

論文の内容の要旨

Structural analysis of Cpf1, a novel nuclease of a CRISPR-Cas system

(CRISPR-Cas 系に関わる新規ヌクレアーゼ Cpf1 の X 線結晶構造解析)

山野 峻

【背景】 CRISPR (clustered regularly interspaced short palindromic repeat) -Cas (CRISPR-associated protein)系は原核生物のもつ獲得免疫機構であり, Cas タンパク質が crRNA (CRISPR RNA) とエフェクター複合体を形成し, crRNA と相補な外来核酸を認識し切断する. CRISPR-Cas 系はクラス 1 とクラス 2 に分類され, 単一の Cas タンパク質がエフェクター分子として機能するクラス 2 はさらに II 型, V 型, VI 型に分類される. II 型 CRISPR-Cas 系で機能するエフェクター分子である Cas9 はゲノム編集ツールとして広く利用されている. V 型 CRISPR-Cas 系に関与する Cpf1 (Cas12a) では, 細菌 *Acidaminococcus sp.* BV3L6 と細菌 *Lachnospiraceae bacterium* ND2006 に由来する Cpf1 (AsCpf1, LbCpf1) が真核生物の細胞内で標的 DNA 切断活性を示すため, ゲノム編集ツールとして用いられている (Zetsche *et al.*, *Cell*, 2015). しかし, Cas9 と Cpf1 は異なる特徴をもつ. まず, Cas9 は標的 DNA を平滑末端で切断する一方, Cpf1 は突出末端で切断する. さらに, 両者とも標的 DNA のターゲティングにはガイド RNA と標的配列との相補性にくわえ, PAM (protospacer adjacent motif) とよばれる配列が標的配列近傍に存在することが必須だが, Cas9 は G リッチな PAM を認識する一方, Cpf1 は T リッチな PAM を認識する. Cpf1 の構造情報が未知であったため作動機構および, 上記の特徴を示す理由は不明であった.

【結果】 本論文では, 第 2 章で AsCpf1 と crRNA および標的 DNA 複合体の構造から明らかになった Cpf1 の作動機構について述べる. 第 3 章では AsCpf1 の PAM の認識を変更した Cpf1 改変体について構造解析を行い, 変化した PAM 認識機構を明らかにした. 第 4 章では Cpf1 は C を含む PAM を認識できることを生化学実験から明らかにし, LbCpf1 と crRNA および標的 DNA 複合体の構造を複数の PAM について決定し, LbCpf1 の C を含む PAM の認識機構を明らかにした.

AsCpf1-crRNA-標的 DNA 複合体の全体構造を 2.8 Å 分解能で決定した. AsCpf1 は REC ロー

ブと NUC ロープからなり、REC ロープは REC1 ドメインと REC2 ドメインから、NUC ロープは PI ドメイン、WED ドメイン、RuvC ドメインおよび Nuc ドメインから構成されていた。crRNA のガイド配列と相補鎖 DNA は RNA : DNA ヘテロ 2 本鎖を形成し、2 つのローブの間に收容されていた。標的 DNA の PAM 2 本鎖は REC1 ドメイン、WED ドメイン、PI ドメインにより形成される空間に收容されていた。標的 DNA の切断に関して、相補鎖および非相補鎖 DNA はヌクレアーゼドメインである RuvC ドメインにより切断されることが示唆されており、RuvC ドメインの活性部位を形成するアミノ酸は種間で保存されていた。しかし、相補鎖 DNA と RuvC ドメインの活性部位は離れており、その切断機構は本構造からは不明であった。RuvC ドメインには機能未知の Nuc ドメインが隣接していた。Nuc ドメインにおいて保存された Arg1226 に変異を導入すると、非相補鎖 DNA のみ切断され相補鎖 DNA は切断されなかった。この結果から、Nuc ドメインは相補鎖 DNA を RuvC ドメインの活性部位へと誘導することが示唆された。この機構は Cpf1 が標的 DNA を突出末端で切断する仕組みをうまく説明できるものであった。AsCpf1 による TTTV PAM (V は A, C, G) の認識機構が明らかになった。Cpf1 に結合した PAM 2 本鎖 DNA は B 型 DNA に比べて、副溝が狭く歪んだ構造をしていた。この狭い副溝に PI ドメインの Lys607 の側鎖が挿入され、PAM 2 本鎖の構造を認識するとともに PAM の塩基を水素結合により認識していた。また、WED ドメインの Lys548 が主溝から PAM の塩基を水素結合により認識していた。この構造から、Cpf1 の PAM 認識機構では水素結合による塩基の認識に加え、PAM 2 本鎖が歪んだ構造をとることの重要性が明らかになった。

認識する PAM が変化した Cpf1 変体 (RVR 変体, RR 変体) が作成された。RVR 変体は TATV PAM と TTTV PAM を認識し、RR 変体は TYCV PAM を認識する。しかし、その認識機構は不明であった。2 つの変体について、Cpf1-crRNA-標的 DNA 複合体の構造を 2.0 Å 分解能で決定した。RVR 変体は TATA PAM と、RR 変体は TCCA PAM との複合体で構造決定した。野生型 AsCpf1 では PAM は主に副溝から認識されていた一方、2 つの変体では副溝からの認識は保存されつつも、主溝からの認識が増えていた。また、PAM の相補な塩基が強く認識されていることが明らかになった。

多くの Cpf1 は TTTV PAM を強く認識するが、C を含む配列も PAM として認識できる。C を含む PAM の認識機構を解明するため、C を含む PAM を比較的強く認識する LbCpf1 と crRNA および TTTA, TCTA, TCCA, CCCA という PAM を含む 4 種の標的 DNA を用いて LbCpf1-crRNA-標的 DNA 複合体を構成し構造解析を行った。各複合体は 2.4-2.5 Å 分解能で構造決定した。LbCpf1 の PAM 認識機構は AsCpf1 と共通しており、Lys595 が PAM 2 本鎖の副溝に挿入され、Lys538 と Tyr542 が主溝から PAM の塩基を認識していた。各構造の PAM 2 本鎖の形状を比較すると、C を多く含むに従い B 型 DNA に近くなっていた。副溝の幅が広がるにつれ PI ドメインの位置が分子の外側へと動き、それに伴い副溝に挿入された Lys595 の側鎖の位置が変化し、Lys595 と PAM 2 本鎖との相互作用が弱くなっていた。これにより、PI ドメインは構造変化を起こし C を含む PAM 2 本鎖と相互作用すること、相互作用の強さが PAM 認識に重要であることが明らかになった。

【考察・展望】 Cpf1-crRNA-標的 DNA 複合体の結晶構造から Cpf1 の標的 DNA 切断機構、詳細な PAM 認識機構が明らかになった。Cpf1 は Cas9 と比較して、各ドメインの役割やヘテロ 2 本鎖の認識は共通していた一方、PAM 認識機構は大きく異なっており副溝からの塩基の認識は特徴的であった。Cpf1 は Cas9 よりも標的 DNA に対する特異性が高く、オフターゲットの切断が少ないことが知られている。また、Cpf1 変体の作成が可能であることが示され、PAM 認識機構の詳細も明らかになった。本研究で得られた知見に基づき、さらなる Cpf1 変体の創出が期待される。