tirofiban, which originated from the disintegrin echistatin. The barbourin is a disintegrin isolated from the venom of Sistrurus miliarius barbouri, which possesses a KGD loop and a high affinity to integrin αIIbβ3 (Tcheng and O’Shea, 2002). Jing and Lu 2005, investigated a mutant proinsulin chimera with eight amino acids from barbourin (CAGGDWNC, respectively). Interestingly, they found that the protein inhibits human platelet aggregation induced by ADP and retains its binding activity to the insulin receptor. Also, Xiao et al., 2004, showed how the epitibatide fits into the Integrin αIIbβ3, demonstrating some hydrophobic contacts with the Lys, based on the binding loop of barbourin, with Phe231 in the αm, β-propeller. In addition, the Asp224 of the residue αIIb may form hydrogen bonds with the epitibatide.

Otherwise, interaction studies between tirofiban (based on echistatin) showed that the co-crystal between the drug and the Integrin αIIbβ3 reveals that the sulfonamide groups (in the tirofiban) interacted with Tyr166 and Arg214 by hydrogen bonds in the β3 subunit and the butyl and pyridyl groups may interact with the Phe160 and Tyr190 in the αIIb subunit by hydrogen bond interaction (Xiao et al., 2004). Interestingly, the interaction with free integrin reveals poor changes in integrin αIIbβ3 structure, changing the globular conformation of the free Integrin to an open conformation in just 4% of the free integrins, seen by transmission electron microscopy (Hantgan, et al., 2002).

**CONCLUSION AND FUTURE DIRECTIONS**

To conclude, disintegrins are divided into five different groups according to their polypeptide length and the number of disulfide bonds. In this work, we propose a classification based on patterns of disulfide bonds, where this classification resulted in the division into seven groups, organized by disulfide bond pattern. Through comparing amino acid sequences by multiple-sequence alignment, a brief phylogenetic analysis, and extensive literature review, we support the view that the different disintegrin subfamilies evolved from a common ADAM (a disintegrin and metalloproteinase-like) and that structural diversification occurred through disulfide bond engineering.

A deep analysis of the conserved cysteine residues in each disintegrin subfamily (Figure 4) strongly indicates that structural diversity of disintegrins was achieved during evolution through selective loss of disulfide bonds. It is well-known that disulfide bonds play a key role in the stability and impose distinct protein folding, with a specific orientation of the loop regions of these proteins. Recognition of integrins by disintegrins is mediated by the specific binding loop of each disintegrin. As well that, the residues flanking the tripeptide motif, the motif loop conformation, and the C-terminus of the disintegrin sequence influence integrin binding and selectivity (Calvete, 2005a). The C-terminal region of disintegrins and the flexibility of the RGD loop (located at the apex of the loop), may be a requirement for the recognition of different integrins, specifically for the β subunit. These two parts mutually affected their conformations, making it crucial to the Integrin modulation. This characteristic modulates the selectivity of disintegrins to different integrins, as mentioned in this work. Furthermore, the modulatory activity of disintegrins depends on the proper pairing of cysteine residues, contributing to the conformation of the disintegrin structure (Niewiarowski et al., 1994). The conservative aspect of cysteine residues and the disulfide bond pattern between disintegrins contribute to the hypothesis of strong selection to maintain the active conformation of these proteins (Juárez et al., 2008).

Finally, few structural studies show the interaction between disintegrins and integrins. In our work, the selective interactions of disintegrins are shown, divided according to each group of our classification, through different receptors, indicating the affinity of each member to different integrins, exhibit the importance of the structural studies to the comprehension of the interaction of the Disintegrins to its receptor targets. This knowledge is fundamental for the designing of new drugs that target integrins, as it is shown that disulfide arrangements of disintegrins have an impact on integrin/disintegrin interaction. Also, it is important to point out that only a small percentage of
the available disintegrins from venoms has been investigated so far and these molecules present opportunities for larger Integrin engagement surface with good stability, increasing the opportunity for the development of new drugs (Trim et al., 2021).

**AUTHOR CONTRIBUTIONS**

Conceptualization, RZ and FA; resources, RZ and FA; writing—original draft preparation, AV, JE, VD, and LW; writing—review and editing, AV, JE, and VD; supervision, LW, RZ and FA. All authors have read and agreed to the published version of the manuscript.

**FUNDING**

This work was supported by FAPERJ (grant no. 202.829/2017; 010.003016/2014 255.940/2020; 202.279/2018; 239.229/2018; 210.361/2015 and 204.432/2014), CNPq (grant no. 308376/2020-0; 309564/2017-4). Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, CAPES. FAPERJ scholarship was granted to AV and VD. CNPq scholarship was granted to JE.
Okuda, D., Nozaki, C., Sekiya, F., and Morita, T. (2001). Comparative Biochemistry of Disintegrins Isolated from Snake Venom: Consideration of the Taxonomy and Geographical Distribution of Snakes in the Genus Echis. *J. Biochem.* 129, 615–620. doi:10.1093/oxfordjournals.jchem.a020298

Olfa, K.-Z., José, L., Salma, D., Amine, R., Najet, S. A., Nicolas, A., et al. (2005). Lebestatin, a Disintegrin from Macrovipera Venom, Inhibits Integrin-Mediated Cell Adhesion, Migration and Angiogenesis. *Lab. Invest.* 85, 1507–1516. doi:10.1038/lab.2005035

Park, D., Kang, I., Chung, K., Kim, D. S., and Yun, Y. (1998). Cloning and Characterization of Novel Disintegrins from Agkistrodon Halys Venom. *Mol. Cell.* 8, 578–584.

Pfaff, M., Mclane, M. A., Bevligia, L., Niewiarowski, S., and Timpl, R. (1994). Comparison of Disintegrins with Limited Variation in the RGD Loop in Their Binding to Purified Integrins αβ3β3, αβ3 and αβ1 and in Cell Adhesion Inhibition. *Cel Adhes. Commun.* 2, 491–501. doi:10.1519/1541960940914213

Phillips, D., Jennings, L., and Edwards, H. (1980). Identification of the Integrin Subunit α5β1 and Its Interaction with α5β1 Integrin Complexes. *Biochem. J.* 135, 247–257. doi:10.1042/bj3350247

Sanz, L., Bazza, A., Marrakchi, N., Pérez, A., Chenik, M., Bel Lasfer, Z., et al. (2006). Molecular Cloning of Disintegrins from Cerasus vipera and *Macroviroma Lebetina* Transmammeana Venom Gland cDNA Libraries: Insight into the Evolution of the Snake Venom Integrin-Inhibition System. *Biochem. J.* 395, 385–392. doi:10.1042/BJ20051678

Sanz, L., Chen, R.-Q., Pérez, A., Hilario, R., Juárez, P., Marcinikiewicz, C., et al. (2005). cDNA Cloning and Functional Expression of Jerdostatin, a Novel RTS-Disintegrin from Trimeresurus Jerdoni and a Specific Antagonist of the α5β1 Integrin. *J. Biol. Chem.* 280, 40714–40722. doi:10.1074/jbc.M509738200

Sanz-Soler, R., Lorente, C., Company, B., Sanz, L., Juárez, P., Pérez, A., et al. (2012). Recombinant Expression of Mutants of the Frankenstein Disintegrin, R特斯-Occultatius. Evidence for the Independent Origin of RGD and KTS/RTS Disintegrins. *Toxicon* 60, 665–675. doi:10.1016/j.toxicon.2012.05.010

Saviola, A. J., Burns, P. D., Mukherjee, A. K., and Mackessy, S. P. (2016). The Evolution of the Snake Venom Integrin-Inhibition System. *Front. Immunol.* 8, 578. doi:10.3389/fimmu.2017.00528

Scarborough, R. M., Rose, J. W., Naughton, M. A., Phillips, D. R., Nannizzi, L., Stahle, M., Jerome, G., Nagaswami, C., Weisel, B. R., and Markland, F. S. (2002). Characterization of the Integrin Species in Chinese Hamster Ovary Cells: Mutants and Variants: Insight into Their Interactions with Integrins. [In Dissertation]. [Tainan (TW)]: National Cheng Kung University of Medicine.

Sinay, E. J., Holden, A., Casselden, E., Bulstrode, H., Thomas, G. J., and Willaime-Morawek, S. (2017). Metalloproteinases ADAM10 and ADAM17 Mediate Migration and Differentiation in Glioblastoma Sphere-Forming Cells. *Mol. Neurobiol.* 54, 3893–3905. doi:10.1007/s12035-016-0053-6

Smith, K. J., Jasea, M., Lu, X., Williams, J. A., Hyde, E. I., and Trayer, I. P. (1996). Three-Dimensional Structure of the RGD-Containing Snake Toxin Albolabrin in Solution, Based on 1H NMR Spectroscopy and Simulated Annealing Calculations. *Int. J. Pept. Protein Res.* 48, 220–228. doi:10.1111/j.1399-3011.1996.tb08035.x

Sohn, Y.-D., Cho, K.-S., Sun, S.-A., Sun, H.-J., Kwak, K.-W., Hong, S.-Y., et al. (2008). Suppressive Effect and Mechanism of Saxatilin, a Disintegrin from Korean Snake (Gloydis Saxatilis), in Vascular Smooth Muscle Cells. *Toxicon* 52, 474–480. doi:10.1016/j.toxicon.2008.06.020

Stahle, M., Jerome, G., Nagaswami, C., Weisel, B. R., and Hanigán, R. (2002). TyrRoid Blocks Platelet Adhesion to Fibrin with Minimal Perturbation of GpIIb/IIIa Structure. *Thromb. Haemost.* 87, 910–917. doi:10.1055/s-0037-1613104

Staniszewska, I., Walsh, E. M., Rothman, V. L., Gaathon, A., Tuszynski, G. P., Calvete, J. J., et al. (2009). Effect of VP12 and Viperin on Inhibition of Collagen Receptors: Dependent Melanoma Metastasis. *Cancer Biol. Ther.* 8, 1507–1516. doi:10.1016/cb.2015.10.0000008

Suurzyński, A., Sienkiewicz, P., Wólczyński, S., and Palka, J. (2005). Differential Effects of Echistatin and Thrombin on Collagen Production and Prolidase Activity in Human Dermal Fibroblasts and Their Possible Implication in β1-integrin-mediated Signaling. *Pharmacol. Res.* 51, 217–221. doi:10.1016/j.phrs.2004.08.004

Szabo, A. M., Howell, N. R., Pellegrini, P., Gregurić, I., and Katsifis, A. (2012). Development and Validation of Competition Binding Assays for Affinity to the Extracellular Matrix Receptors, αβ3 and αβδβ3 Integrin. *Anal. Biochem.* 423, 70–77. doi:10.1016/j.ab.2011.12.046

Takegaki, J., Betrie, B. M., Walz, T., and Springer, T. A. (2002). Global Conformational Rearrangements in Integrin Extracellular Domains in Outside-In and Inside-Out Signaling. *Cell* 110, 599–611. doi:10.1016/S0092-8674(02)00935-2

Takeda, S., Takeya, H., and Iwanaga, S. (2012). Snake Venom Metalloproteinases: Structure, Function and Relevance to the Mammalian ADAM/ADAMTS Family Proteins. *Biochim. Biophys. Acta (Bba) - Proteins Proteomics* 1824, 164–176. doi:10.1016/j.bbapap.2011.04.009

Takeya, H., Nishida, S., Miyata, T., Kawada, S., Saisaka, Y., Morita, T., et al. (1992). Coagulation Factor X Activating Enzyme from Russell’s viper venom (RVV-X). A Novel Metalloproteinase with Disintegrin (Platelet Aggregation Inhibitor)-like and C-Type Lectin-Like Domains. *J. Biol. Chem.* 267, 14109–14117. doi:10.1016/S0021-9954(19)49685-3

Tcheng, J. E., and O’Shea, J. C. (2002). Epitibatide: A Potent Inhibitor of the Platelet Receptor Integrin Glycoprotein IIb/IIIa. *Expert Opin. Pharmacother.* 3, 1199–1210. doi:10.1517/14656566.3.8.1199

Trikha, M., Rote, W. E., Manley, P. L., Lucchesi, B. R., and Markland, F. S. (1994). Purification and Characterization of Platelet Aggregation Inhibitors from Snake Venoms. *Thromb. Res.* 73, 39–52. doi:10.1016/0049-3848(94)90052-3

Trim, C. M., Byrne, L. J., and Trim, S. A. (2021). Utilisation of Compounds from Venoms in Drug Discovery. 1st ed. Amsterdam: Elsevier B.V, 1–66. doi:10.1016/bs.pmch.2021.01.001

Tselepis, V. H., Green, L. J., and Humphries, M. J. (1997). An RGD to LDV Motif Conversion within the Disintegrin Kistinin Generates an Integrin Antagonist That Retains Potency but Exhibits Altered Receptor Specificity. *J. Biol. Chem.* 272, 21341–21348. doi:10.1074/jbc.272.34.21341

Tur-Fu, H., Chao-Zong, L., Chao-Ho, O., and Che-Ming, T. (1991). Halyasin, an Antiplatelet Arg-Gly-Asp-Containing Snake Venom Peptide, as Fibrinogen Receptor Antagonist. *Biochem. Pharmacol.* 42, 1209–1219. doi:10.1016/0006-2952(91)90256-5
Uzair, B., Atlas, N., Malik, S. B., Jamil, N., Ojuolape, S. T., Rehman, M. U., et al. (2018). Snake Venom as an Effective Tool Against Colorectal Cancer. *Pp* 25, 626–632. doi:10.2174/0929866525666180614112935
Van Aghoven, J. F., Xiong, J.-P., Alonso, J. L., Rui, X., Adair, B. D., Goodman, S. L., et al. (2014). Structural Basis for Pure Antagonism of Integrin αVβ3 by a High-Affinity Form of Fibronectin. *Nat. Struct. Mol. Biol.* 21, 383–388. doi:10.1038/nsmb.2797
Walsh, E. M., and Marcinkiewicz, C. (2011). Non-RGD-Containing Snake Venom Disintegrins, Functional and Structural Relations. *Toxicon* 58, 355–362. doi:10.1016/j.toxicon.2011.07.004
Wermelinger, L. S., Geraldó, R. B., Frattani, F. S., Rodrigues, C. R., Juliano, M. A., Castro, H. C., et al. (2009). Integrin Inhibitors from Snake Venom: Exploring the Relationship between the Structure and Activity of RGD-Peptides. *Arch. Biochem. Biophys.* 482, 25–32. doi:10.1016/j.abb.2008.11.023
Wu, P.-L., Lin, C.-C., Lin, T.-H., Lee, M.-S., and Wu, W.-G. (2016). Distal M Domain of Cobra ADAM-like Metalloproteinase Mediates the Binding of Positively Charged Cysteine-Rich Domain to αvβ3 Integrin in the Suppression of Cell Migration. *Toxicon* 118, 1–12. doi:10.1016/j.toxicon.2016.04.034
Wu, W. B., Chang, S. C., Liu, M.-Y., and Huang, T.-F. (2001). Purification, Molecular Cloning and Mechanism of Action of Graminellysin I, a Snake-Venom-Derived Metalloproteinase That Induces Apoptosis of Human Endothelial Cells. *Biochem. J.* 357, 719–728. doi:10.1042/bj3570719
Xia, X., Ma, Y., Xue, S., Wang, A., Tao, J., Zhao, Y., et al. (2013). Cloning and Molecular Characterization of BumaMPs1, a Novel Metalloproteinases from the Venom of Scorpion Buthus Martensi Karsch. *Toxicon* 76, 234–238. doi:10.1016/j.toxicon.2013.10.006
Xiao, T., Takagi, J., Coller, B. S., Wang, J.-H., and Springer, T. A. (2004). Structural Basis for Allostery in Integrins and Binding to Fibrinogen-Mimetic Therapeutics. *Nature* 432, 59–67. doi:10.1038/nature03976
Xiong, J.-P., Stehle, T., Diefenbach, B., Zhang, R., Dunker, R., Scott, D. L., et al. (2001). Crystal Structure of the Extracellular Segment of Integrin Alpha Vbeta3. *Science* 294, 339–345. doi:10.1126/science.1064535
Yang, R.-S., Tang, C.-H., Chuang, W.-J., Huang, T.-H., Peng, H.-C., Huang, T.-F., et al. (2005). Inhibition of Tumor Formation by Snake Venom Disintegrin. *Toxicon* 45, 661–669. doi:10.1016/j.toxicon.2005.01.013
Yeh, C.-H., Peng, H.-C., Yang, R.-S., and Huang, T.-F. (2001). Rhodostomin, A Snake Venom Disintegrin, Inhibits Angiogenesis Elicited by Basic Fibroblast Growth Factor and Suppresses Tumor Growth by a Selective αvβ3 Blockade of Endothelial Cells. *Mol. Pharmacol.* 59, 1333–1342. doi:10.1124/mol.59.5.1333

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher’s Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Vasconcelos, Estrada, David, Wermelinger, Almeida and Zingali. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.