

## 論文の内容の要旨

### 論文題目

Crystal structures of CRISPR-Cas9 and RNA methyltransferase  
(CRISPR-Cas9 および RNA メチル化酵素の結晶構造解析)

氏名 平野 清一

生命が設計図であるゲノムから実際に成り立つ過程において、DNA および RNA はタンパク質と相互作用し変化する。本研究では X 線結晶構造解析によって、RNA 依存性 DNA 切断酵素 Cas9 および RNA メチル化酵素 CAPAM が標的の核酸を認識し作用する分子メカニズムを明らかにした。

Cas9 は原核生物の獲得免疫機構に関与する RNA 依存性 DNA エンドヌクレアーゼである。Cas9 とガイド RNA を細胞に導入することで、ゲノムの特定の領域を切断し DNA 配列を改変することができるため、現在ゲノム編集ツールとして広く利用されている。Cas9-ガイド RNA 複合体は、PAM (protospacer adjacent motif) とよばれる特定の 2-5 塩基を認識した後、20 塩基程度のガイド配列と相補的な標的配列をほどこき、二本鎖 DNA 切断を引き起こす。

ゲノム編集に広く利用されている SpCas9 は NGG という配列を PAM として認識するため、ゲノム編集可能なゲノム領域には制限が存在するという問題が残されている。ゲノム編集の標的部位を拡張するために、分子進化的手法によって、異なる PAM を認識する SpCas9 改変体 (VQR 改変体および VRER 改変体) が開発された。VQR 改変体は 3 重変異 (D1135V/R1335Q/T1337R) をもち NGA を PAM として認識する一方、VRER 改変体は 4 重変異 (D1135V/ G1218R/R1335E/T1337R) をもち NGCG を PAM として認識する。

しかし、複数の変異が導入された SpCas9 改変体が異なる PAM を認識する分子機構は不明であった。そこで本研究では、VQR 改変体-ガイド RNA-標的 DNA (TGA PAM)、および、VRER 改変体-ガイド RNA-標的 DNA (TGCG PAM) 複合体の結晶構造を決定した。野生型 SpCas9 の結晶構造では PAM の 2、3 文字目の G が Arg1333、Arg1335 と水素結合を形成する一方で、VQR 改変体では PAM の 2、3 文字目の G、A が Arg1333、Gln1335 (R1335Q) と水素結合を形成しており、VRER 改変体では PAM の 2、3、4 文字目の G、C、G が Arg1333、Glu1335 (R1335E)、Arg1337 (T1337R) と水素結合を形成していた (図 2)。野生型 SpCas9 と SpCas9 改変体の構造の比較から、DNA の糖-リン酸骨格が予想外の構造変化を起こし、PAM の 3 文字目の塩基が 1335 番目の残基に近づいていることが明らかになった。Val1135 (D1135V)、Arg1218 (G1218R) は DNA の糖-リン酸骨格と相互作用し、この DNA の構造変化を促進していた。

SpCas9 は NGG という配列を PAM として認識するのに対し、CdCas9 は NNRHHHY (R は A/G、H は A/T/C、Y は T/C) という配列を PAM として認識する。CdCas9 は 5 文字の塩基が組み合わさった 108 通り ( $2 \times 3 \times 3 \times 3 \times 2$ ) の PAM を認識できるため、標的範囲の広いゲノム編集ツールとして期待されている。しかし、CdCas9 による複雑な PAM 認識機構は不明であった。そこで本研究では、CdCas9-ガイド RNA-標的 DNA (TGGTAAT PAM) 複合体の結晶構造を決定した。CdCas9 は、SpCas9 と類似構造をもつドメインを用いてガイド RNA や標的 DNA を認識する一方で、SpCas9 とは異なるアミノ酸残基を用いて PAM 配列を認識していることが明らかになった。PAM の 3 文字目の G は Arg1017 と水素結合を形成し、PAM の 7 文字目の T と相補する A は Lys1015 と水素結合を形成していた。PAM の 5、6 文字目の A と相補する T は、Phe1011、Pro1043、Leu1046 によって形成される疎水性パッチと相互作用していた。これらのことから、他の Cas9 とは異なり、CdCas9 は疎水性相互作用と水素結合を組み合わせて利用することにより、複雑な 108 通りの PAM を認識することが明らかになった。

真核生物の mRNA が転写後修飾をうけることは古くから知られていたが、その全容は謎に包まれていた。最近のエピトランスクリプトーム解析技術の進歩により、mRNA の転写後修飾が数多く同定され、発生の各段階やストレス応答時など特定の状況下において、転写後修飾をもつ mRNA が一斉に制御を受けるという新しい遺伝子発現制御の概念が提唱されている。mRNA の転写後修飾のなかで、アデニン 6 位のアミノ基がメチル化された m<sup>6</sup>A 修飾は最も注目されており、遺伝子の発現制御を通して、がんや神経変性などの疾患、および、性決定などの生命現象に関わっていることが明らかになってきた。真核生物の mRNA は 5' 端に m<sup>7</sup>G 塩基が三リン酸や反転したリボースで結合した特徴的な構造をもつ (5' キャップ)。Cap-specific adenosine-*N*<sup>6</sup>-methyltransferase (CAPAM) は 5' キャップに隣接する

A を m<sup>6</sup>A 修飾し、CAPAM のノックアウト細胞ではストレス応答遺伝子などの翻訳が抑制され、ストレス条件下での成育が阻害されることが明らかにされていた。しかし、CAPAM がどのように mRNA の 5' キャップを認識し、隣接する A をメチル化するかは不明であった。そこで本研究では、ゼブラフィッシュ由来 CAPAM、5' キャップ、メチル基供与体アナログ (SAH) からなる複合体の結晶構造を決定した。

結晶構造から、CAPAM はメチル化転移酵素に共通するロスマンフォールドをもつメチル化ドメイン、および、 $\alpha$ ヘリックスが入り組んだ新規フォールドをもつヘリカルドメインから構成されることが明らかになった。5' キャップはメチル化ドメインとヘリカルドメインの間のポケットで認識され、メチル基供与体アナログ (SAH) はメチル化ドメインの触媒ポケットで認識されていた。変異体解析から、標的アデニン塩基は他の m<sup>6</sup>A 修飾酵素と同様にメチル化ドメインの活性部位で認識され、RNA の下流部分はヘリカルドメインの正に帯電した溝に結合することが示唆された。さらに、CAPAM と他の m<sup>6</sup>A 修飾酵素の構造比較から、これらの m<sup>6</sup>A 修飾酵素は保存されたメチル化ドメインを用いてアデニンをメチル化する一方で、多様なフォールドをもつ付加ドメインを用いて RNA に対する特異性を獲得していることが明らかになった。