CellPress

Feature Review

Discoidin Domains as Emerging Therapeutic Targets

Bruno O. Villoutreix[1,2,*] and Maria A. Miteva[1,2]

Discoidin (DS) domains are found in eukaryotic and prokaryotic extracellular and transmembrane multidomain proteins. These small domains play different functional roles and can interact with phospholipids, glycans, and proteins, including collagens. DS domain-containing proteins are often involved in cellular adhesion, migration, proliferation, and matrix-remodeling events, while some play a major role in blood coagulation. Mutations in DS domains have been associated with various disease conditions. This review provides an update on the structure, function, and modulation of the DS domains, with a special emphasis on two circulating blood coagulation cofactors, factor V and factor VIII, and the transmembrane neuropilin receptors that have been targeted for inhibition by biologics and small chemical compounds.

Trends

DS domains can naturally interact with different types of molecules (glycans, proteins, lipids) and are present in several proteins important in health and in disease states including cancer and blood coagulation.

The DS fold is now well-characterized at the atomic level, this should accelerate structure-guided design of therapeutic molecules.

Designed DS binders (low molecular weight molecules, peptides, and proteins) are so far protein–protein interaction inhibitors or transient protein–membrane inhibitors. These molecular mechanisms are expected to have untapped potential in terms of therapeutic intervention. Small oral drugs acting on DS domains should therefore be truly novel.

Many DS domains have still not been targeted by biologics or small compounds; exploration of these proteins could open new avenues for the development of new treatments.

The DS Domains

DS domain (also known as F5_8 type C domain) is a module found in various eukaryotic and prokaryotic multidomain proteins as a single domain or as repeats. This domain was originally described in the protein discoidin 1, a lectin with some affinity for galactose that mediates cell adhesion and migration in the slime mold *Dictostelium discoideum* [1,2]. The domain is not directly related to the so-called C2 domain involved in targeting some proteins to the cell membranes (often requiring the presence of Ca^{2+}). DS domains have been detected in over 100 eukaryotic and over 300 prokaryotic extracellular and membrane proteins, including (i) the neuronal cell surface antigen A5 (renamed neuropilin; NRP, NRP1, and NRP2 in higher eukaryotes), a potential therapeutic target in cancer [3–10], and (ii) retinoschisin (RS1), a protein expressed in the retina, and mutations in the gene coding for RS1 have been described as responsible for X-linked juvenile retinoschisis [11]. DS domains are also found in (iii) Del-1 protein, an embryonic endothelial cell protein that binds to some integrins [12], (iv) tyrosine kinase receptor DDRs (DDR1 and DDR2), the DS domain receptor family that binds collagen molecules, the most abundant proteins in mammals and that regulates cell adhesion, proliferation, and extracellular matrix remodeling, with potential therapeutic target for several diseases [13–17], (v) endothelial and smooth muscle cell-derived neuropilin-like protein (ESDN), a potential therapeutic target for angiogenesis [18], and (vi) two blood coagulation cofactors, factor VIII (FVIII), the anti-hemophilic factor, and factor V (FV), which are crucial for the generation of the serine protease thrombin [19–22]. A Pfam search (December 2015, Pfam version 29 [23]) performed on the DS domain family (PF00754) identified 999 different architectures (domain organizations of proteins containing DS domains), 6622 sequences (80 sequences in human, 3043 sequences in eukaryotes) and 717 species. The DS domains comprise about 150 amino acids and have many different biological roles that are in part linked to their interacting partner molecules. These domains can bind to lipids (e.g., the cell membrane [20,24–26]), glycans (e.g., heparin-like molecules [27]) and proteins (including various types of collagens [14,28–30]). DS domain-containing proteins in human are essentially involved in cellular adhesion, migration, or aggregation events, mostly associated with organogenesis (vasculogenesis and angiogenesis) and other developmental processes, as well as in blood coagulation [1,2]. The pathological

[1]Université Paris Diderot, Sorbonne Paris Cité, Molécules Thérapeutiques In Silico, Paris, France
[2]Institut National de la Santé et de la Recherche Médicale (INSERM) Unité 973, Molécules Thérapeutiques In Silico, Paris, France

*Correspondence:
bruno.villoutreix@inserm.fr
(B.O. Villoutreix).

Box 1. An Overview of Blood Coagulation

Normal hemostasis is crucial to maintain blood fluidity, to protect against inappropriate clotting and bleeding, and to re-establish vascular continuity [21,22,138]. The plasma coagulation system consists of a cascade of enzyme activation events in which multidomain serine proteases activate proenzyme zymogens and pro-cofactors for the next step of the cascade via limited proteolysis, culminating in the formation of a fibrin clot at the site of vascular injury. The revised version of the blood coagulation cascade, known as the cell-based model of fibrin formation, is shown in Figure 3, main text [80]. Thrombin is a key enzyme that contributes to the formation of the clot, but it also has many other biological functions [139]. Among the proteins involved in blood coagulation, two crucial non-enzymatic cofactors, FV and FVIII, circulate as pro-cofactors. These two proteins are proteolytically activated to activated FV (FVa) and activated FVIII (FVIIIa). FVa and FVIIIa are homologous cofactors acting in two macromolecular complexes: the prothrombinase or PTase (containing: factor Xa or FXa, FVa, anionic membranes, and Ca^{2+}) and the intrinsic tenase or Xase (Factor IXa or FIXa, FVIIIa, Ca^{2+}, and anionic membranes). The prothrombinase complex catalyzes the conversion of the zymogen prothrombin to the active thrombin, whereas the intrinsic Xase activates FX to FXa. FXa and FIXa can catalyze their protein substrates in the absence of FVa and FVIIIa, but their enzymatic activity is too weak to prevent excessive blood loss following injury. However, in the presence of the appropriate membrane surface and of the cofactors (FVa and FVIIIa), the catalytic efficiency of FIXa and FXa increases significantly, providing the physiologic pathway for thrombin generation.

Box 2. NRPs–Roles in Signaling Pathways

NRP1 and NRP2 are transmembrane glycoproteins that contain two DS domains and function mainly as multifunctional coreceptors [3–10,62]. The proteins encoded by the NRP genes are essentially membrane-bound, although splice forms encoding soluble extracellular domains have also been identified. The extracellular region of NRPs is essential for the binding of two main types of unrelated ligands, the class 3 semaphorins (a family of axon guidance molecules that incorporates NRPs as coreceptors) and several growth factors such as VEGF or PDGF (see Figure 4 in main text). The CUB a1a2 domains of NRPs interact with the Sema domain of semaphorins, and the b1 domain interacts with the semaphorin integrin (PSI) and immunoglobulin (Ig)-like domains. The best-characterized interaction between NRPs and growth factors involves the VEGF family and more specifically VEGF-A (e.g., VEGF165). VEGF interacts essentially with several loops present on the b1 domain, although the b2 domain is necessary for optimal binding. This interaction has been analyzed at the atomic level via the crystal structure of a tetrapeptide mimetic of the VEGF tail, named Tuftsin, in complex with the NRP1 b1b2 domains [27]. Of importance, the b1b2 domains of NRPs interact with heparin and heparin binds directly to VEGF-A. The MAM domain is thought to play a role in oligomerization. The VEGF-A–NRP1 interaction is thought to be essential for cardiovascular development (e.g., vessels) while the semaphorin–NRP1 signaling axis is crucial for the formation of cranial and spinal nerve projections, among others. With regard to signaling via the VEGFRs, NRP1 acts as a coreceptor (for instance) for VEGFR2, and enhances VEGF binding and VEGFR phosphorylation, and thus seems important to promote endothelial cell proliferation, migration, and angiogenesis. The current view is that a tertiary complex is formed between VEGF, VEGFR, and NRP, with VEGF acting as bridge between NRP and VEGFR. Further, it has been suggested that NRP1 might display functions through other receptor tyrosine kinases in response to other growth factors. The implication of NRPs in the immune system is tissue-dependent and not yet fully understood, but it is known that interactions of NRP1 with transforming growth factor β (TGFβ) exerts immunosuppressive effects through Smad activation [110]. Inhibition of NRPs is expected to overcome the limitations of individually inhibiting the VEGF–VEGFR pathway in cancer therapy, and should provide new ideas for cancer treatment.

consequences of mutations or variations in human DS domain proteins or changes in their concentration levels further highlight the importance of this structural module in health and disease [2]. Together, it would seem interesting to design molecules that modulate DS domains with the aim of gaining new insights into their pathophysiological roles and to develop new therapies. The most recent structure–function review on the DS domain appeared in 2007 [2] and, at that time, very few studies had addressed the modulation of DS domains for therapeutic intervention. In the present review we briefly introduce the field and focus on the DS domains of FV, FVIII (Box 1), and neuropilins (Box 2) because they are the only DS domains that have been rationally modulated to date by both biologics and small chemical compounds.

Main Structural Features of the Discoidin Domains

The first 3D structural insights into the DS domains were reported for FVIII [1,31–33] and FV [1,34,35]. Tandem DS domains was discovered in the C-terminal region of FVIII and FV when the sequences of these two proteins were determined [20,24,25,36]. FV and FVIII have an identical domain organization, A1–A2–B–A3–C1–C2, and the two C domains (C1 and C2) are DS domains (Figure 1, Key Figure). The FV/FVIII C domains are important for protein–protein and protein–membrane interactions. The C2 domain is known to be crucial for transient binding