

【第108回生涯教育講座】

RNAスプライシングの構造生物学的研究

お ばやし えい じ うら の たけし
尾 林 栄 治 浦 野 健キーワード：RNA スプライシング，イントロン認識，
U2AF タンパク質，立体構造解析

要 旨

DNA から転写された RNA は，スプライシングによりイントロンが除かれた後，タンパク質に翻訳可能な mRNA となる。スプライシングは一塩基のずれも許されない厳密さが要求され，スプライシングの異常を伴う病気も数多く報告されている。筆者らは，スプライシングの分子機構の理解を基に，病気の診断や治療への応用を目的に研究を行っている。特に，スプライシングに関わるタンパク質の立体構造解析を行うことで，スプライシングの分子機構およびスプライシング不全を引き起こす機構を原子レベルで解明していく。ここでは，RNA スプライシングと立体構造解析について簡単に紹介した後，筆者らが行っているイントロンを認識する U2AF タンパク質に関する研究について概説する。

私たち人間を形作るために必要な遺伝情報は，DNA というかたちで細胞内に保存されている。この遺伝情報データベースには，主に人間の体の中で働くタンパク質の設計図が記録されている。そしてその情報は，必要なときに，必要な部分だけが RNA として写しとられる。この DNA から RNA，そしてタンパク質が作られるという流れは，生物にとって普遍的な定義「セントラルドグマ」として1950年代に提唱された。一つの遺伝子から一つのタンパク質ができると長い間考えられてきた。しかし近年，様々な生物の全遺伝子配列

が明らかになり，人間の遺伝子数（約25,000個）は，ハエ（約14,000個）や線虫（約19,000個）と比べて大きな差がないことが判明した。遺伝子の数だけでは，どのように人間のような複雑な生物が形成されるのか，説明が難しかった。その答えの一つが，RNA スプライシングである。1970年代後半に人間を含む真核細胞の遺伝子には，所々にタンパク質情報をコードしない分断部位があることが報告された。この発見は，遺伝子の機能に対する従来の分子生物学の常識を覆す衝撃的なものであった。発見者であるロバーツ博士とシャープ博士には1993年にノーベル医学・生理学賞が授与されている。その後，この分断部位はイントロン，タンパク質を作るために必要な部位はエキソ

Eiji OBAYASHI et al.

島根大学医学部病態生化学講座

連絡先：〒693-8501 出雲市塩冶町89-1

ンと命名された。新生 RNA からイントロンを取り除き、そしてエキソンをつなぎ合わせるという一連の反応は、RNA スプライシングと呼ばれている¹⁾。スプライシングにおいて、イントロンは必ず認識されるわけではなく、存在する場所や刺激により認識されずに、その後ろのエキソンと共にスキップされる場合がある。これは選択的スプライシングと呼ばれ、高等真核生物において多種多様なタンパク質を産み出す原動力となっている²⁾。このようなスプライシングによる遺伝情報の組み替えは、一つの遺伝子から多様なタンパク質を産み出すことが可能となると同時に、その機能不全により細胞にとって有害な情報に変化する危険性も持ち合わせている。そのため、スプライシングには厳密な正確さが要求されている。著者らは、このスプライシングがどのようにして正確に行われているのか、その反応機構を、スプライシングを行うタンパク質と RNA の立体構造から理解し、スプライシング異常による病気に対する治療薬開発への足がかりを得るために、研究を進めている。

タンパク質の立体構造解析

タンパク質は、アミノ酸が直鎖状に結合したペプチド鎖が、三次元的に折りたたまることで機能的に働いている。そのため、タンパク質の機能を正確に理解するためには、タンパク質を構成するアミノ酸、ひいてはアミノ酸を構成する炭素や酸素、窒素原子、また各種金属原子などの空間的な立体配置を知る必要がある。しかしながら、タンパク質はアミノ酸が100個以上結合した、分子量が一万を超える大きな高分子であり、そのアミノ酸の組成からタンパク質全体の立体構造を予測することは不可能である。このタンパク質の立体構造を初めて明らかにしたのは、イギリス MRC 分子生物学研究所のケンドリュー博士とペルーツ博士である。彼らは1958年に X 線結晶構造解析という技術を使って、筋肉中に存在するミオグロビンというタンパク質の立体構造を世界で初めて明らかにした(図1)^{3,4)}。彼らは後の1962年に共にノーベル賞を受賞し、タンパク質構造解析の父と呼ばれている。その後半世紀にわたって様々なタンパ

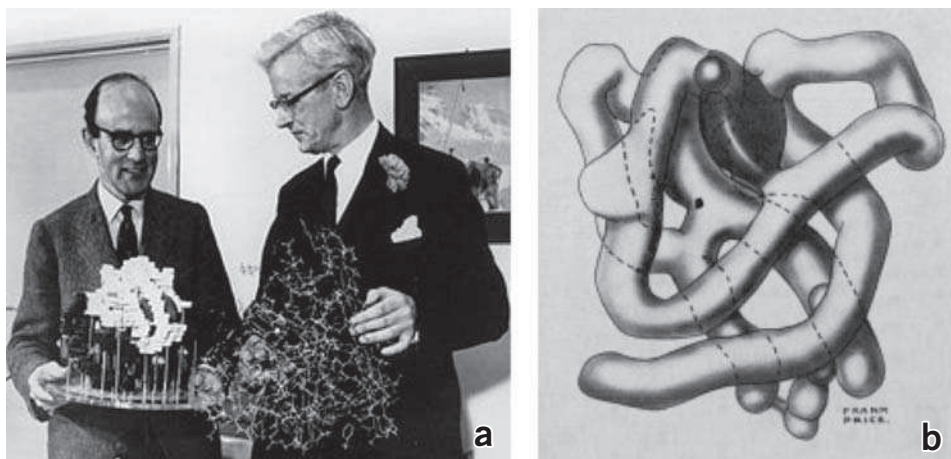


図1

(a) 世界で初めてタンパク質の原子レベルでの立体構造解析に成功したペルーツ博士(左)とケンドリュー博士(右)

(b) ミオグロビンの立体構造 (PDB: 1MBN)

ク質の立体構造が明らかにされ、現在では六万個以上のタンパク質の立体構造が、研究者が作るタンパク質構造データベース (Protein Data Bank; PDB) に登録されている。

また、立体構造を解析する手法も核磁気共鳴法 (NMR) や電子顕微鏡法などが開発された。NMR は溶液状態での構造解析が可能である反面、分子量に制限があり、現在の技術でも分子量が3万を超えると解析は困難を極める。また、電子顕微鏡法では原子の位置を正確に解析するような原子レベルでの構造解析を行うことはできない。そのため現在でも、分子量が大きいタンパク質の立体構造解析に対しては、X線結晶構造解析が唯一の研究手段となっている。実際に PDB に登録されている構造のうち、80%以上がX線結晶構造解析によるものである。しかしX線結晶構造解析に短所がないかと言われるとそうではない。X線結晶構造解析を行うためには、その目的とするタンパク質の単結晶を作る必要がある。小学生の頃試した、食塩やミョウバンの結晶と同様、そのタンパク質が整然と並んでできる結晶である。しかしタンパク質の結晶を作製するのは、食塩やミョウバンとは比較にならないほど難しい。さらに結晶を作製するためには純度の高いタンパク質が大量に必要となり、その単離・精製もX線結晶構造解析を行う上での「壁」となっている。

さらに近年では、明らかにされたタンパク質の立体構造からその機能を理解するだけでなく、その機能を阻害したり促進したりするような化合物をデザインすることで、新規薬剤を開発する「ドラッグデザイン」という観点からもタンパク質の立体構造解析は注目されている。実際に、新規抗生物質開発に有用なリボソームの立体構造解析の成果には2010年、また肥満や糖尿病、高血圧に対

する薬剤ターゲットとして注目されているGタンパク質共役型受容体 (GPCR) の立体構造解析の成果には2012年、それぞれノーベル賞が授与されており、基礎研究のみならず、製薬業界においても積極的に研究が進められている分野である。

スプライソソームタンパク質

スプライシングは、スプライソソームと呼ばれる100個以上のタンパク質とRNAの複合体によって行われる⁵⁾。スプライソソームを構成するタンパク質は、U1, U2, U4/U6, U5 snRNP タンパク質, Prp19 複合体タンパク質, そしてその他のタンパク質に大きく分けられる。核内低分子リボ核タンパク粒子 (snRNP) は数百塩基からなる短いRNA (核内低分子RNA, snRNA) にタンパク質が結合した複合体で、主に核内で働いている。U1 snRNP はU1 snRNA に10個のタンパク質が結合した複合体であり、U2 snRNP はU2 snRNA に20個のタンパク質が結合した複合体、U5 snRNP はU5 snRNA に14個のタンパク質が結合した複合体である。U4/U6 snRNP は、U4 snRNA とU6 snRNA の二本のsnRNA に19個のタンパク質が結合した複合体である。これら以外にも、RNAの二次構造を変えることのできるDNA/RNAヘリカーゼであるDEAD boxタンパク質や、選択的スプライシングを調節するタンパク質など、様々なタンパク質がスプライシング反応を行っている。さらに、スプライソソーム複体に含まれていることは確認されているものの、既知のアミノ酸配列モチーフを持たない、また相互作用するタンパク質やRNAが不明といった理由から、その働きが未だに明らかでないタンパク質が50個以上存在し、今後のさらなる研究が待たれる。

スプライシング機構

これまでの研究から、スプライシングはスプライソソームにより以下のように進められると考えられている (図2)¹⁾。スプライシング反応において最初の段階は、イントロンの認識である。イントロンには3カ所の特徴的な配列が存在し、スプライソソームタンパク質によって認識される。イントロンRNAの5'側にはGC配列を含んだ5'スプライス部位があり、U1 snRNPにより結合・認識される。その反対側の3'側にはAG配列を持つ3'スプライス部位があり、U2AFタンパク質により認識される。U2AFタンパク質はU2AF65とU2AF35の2種類のサブユニットからなり、図2ではそれぞれ“65”と“35”と表記した。また3'スプライス部位の少し上流にはブランチ部位 (イントロン中に保存されたアデノシン塩基で、図2では“A”としている)が存在し、SF1/BBPタンパク質により結合・認識される。このイントロンの認識が終わると、U2AFがブランチ部位にU2 snRNPをリクルートして、ブランチ部位に結合しているSF1/BBPがU2 snRNPに置き換わることでスプライシング反応が進んでいく。しかしこの時、U2AFの働きが阻害され、U2 snRNPがブランチ部位に来ることができないと、その3'スプライス部位の下流にあるエキソンはスキップされて選択的スプライシングが起こる²⁾。選択的スプライシングが起こる機構や条件は多数存在し、現在も様々な研究が進められているが、まだ不明な点が多い。U2 snRNPがブランチ部位に結合すると、Prp5やいくつかのSRタンパク質を介してU1 snRNPとU2 snRNPがブリッジされ、5'スプライス部位とブランチ部位が近くに配置された“A complex”が形成される。する

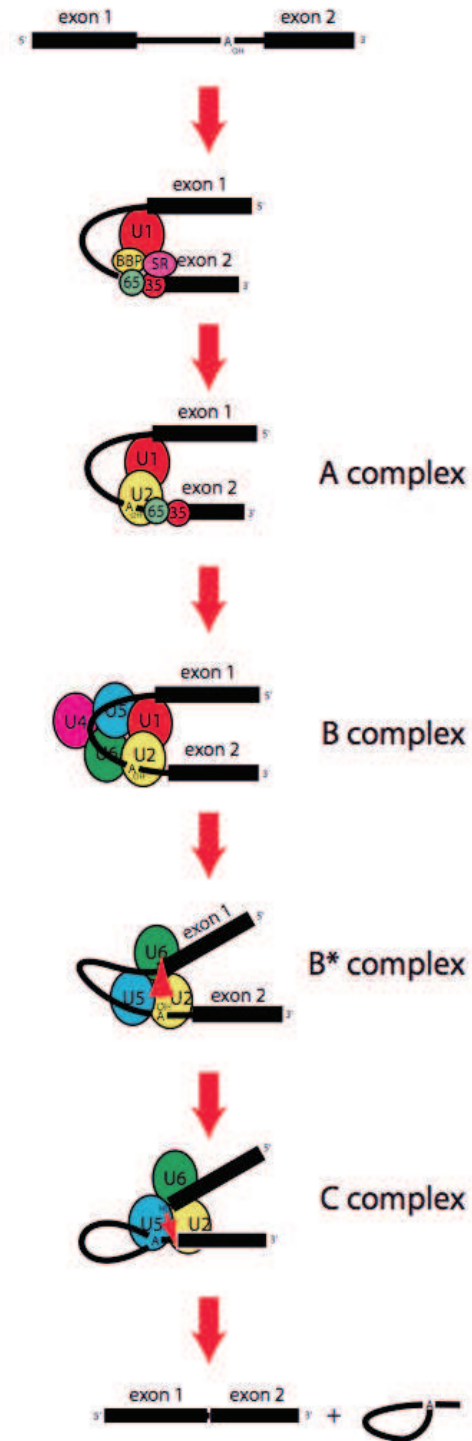


図2. スプライシングの反応機構。

U1, U2, U4, U5 および U6 はそれぞれの snRNP を、また BBP は BBP/SF1 タンパク質を、65, 35 はそれぞれ U2AF65, U2AF35 を、そして SR は SR タンパク質を示している。また RNA 中の A はイントロン中に保存されたブランチ部位のアデノシン塩基を示している。図中の“B* complex”および“C complex”に示した矢頭のようにエステル転移反応が2段階にわたって起こり、エキソン同士がつなぎ合わされる。

と U4/U6・U5 tri-snRNP が導かれ、いくつかのタンパク質の働きにより非常に大きな RNA・タンパク質の構造再配置が起こることで、活性型のスプライソーム複合体 (B complex) が形成される。具体的には、DEAD/H box タンパク質 U5-200k が U4/U6 snRNA 二本鎖を開裂させ、U6は U2 snRNA と塩基対を形成する。また、U1 snRNP と 5'スプライス部位の相互作用は、別の DEAD/H box モチーフを含む U5-100k タンパク質によって開裂されて、5'スプライス部位は U6 snRNP によって再認識される。この反応を通して、U1と U4 snRNP は “B complex” から解離し、スプライソームは “B* Complex” となる。さらに、Prp19 複合体が “B* complex” に結合する。Prp19 複合体の組成ははっきりと理解されていないが、7個のタンパク質が中心に存在することが明らかにされている。Prp19 複合体はスプライシング反応に必須であり、最初のエステル転移反応に必要とされる。また、U5, U6 snRNP の複合体への結合の安定化、U6 snRNA-5'スプライス部位塩基対の安定化に寄与している。“B* complex” が生成されると、最初のエステル転移反応が起こり、mRNA は 5'エキソンとラリアット中間体に分けられ、複合体はまた少し形を変えて “C complex” になる。このときに重要な役割を果たすのが U5 snRNP だと考えられている。U5 snRNP は、5'および 3'スプライス部位をそれぞれ認識して、DEAD/H box タンパク質 Prp16 の助けを借りて、2番目のエステル転移反応が起きやすいようにスプライソーム中で mRNA を再配置する。その後、近くに配置された 5'と 3'スプライス部位間で 2番目のエステル転移反応が起こり、5'スプライス部位上流のエキソンと 3'スプライス部位下流のエキソンがつなぎ合わされて、

スプライシング反応は完了する。活性型のスプライソーム複合体である “B* complex” および “C complex” は U2, U5 および U6 snRNP を含む数十個のタンパク質からなる非常に大きな複合体である。しかも二段階のエステル転移反応を行う過程で、その組成や立体配置を変化させる複雑さも相まって、その立体構造はおろか、どのタンパク質がどのような役割を持っているのかさえも不明であり、今後の研究が待たれる。

スプライシングと疾患

DNA もしくは RNA 上に記録されているタンパク質のアミノ酸配列情報は、アデニン、シトシン、グアニン、チミン (RNA の場合はウラシル) の 4 種類の塩基によって構成されている。しかしながら、アミノ酸は 20 種類存在するため、3つの塩基の組み合わせ (並び) で一つのアミノ酸をコードしている。例えば、ATG の並びはアミノ酸のメチオニンを、CAC はヒスチジンを表している。そのため、スプライシングには一塩基の「ずれ」も許されない正確さが要求される。なぜなら、エキソンのつなぎ合わせの際の一塩基の「ずれ」は、一アミノ酸の「ずれ」を生じるだけでなく、その後にコードされているアミノ酸をすべて変えてしまうからである。実際にこれまでの研究により、この「ずれ」からくるスプライシング異常によって筋ジストロフィー、メンケス病など様々な病気が引き起こされることが報告されてきた。最近では、骨髄異形成症候群やがんとの関連が注目されている^{6,7)}。両患者の細胞のゲノム解析の結果、高い確率でスプライソームタンパク質にアミノ酸変異があることが報告された (図 3)。興味深いことに、アミノ酸変異は特定のタンパク質の特定の部位に偏って存在していた。注

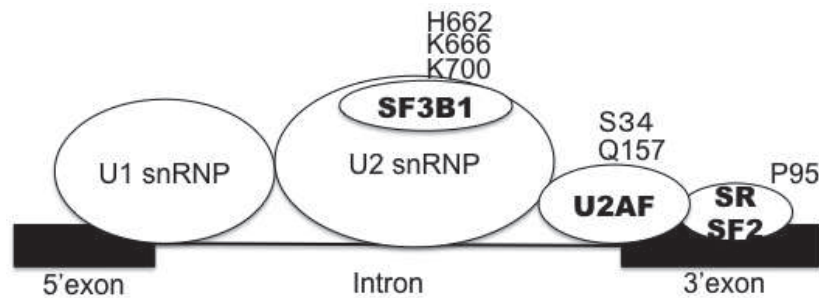


図3. スプライシング初期複合体と骨髄異形成症候群患者に見られるアミノ酸変異
イントロンの5'側はU1 snRNPに、3'側はU2AFタンパク質により認識される。
U2 snRNPはブランチ部位を認識する。骨髄異形成症候群患者に見られるアミノ酸変異は主にイントロンに直接結合するタンパク質に多く見られている。

目すべきは、それらのほとんどが、スプライシングの初期段階でイントロンを認識するタンパク質群に見られるということである。すなわち、スプライソソームタンパク質によるイントロンの誤った認識が、スプライシング不全を引き起こし、正常な生命反応を阻害し病気を引き起こしているのである。そこで著者らは、骨髄異形成症候群やがんにおいて見られるアミノ酸変異を持ったスプライソソームタンパク質の働きを正確に理解し、また変異部位の役割から、病気が引き起こされるメカニズムを明らかにすることを目的に研究を進めている。

U2AF タンパク質

骨髄異形成症候群やがんで見つかった特異的なアミノ酸変異を持つスプライソソームタンパク質の一つが、U2AFである。U2AFは、イントロンの3'側を認識するスプライシングに必須なタンパク質複合体である⁸⁾。選択的スプライシングもU2AFの働きを制御することで起こる例が多い。また、U2AF小サブユニット(U2AF35)は3'スプライス部位の認識に強く関わっているため、U2AF35の正確な機能に関して、世界中で研究が進められている。しかし、U2AF35はその物

性の悪さから大サブユニット(U2AF65)に比べ研究の進みが遅く、その機能に関する報告例も少ないのが現状である。実際、U2AF35単独で3'スプライス部位に結合するのか、また3'スプライス部位をRNA塩基配列特異的に認識できるのか、さらにはU2AF35のどの部位でそれが行われるのかなど、不明な点が多かった。U2AF35はRNA結合モチーフ(RRM)と呼ばれる領域を持つことから、古くからこの領域がイントロンRNAを認識・結合すると考えられてきた⁹⁾。しかし、RRMの中でもRNA結合に関与しないものも多く見つかかり、U2AF35の持つRRMもそのグループに属すること、またNMRや他の生化学的報告から、U2AF35はRRM以外の部位でRNAを認識・結合しているのではないかと考えられるようになってきている。実際に、骨髄異形成症候群やがんにおける変異もRRMではなく二つの亜鉛結合(Znフィンガー)領域に見つかっている。そのため筆者らはこれまでに、U2AF35の持つ二つのZnフィンガー領域を含む形での立体構造解析を進めてきた。筆者らは酵母のU2AF小サブユニット(U2AF23)がヒトのU2AF35に非常に良く似たアミノ酸配列を持つことに注目し、これを安定に単離・精製すること

に成功し、その立体構造を解明した。また、U2AF23 単独で配列特異的に RNA に結合することを明らかにした (論文投稿中)。しかしながら、本構造からだけでは、U2AF23 がどの部位で、またどのように配列特異的に RNA に結合するのかまでは説明することができなかった。さらに、U2AF35 において骨髄異形成症候群やがんを引き起こすと考えられているアミノ酸変異部位の役割も不明なままである (図3)。そのため筆者らは、U2AF23 および U2AF35 の RNA の認識機構を原子レベルで明らかにすることを目的に今後さらに研究を進めるとともに、アミノ酸変異がスプライシング反応に与える影響を解明し病気発症のメカニズムに迫って行く。

ま と め

RNA スプライシングは、我々ヒトが生きていく上で必須の基本的な生命反応であり、その破綻

は病気の発症、さらには死につながる。しかしながら、選択的スプライシング機構や、病気発症のメカニズムを含め、まだまだ不明な点が多い。そのため、このスプライシング機構の正確な理解は、生命反応を深く知る上で重要なだけでなく、様々な病気の治療法の開発など医学の発展に貢献する。特に、骨髄異形成症候群は、急性骨髄性白血病の発症を促し、人を死に至らしめる病気であり、厚生労働省により難病指定されている。この病気の原因はよくわかっていない。年齢と共に発症率が増加するため、今後の高齢化社会の進展により患者数の増加が容易に予想され、その対策は急務である。特に島根県は高齢者率が日本一であり、医学部付属病院には毎年十数人の患者が治療に訪れることから、本研究を進めることの意義は大きい。筆者らの研究から、骨髄異形成症候群を簡便に同定するキットの開発、また新たな治療法を模索するなど、医療への応用が待たれる。

参 考 文 献

- 1) Will, C. L. & Luhrmann, R. Spliceosome structure and function. *Cold Spring Harbor perspectives in biology* 3, (2011)
- 2) Nilsen, T. W. & Graveley, B. R. Expansion of the eukaryotic proteome by alternative splicing. *Nature* 463, 457-463, (2010).
- 3) Kendrew, J. C., Dickerson, R. E., Strandberg, B. E., Hart, R. G., Davies, D. R., Phillips, D. C. and Shore, V. C. Structure of myoglobin: A three-dimensional Fourier synthesis at 2 Å resolution. *Nature* 185, 422-427, (1960)
- 4) Perutz, M. F., Rossmann, M. G., Cullis, A. F., Muirhead, H., Will, G. and North, A. C. Structure of haemoglobin: A three-dimensional Fourier synthesis at 5.5-Å resolution, obtained by X-ray analysis. *Nature* 185, 416-422, (1960)
- 5) 尾林栄治, 長井 潔: スプライソソーム, 実験医学増刊 22, 2385-2392 (2004)
- 6) Yoshida, K. *et al.* Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature* 478, 64-69, (2011).
- 7) Brooks, A. N. *et al.* A pan-cancer analysis of transcriptome changes associated with somatic mutations in U2AF1 reveals commonly altered splicing events. *PloS one* 9, e 87361 (2014)
- 8) Zamore, P. D. & Green, M. R. Identification, purification, and biochemical characterization of U2 small nuclear ribonucleoprotein auxiliary factor. *Proceedings of the National Academy of Sciences of the United States of America* 86, 9243-9247 (1989)
- 9) Kielkopf, C. L., Rodionova, N. A., Green, M. R. & Burley, S. K. A novel peptide recognition mode revealed by the X-ray structure of a core U2AF35/U2AF65 heterodimer. *Cell* 106, 595-605 (2001)