Crystallographic analysis of the laminin β2 short arm reveals how the LF domain is inserted into a regular array of LE domains

David Pulido, David C. Briggs, Jinwen Hua and Erhard Hohenester

Department of Life Sciences, *Imperial College London, UK*

Correspondence to Erhard Hohenester: *e.hohenester@imperial.ac.uk*
http://dx.doi.org/10.1016/j.matbio.2016.06.006

Abstract

Laminins are a major constituent of all basement membranes. The polymerisation of laminins at the cell surface is mediated by the three short arms of the cross-shaped laminin heterotrimer. The short arms contain repeats of laminin-type epidermal growth factor-like (LE) domains, interspersed with globular domains of unknown function. A single LF domain is inserted between LE5 and LE6 of the laminin β1 and β2 chains. We report the crystal structure at 1.85 Å resolution of the laminin β2 LE5-LF-LE6 region. The LF domain consists of a β-sandwich related to bacterial family 35 carbohydrate binding modules, and more distantly to the L4 domains present in the short arms of laminin α and γ chains. An α-helical region mediates the extensive interaction of the LF domain with LE5. The relative arrangement of LE5 and LE6 is very similar to that of consecutive LE domains in uninterrupted LE tandems. Fitting atomic models to a low-resolution structure of the first eight domains of the laminin β1 chain determined by small-angle X-ray scattering suggests a deviation from the regular LE array at the LE4–LE5 junction. These results advance our understanding of laminin structure.

Introduction

Heterotrimeric (αβγ) laminins are a major constituent of all basement membranes [1,2]. They are present in even the simplest animals and presumed to be essential for multicellular life [3,4]. In mammals, at least 16 laminins are assembled from one of five α chains, one of three β chains, and one of three γ chains [5]. The three chains associate through an α-helical coiled coil, the so-called long arm (80 nm length). At the N-terminus of the long arm, the three chains separate and form three distinct short arms (35–50 nm length), lending the laminin heterotrimers their characteristic cross-shaped appearance in electron micrographs (Fig. 1). In the α3A, α4, β3 and γ2 chains, the short arms are truncated or altogether absent. At the C-terminus of the long arm, the α chain continues for another ~1000 residues, which are folded into five consecutive laminin G-like (LG) domains. A major function of laminins is to form cell-associated polymers, which provide structural support as well as platforms for signalling [1,2]. Cell attachment is mediated by the LG domains and the very C-terminus of the coiled coil, whereas polymer formation is mediated by the tips of the three short arms [6].

The three short arms of the laminin heterotrimer are composed of long repeats of laminin-type epidermal growth factor-like (LE) domains, capped by laminin N-terminal (LN) domains that mediate the self-interaction of laminins. The LE repeats are interrupted once (β and γ chains) or twice (α chains) by globular domains of unknown function (Fig. 1). The inserted domains in the β and γ chains are called LF and L4, respectively [5] (or IVB and IVA in UniProt); the two types are not related at the sequence level. In the α1 and α2 chains, both inserts are L4 domains, whereas in the α3 and α5 chains only the second insert domain is an L4 domain. The first insert in α3 and α5 is much longer (~580 residues) and has weak homology to the LF domain in the first ~200 residues.

The LE fold was first described by Stetefeld et al. [7], who determined the crystal structure of LE domains 7–9 of the laminin γ1 chain (we follow the UniProt convention of numbering the LE domains consecutively from the N-terminus; Stetefeld et al. termed this region γ1III3–5). Subsequent structures of laminin short arm tips [8,9], netrins [10–12], and netrin G proteins [13,14] have confirmed that LE domains are largely devoid of secondary structure and are lacking a conventional hydrophobic core. Instead, the LE domain is held together by eight conserved cysteines, which are linked 1–3, 2–4, 5–6 and 7–8. This disulphide bonding pattern creates four loops: loop a (Cys1–Cys3), loop b (Cys2–Cys4), loop c (Cys5–Cys6), and loop d (Cys7–Cys8). The L4 domains in laminin α and γ chains are inserted between cysteines 3 and 4 of a standard 8-cysteine LE domain. The LF domains in laminin β chains, and the large inserts in α3 and α5 chains, are inserted after a truncated LE domain lacking the 7–8 disulphide bond (LE5 in β chains).

A recent crystal structure of the second L4 domain of the laminin α2 chain (α2 L4b) revealed an irregular β-sandwich similar to bacterial carbohydrate binding modules, ephrin-binding modules and MAM domains [15]. In the present study, we report that the LF domain of the laminin β2 chain has a similar fold, despite sequence identity of only ~10%, but that the LF domain additionally contains a unique α-helical region that makes extensive interactions with the preceding LE domain, LE5. Although separated by a 220-residue insert, LE5 and LE6 are observed in the rod-like arrangement that is typical of tandem LE domains. Thus, the globular LF domain in laminin β short arms is accommodated without interrupting the regular array of LE domains.

Results

Crystal structure of laminin β2 LE5-LF-LE6

We produced a panel of recombinant laminin fragments containing the β1 or β2 LF domain flanked by one or more LE domains for crystallisation trials. Crystals of the laminin β2 LE5-LF-LE6 fragment were suitable for structure determination. The β2 LE5-LF-LE6 structure was solved by the multiple isomorphous replacement method and refined to R_{free} = 0.210 at 1.85 Å resolution (Table 1). The asymmetric unit of the monoclinic crystals contains two structurally very similar copies of β2 LE5-LF-LE6 (r.m.s. deviation of 0.62 Å for 301 Cα atoms). The β2 LE5-LF-LE6 structure is complete except for residues 705–712 in the LF domain, which are presumed to be disordered.

The β2 LE5-LF-LE6 fragment has a surprisingly compact structure with approximate dimensions of 60 Å × 50 Å × 40 Å (Fig. 2A). The LE5 and LE6 domains are aligned with their long axes on one side of the structure. Even though LE5 and LE6 are separated by 220 residues in sequence, they interact similarly to consecutive LE domains in laminin short arms (see below for details). The 220 residues inserted between LE5 and LE6 are folded into two distinct regions, a 10-stranded β-sandwich (residues 564–726) and an α-helical region (727–783). To be consistent with the established nomenclature [5], we refer to the entire insert as the LF domain.

The β-sandwich of the LF domain consists of two antiparallel sheets of complicated topology, β1–(β2–β3)–β10–β5–β8 and β4–β9–β6–β7. The long central strand β10 interacts with β1 and with β3. A disulphide bond, Cys657–Cys685, links the β6–β7 and β8–β9 loops. Strong spherical electron density indicated that a metal ion is bound to the β1–β2 and β2–β3 loops. This ion was assigned as Ca^{2+} based on the nature of its ligands, the coordination geometry, and ion-ligand distances of 2.4–2.5 Å. The Ca^{2+} ion is coordinated by the side chains of Glu573 (monodentate), Glu575 (bidentate) and Asp719 (monodentate), as well as the carbonyl oxygen atoms of residues 598, 601 and 719; the resulting coordination geometry is a pentagonal bipyramid (Fig. 2B). Sequence comparison predicts that the Ca^{2+} ion is present in all LF domains (Fig. 2C).

The α-helical region of the LF domain consists of three helices held together by a disulphide bond, Cys752–Cys768, and a substantial hydrophobic core (Val727, Leu730, Met732, Phe733, Phe747, Leu771, Leu772, Ala775, Val779) that is conserved in other LF domains (Fig. 2C). Helix α3 plays an important role in mediating the apparently stable association of the LF domain with LE5: α3 residues Ser774 and Ser776 interact with loop a of LE5 (residues 528–531), and Gly782 and Val783 interact with loop c of LE5 (residues 547–550). In contrast, the LF domain makes only one tenuous contact with LE6, and there is a large solvent-filled cavity between the two domains (Fig. 2A). Altogether, the interface between the LF domain and LE domains 5 and 6 buries ~1600 $Å^2$ of solvent-accessible surface.

A search for related structures using PDBeFold [16] revealed that the β-sandwich of the LF domain is related to the family 35 carbohydrate binding module (CBM35) structure, despite pairwise sequence identities of just over 10% (Fig. 3); the α-helical region of the LF domain has no counterpart in CBM35 or any other protein. The galactose-binding CBM35 of a cell wall-degrading enzyme of *Clostridium thermocellum* [17] can be superimposed onto the LF domain with a r.m.s. deviation of 2.2 Å for 107 aligned Cα atoms (Z-score of 8.0 in PDBeFold), and even the Ca^{2+} site is conserved in the two domains. The loops connecting the strands of the β-sandwich