

◆トピックス◆

立体構造でみる凝固線溶蛋白質の作用機構

—EGF ドメイン—

神 田 大 輔*

EGF Domain

Daisuke KOHDA*

Key words : EGF domain, three-dimensional structure, crystallography, NMR

ドメインとモジュール

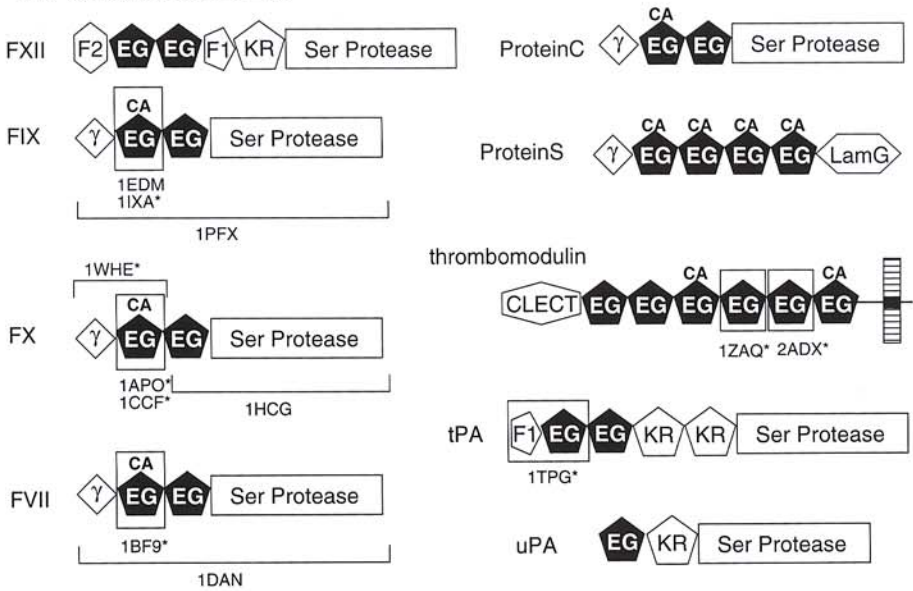
EGF ドメインは 30~40 残基のアミノ酸配列であり、非常に多くの蛋白質中に見いだされる蛋白質モジュールの一つである¹⁾。“モジュール”と言う言葉は“ドメイン”とともに、いろいろな場面で便利に使われるが、その定義はそれほど明確でない。蛋白質ドメインは、「蛋白質の立体構造上の割合独立したユニット」であることを意味し、それが一見無関係な蛋白質に広く見られるときに、「ポータブルな構造単位」と言う意味でモジュールと呼ぶことが多い²⁾。EGF ドメインはそのほとんどが動物由来の蛋白質、特に細胞外に存在する蛋白質中に、1,000 個以上見いだされている。しばしば、EGF ドメインは一次構造上連続して存在することが多く、中には 10 個以上も並んでいるものがある(図 1)。蛋白質モジュールの存在は、構造生物学的なアプローチを極めて容易にする。ある一つの EGF ドメインの立体構造を知ることができれば、その他の EGF ドメインの構造をホモロジーモデリングの手法を用いて推定することが可能となる。

上皮増殖因子の一次構造と立体構造

EGF ドメインの最初のアミノ酸配列は、もちろん、その名前の由来となった上皮増殖因子 (Epidermal Growth Factor, EGF) のものであり、1972 年に決定された。上皮増殖因子は細胞表面にあるレセプターに結合し、細胞の増殖を促す“蛋白質性の増殖因子”である³⁾。ヒトおよびマウスの上皮増殖因子はともに 53 残基であり、3 本の SS 結合を含んでいる。SS 結合の組み合わせは、図 1 の注に示すように、1-3, 2-4, 5-6 のパターンをとる。EGF ドメインの最初の立体構造例も実は上皮増殖因子であり、NMR (核磁気共鳴法) を用いて決定された (図 2A と B)。その後、NMR や X 線結晶解析を用いて現在までに、10 種類あまりの EGF ドメイン (あるいは EGF ドメインを含む蛋白質) の立体構造が決定されている。EGF ドメインは 40 残基程度と、かなり小さなモジュールであるが、立体構造を見るとさらに 2 つのサブドメインに分けることができる。分子全体の形がかなり細長いことに加えて、N 末端と C 末端は逆方向に

* 生物分子工学研究所・構造解析研究部門 (〒 565-0874 大阪府吹田市古江台 6-2-3)
Department of Structural Biology, Biomolecular Engineering Research Institute (6-2-3, Furuedai 6-chome, Suita, Osaka 565-0874, Japan.)

(A) 凝固系線溶系の蛋白質



(B) その他の蛋白質

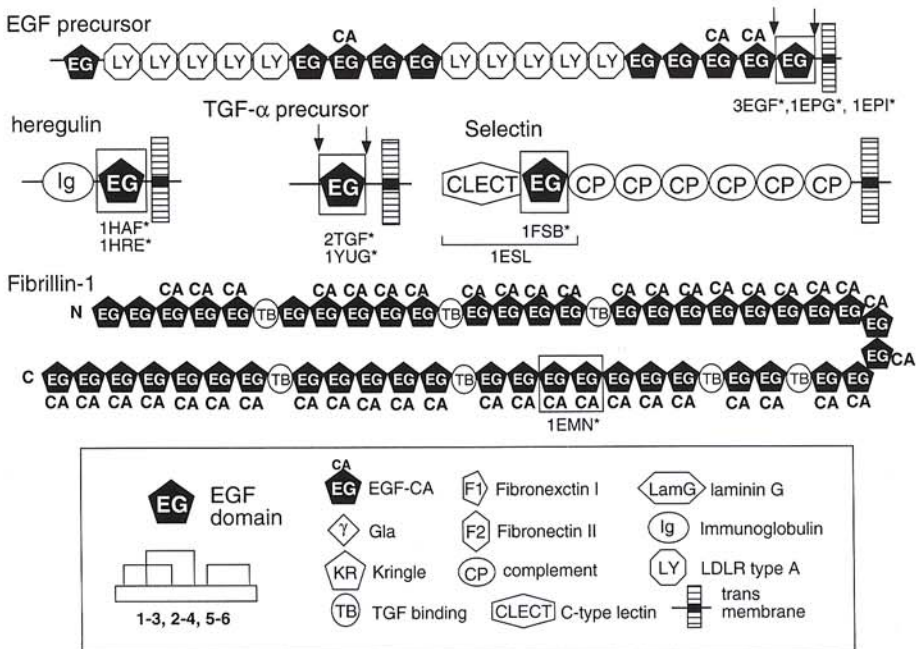


図 1 EGF ドメインを含む蛋白質のドメイン構成.

数字+アルファベット 3 文字は、プロテインデータベースの登録名。* はそれが NMR (核磁気共鳴法) によって決定されたことを示す。無印は X 線結晶解析による。プロテインデータベース (<http://www.pdb.bnl.gov>) へアクセスして、登録名を指定すると座標を得ることができる。↓ はプロテアーゼで切り出されることを示す。

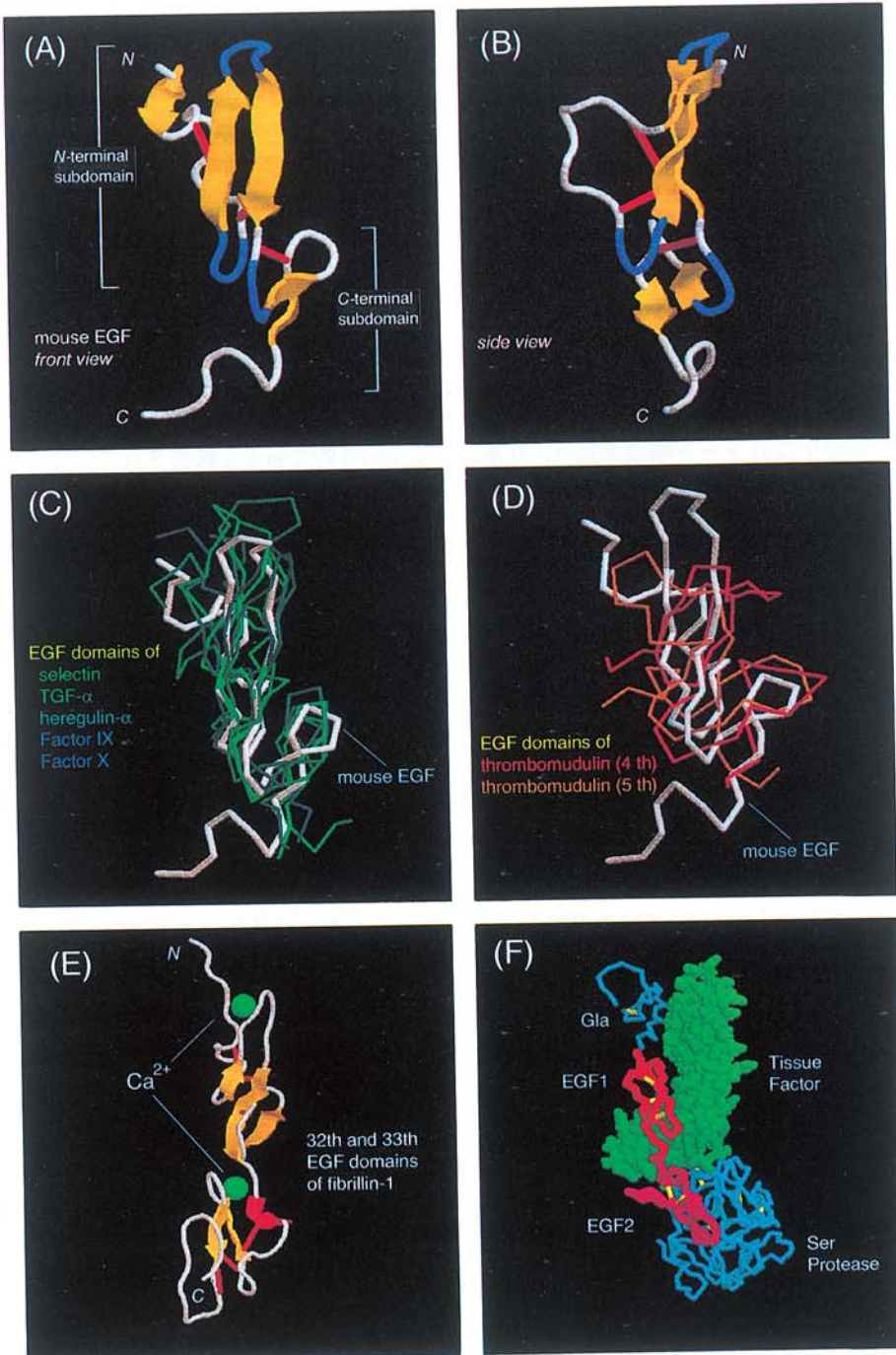


図 2 (A) と (B) はマウスの EGF の立体構造 (PDB 名は 1 EPG), (C) と (D) はマウスの EGF と他の EGF ドメインとの重ね合わせ, 6 個のシステインの α 炭素原子を重ね合わせた, セレクチン(1 FSB), TGF α (1 YUG), ヘレグリン α (1 HRE), IX 因子(1 IXA), X 因子 (1 APO), トロンボモジュリンの 4 番目 (1 ZAQ) と 5 番目 (2ADX) の EGF ドメイン. (E) はフィブリリンの 32 番目と 33 番目の連続した EGF ドメインの NMR 構造 (1 EMN). (F) 組織因子と VII 因子との複合体の X 線結晶構造 (1 DAN).

伸びているため、EGF ドメインの上流側にあるアミノ酸配列と下流側のアミノ酸配列は、空間的に最大限離れた位置に置かれることになる。したがって、EGF ドメインが連続して並ぶと、全体がロッド状のかたい構造体になることが予想される。特定の機能（後述）を持たない EGF ドメインの働きはこうした物理的なものであろう。たとえば、細胞膜から長く突き出して、機能ドメインを細胞膜から離して固定するのに役立つ。

N 端側サブドメインは2つの SS 結合で支えられた3本鎖逆平行 β シート構造を含み、芳香族残基に富む疎水性のコアが存在する。一方、C 端側サブドメインには1つの SS 結合が含まれている。小さな β シートが存在するが、短いため β シートとして記述されないことも多い。C 端側サブドメインの SS 結合をつくる、2つのシステインの間の残基数は8残基であることが多い。8残基という短い配列が2つのターン構造をとって、両端を SS 結合で保持されるため、構造上の自由度が少なく、堅い構造となっている（ダブルヘアピン構造と呼ぶことがある）。これを反映して、複数の EGF ドメイン構造を重ねたとき、C 端側サブドメインの SS 結合に挟まれたアミノ酸配列の重なりは特に良い（図 2C）。トロンボモジュリンの4番目の EGF ドメインはあまり重なりがよくないので図 2D に分けて示す。トロンボモジュリンの5番目の EGF ドメインは SS 結合のパターンが標準と異なっている（後述）ためさらに似ていない⁴⁾。

EGF ドメインの2つのサブドメインの間の相互作用はそれほど多くはない。2つのサブドメインの間の相対配置は固定されていないで、揺らぐことができる可能性がある。

EGF ドメインのコンセンサス配列

あるアミノ酸配列が EGF ドメインであるかどうかを判定することは、それほど容易ではない。EGF ドメインはいわゆるシステインリッチ

配列の一種であるが、システインリッチ配列（10種以上ある）では、システイン残基の保存性は高いが、その他のアミノ酸残基の保存性は低い。また、システインとシステインの間にあるアミノ酸残基の数がほぼ決まっているという特徴がある。これに対し、システイン残基間の SS 結合の組み合わせは高度に保存されている。したがって、SS 結合の組み合わせが判明すれば、かなりの確度でドメインの分類ができる。ただし、トロンボモジュリンの5番目の EGF ドメインの SS 結合の組み合わせは、1-2, 3-4, 5-6 であって例外的であると報告されている⁴⁾。

PROSITE (<http://expasy.hcuge.ch/sprot/prosite.html>) では、EGF ドメイン同定のために、C 端側のサブドメインに対応したアミノ酸配列のコンセンサスパターンが2つ提案されている。

EGF_1: C⁴-x-C⁵-x(5)-G-x(2)-C⁶

EGF_2: C⁴-x-C⁵-x(2)-[GP]-[FWY]-
x(4, 8)-C⁶

ただし、x (m, n) は m 個から n 個の任意のアミノ酸残基、x (n) = x (n, n)、x = x (1) であることを示す。[FWY] はフェニルアラニン (F) かトリプトファン (W) かチロシン (Y) のどれかであることを示す。EGF_1 あるいは EGF_2 のどちらかのパターンに合えば、EGF ドメインである可能性が高い。ただし、これらのパターンには当てはまらない EGF ドメインもあるし、逆に、EGF ドメインとすべきではないアミノ酸配列も少数ではあるが検出されてしまう。このような判断が難しいケースでは、SS 結合の情報や立体構造決定なくしては最終的な判断はできない。

EGF ドメインのサブセットとして、カルシウムイオンを結合できる EGF ドメインがある（図 2E）。これを EGF-CA として、カルシウムイオン結合部位を含まない EGF ドメインと区別する場合もある。カルシウムイオン結合部位は N 端側サブドメインに含まれていて、そのコン

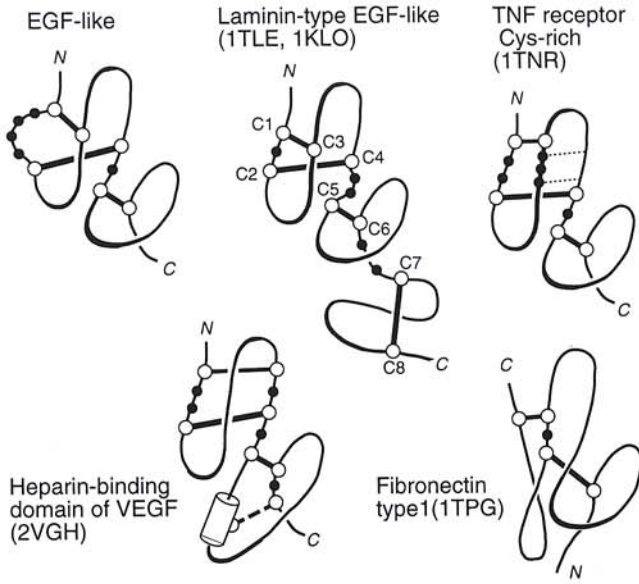


図3 EGFドメインのフォールドの模式図(左上)と、EGFドメインとフォールドに共通性があるシステインリッチフォールド。括弧内はプロテインデータバンク名。

センサパターンは、
 EGF_CA: [DEQN]-x-[DEQN] (2)-C¹-x (3, 14)-C²-x (3, 7)-C³-x-[DN]-x (4)-[FY]-x-C⁴-x-C⁵である。このうち、1番目、3番目、4番目のアスパラギン(D)、グルタミン酸(E)、グルタミン(Q)、またはアスパラギン(N)残基がカルシウム結合に関与している。4番目のDまたはNは、酵素的に水酸基が付加されてエリスロ-β-ヒドロキシル化されていることがある。このEGF_CAパターンは、凝固系の蛋白質に多く見つかる。たとえば、VII, IX, X因子、プロテインC, S, Z, トロンボモジュリンなどである。IX, X因子では実際にカルシウムイオンが結合することが示されている。一般的には、これらのEGF_CAドメインが本当にカルシウムイオンを結合するかどうかは実験的に示す必要がある。カルシウムイオンが結合しても、蛋白質部分の立体構造はほとんど変わらない。図2CはEGF_CAであるIX, X因子の立体構造の重なり具合は、カルシウムを結合部位を持たない立体構造のそれと同程度であることを示している。

EGFドメインと関係が深いほかのドメイン

EGFドメインと関係の深いドメインが存在する(図3)。ラミニン型EGF様ドメインは約60残基からなり、6個ではなく8個のシステインを含む(1-3, 2-4, 5-6, 7-8)。アミノ酸配列上でも、立体構造的にも、EGFドメインのC端側にもう一つC端側サブドメインを重複させたような構造をしている。今の所、EGFドメインとラミニン型EGF様ドメインが進化的に関係しているのか否かは不明である。少々混乱を招くことであるが、EGFのレセプターには2つのシステインリッチ配列が存在し、これらはラミニン型EGF様ドメインの複数回の繰り返しからなると予想されている⁵⁾。リガンドと良く似たフォールドがそのレセプターにあることになる。ただし、EGFはシステインリッチ配列に結合するわけではない。

もう少し、類似性を甘く見ると、蛋白質の主鎖のフォールディングがEGFドメインと同じものが結構見つかる。TNF(腫瘍壊死因子)レセプターのシステインリッチリピート、VEGF(血管内皮増殖因子)のヘパリン結合ドメインの

フォールドはEGFドメインと同じである。また、意味があるかどうかは不明であるが、tPAやuPAにも含まれているフィブロネクチンタイプIドメインは、N端とC端を逆にすると良く似たフォールドとなる。

EGFドメインの機能

EGFドメインは普遍的に見いだされる、おそらくは自然界に最も多く存在するモジュールの一つであるが、その機能についてはよくわかっていない。前述した堅い構造体をつくるスペーサー的な役割や、カルシウム結合部位を提供する以外に、蛋白質-蛋白質相互作用を仲介する働きがあると考えられている。上皮成長因子とそのレセプターの結合はその典型的な例である。凝固系蛋白質のなかでも、組織因子(TF)とVII因子の複合体中にEGFドメインが蛋白質-蛋白質相互作用に関与している例が、X線結晶解析を用いて可視化されている(図2F)。1番目のEGFドメインが組織因子蛋白質と β シートが重なり合うようにして相互作用している。2番目のEGFドメインはセリンプロテアーゼドメインと主に相互作用している。

おわりに

蛋白質モジュールの存在は、すべての蛋白質の立体構造決定をしなくても、ホモロジーモデ

リングで立体構造を推定すれば、多くのドメインについて分子生物学的・生化学的な実験の解釈が、立体構造にもとづいて議論することを可能にした。しかし、あくまでその蛋白質の立体構造決定が行われる当面の間という保留付きである。個々の蛋白質の機能は原子レベルでの立体構造に依存しているので、最終的にはその目的とする蛋白質の立体構造決定が必要である。凝固系・線溶系の蛋白質についても、さらに立体構造決定が進むことが望まれる。

文 献

- 1) Campbell ID, Bork P: Epidermal growth factor-like modules. *Curr. Opin. Struct. Biol.* **3**: 385-392, 1993.
- 2) Bork P: Mobile modules and motifs. *Curr. Opin. Struct. Biol.* **2**: 413-421, 1992.
- 3) Ullrich A, Schlessinger J: Signal transduction by receptors with tyrosine kinase activity. *Cell* **61**: 203-212, 1990.
- 4) Sampoli Benitez BA, Hunter MJ, Meininger DP, Komives EA: Structure of the fifth EGF-like domain of thrombomodulin: An EGF-like domain with a novel disulfide-bonding pattern. *J. Mol. Biol.* **273**: 913-926, 1997.
- 5) Abe Y, Odaka M, Inagaki F, Lax I, Schlessinger J, Kohda D: Disulfide bond structure of human epidermal growth factor receptor. *J. Biol. Chem.* **273**: 11150-11157, 1998.