

長期利用課題報告 2

ゲノム編集ツール Cas9 エンドヌクレアーゼの X 線結晶構造

東京大学大学院 理学系研究科

西増 弘志、濡木 理

Abstract

原核生物のもつ CRISPR-Cas 獲得免疫機構に関する RNA 依存性 DNA ヌクレアーゼ Cas9 はガイド RNA と相補的な 2 本鎖 DNA を選択的に切断する性質をもつ。近年、Cas9 を利用したゲノム編集技術は基礎研究から臨床応用にいたる幅広い分野において急速に普及した。本長期利用課題では異なる細菌に由来する多様な CRISPR-Cas 酵素の結晶構造を決定し、その RNA 依存性 DNA 切断機構を原子レベルで明らかにすることに成功した。さらに、構造情報を基にした新規のゲノム編集ツールの開発にも成功した。

1. CRISPR-Cas 系

原核生物は外来核酸に対する防御機構として CRISPR-Cas (clustered regularly interspaced short palindromic repeat-CRISPR-associated) とよばれる獲得免疫機構をもつ¹⁾。CRISPR-Cas 系では Cas (CRISPR-associated) タンパク質と crRNA (CRISPR RNA) が複合体を形成し、crRNA のガイド配列と相補的な外来核酸を認識・切断する。CRISPR-Cas 系は 2 つのクラスに分類される^{2,3)}。クラス 1 の CRISPR-Cas 系には複数の Cas タンパク質からなる Cas-crRNA 複合体が関与する。一方、クラス 2 の CRISPR-Cas 系では単一の Cas タンパク質が外来核酸を切断する。クラス 2 の CRISPR-Cas 系は II 型、V 型、VI 型に分類される。II 型 CRISPR-Cas 系においては RNA 依存性ヌクレアーゼ Cas9 が crRNA および tracrRNA (*trans*-activating crRNA) と複合体を形成し、crRNA

のガイド配列と相補的な 2 本鎖 DNA を切断する^{4,5)} (図 1)。Cas9 は 2 つのヌクレアーゼドメイン (RuvC と HNH) をもち、HNH ドメインはガイド RNA と相補的な DNA 鎖 (相補鎖) を切断する一方、RuvC ドメインはもう一方の DNA 鎖 (非相補鎖) を切断する。Cas9 による DNA 認識にはガイド RNA との相補性に加え、PAM (protospacer adjacent motif) とよばれる特定の塩基配列が必要である。crRNA と tracrRNA を人工的に連結した sgRNA (single-guide RNA) も同様の機能をもち⁶⁾、20 塩基のガイド配列は自由に変更できるため、Cas9-sgRNA 複合体を用いることにより、ゲノム DNA の狙った位置を特異的に切断することが可能である。したがって、2013 年以降、Cas9 は革新的なゲノム編集ツールとして広く普及した⁶⁾。

2014 年、本研究グループはゲノム編集に広く利用されている *Streptococcus pyogenes* 由来 Cas9 (SpCas9) に着目し、SpCas9-sgRNA-DNA 複合体の結晶構造を世界にさきがけて決定し、Cas9 による RNA 依存性 DNA 切断の分子機構を明らかにした⁷⁾。さらに、他の研究グループにより報告された SpCas9 単体⁸⁾、SpCas9-sgRNA 複合体⁹⁾、SpCas9-sgRNA-標的 DNA 複合体^{10,11)}の結晶構造から、Cas9 による DNA 切断機構が明らかになってきた。しかし、異なる細菌に由来する Cas9 のアミノ酸配列の相同性は低く、ガイド RNA や PAM の塩基配列は大きく異なるため、Cas9 の作動機構には不明な点が残されていた。SpCas9 は NGG という配列を PAM として認識する

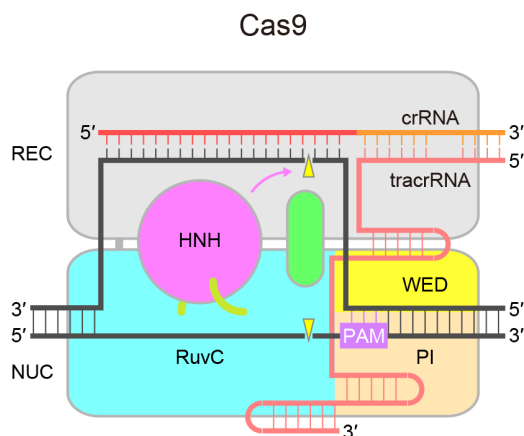


図 1 Cas9 による DNA 切断機構

一方、*Staphylococcus aureus* 由来 Cas9 (SaCas9)、*Francisella novicida* 由来 Cas9 (FnCas9)、*Campylobacter jejuni* 由来 Cas9 (CjCas9) はそれぞれ NNGRRT (R は A または G)、NGG、NNNVRYAC (V は A、G または C、Y は T または C) という配列を PAM として認識する^{[12][13]}。さらに、2015 年には V 型 CRISPR-Cas 酵素である Cpf1 (Cas12a) が発見された^[14]。Cpf1 は Cas9 と同様に RNA 依存的に 2 本鎖 DNA を切断するが、(1) tracrRNA を必要としない、(2) TTTV という配列を PAM として認識する、(3) PAM から離れた位置で DNA を切断する、(4) HNH ドメインをもたない、などの特徴をもつため、*Acidaminococcus sp.* 由来 Cpf1 (AsCpf1) や *Lachnospiraceae bacterium* 由来 Cpf1 (LbCpf1) は Cas9 と相補的なゲノム編集ツールとして利用されている。

本長期利用課題では異なる細菌に由来する CRISPR-Cas 酵素の結晶構造を決定することにより、その多様な RNA 認識機構、PAM 認識機構、DNA 切断機構を明らかにすることに成功した。

2. SaCas9 の結晶構造

SaCas9 (1053 残基) は SpCas9 (1368 残基) よりもサイズが小さいため、ウイルスベクターへの導入効率の高いゲノム編集ツールとして利用されている^[12]。SaCas9 の作動機構の理解を目指し、SaCas9-sgRNA-DNA 複合体を結晶化し、BL41XU において X 線回折データを収集し、SeMet 置換体を用いた SAD 法により結晶構造を決定した^[15] (図 2)。結晶構造から、SaCas9 は SpCas9 と同様に、REC ロープと NUC ロ

ープからなる構造をもつことがわかった。NUC ロープは RuvC ドメイン、HNH ドメイン、WED ドメイン、PI ドメインから構成されていた。WED ドメインは新規フォールドをもつことが明らかとなった。sgRNA のガイド配列は標的 DNA と RNA : DNA ヘテロ 2 本鎖を形成し、2 つのローブの間に結合していた。SaCas9 と SpCas9 の REC ドメインおよび WED ドメインは構造が異なり、それぞれの sgRNA を特異的に認識していた。PAM (TTGAAT) は 2 本鎖を形成し WED ドメインと PI ドメインの間に結合していた。PAM の 1 文字目、2 文字目の T 塩基 (T1 と T2) は SaCas9 と相互作用していなかった。一方、PAM の 3 文字目が G であることと一致して、G3 は Arg1015 と 2 本の水素結合を形成していた。A4 は Asn985 と水素結合し、A5 は Asn985、Asn986、Arg991 と水分子を介した水素結合により認識されていた。したがって、SaCas9 は A と G に共通の N7 と相互作用することにより、PAM (NNGRRT) の 4 文字目および 5 文字目の R (A または G) を認識していることが明らかになった。さらに、PAM の 6 文字目の T に対する嗜好性と一致して、T6 は Arg991 と水素結合していた。PAM 認識残基を変異させると、DNA 切断活性が低下したことから、これらの相互作用の重要性が確認された。配列相同性は低いにもかかわらず、SaCas9 と SpCas9 の PI ドメインは類似の構造をもっていた。しかし、PAM 認識残基は異なっており、SpCas9 では Arg1333 と Arg1335 が PAM (NGG) を認識していた。これらの構造比較から、SaCas9 と SpCas9 が異なる配列を PAM として認識する分子機構が明らかになった。

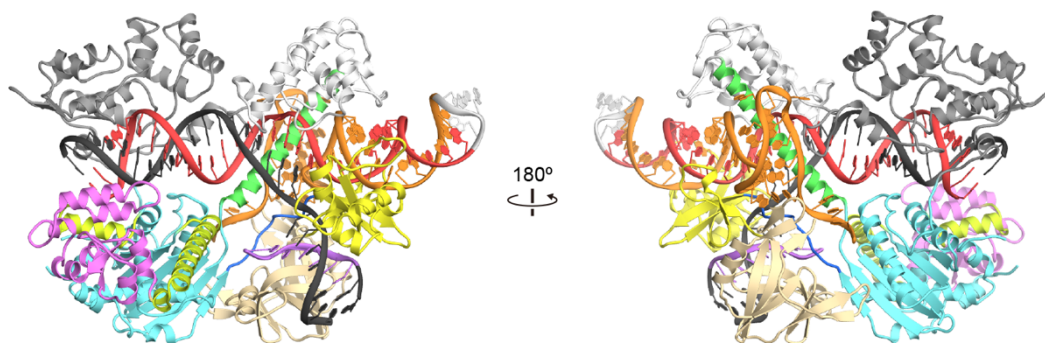


図 2 SaCas9-sgRNA-DNA 複合体の結晶構造

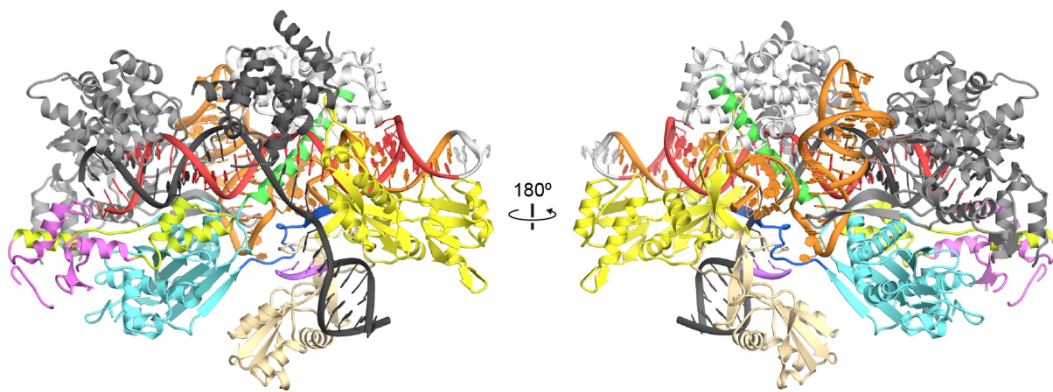


図3 FnCas9-sgRNA-DNA 複合体の結晶構造

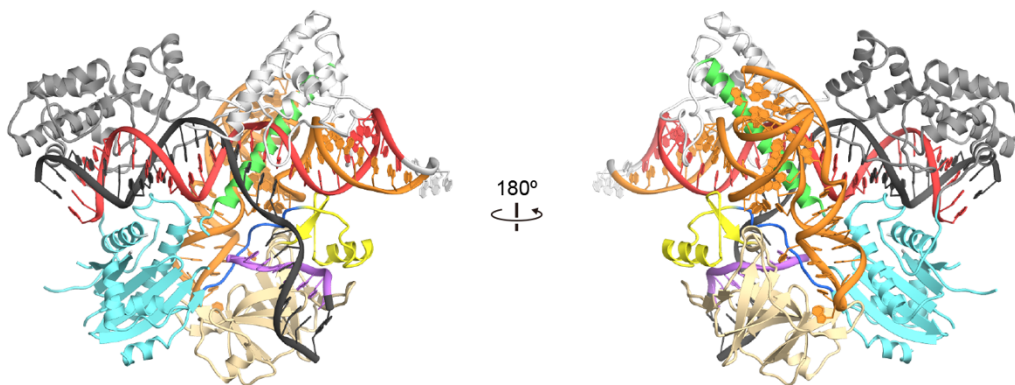


図4 CjCas9-sgRNA-DNA 複合体の結晶構造

3. FnCas9 の結晶構造

FnCas9 は 1629 残基からなり、Cas9 の中でも最もサイズが大きく、SpCas9 や SaCas9 との配列相同性が低い。FnCas9 の作動機構の理解を目指し、FnCas9-sgRNA-DNA 複合体を結晶化し、BL41XU において X線回折データを収集し、SeMet 置換体を用いた SAD 法により結晶構造を決定した^[16] (図 3)。FnCas9 は RuvC ドメイン、REC1 ドメイン、REC2 ドメイン、REC3 ドメイン、HNH ドメイン、WED ドメイン、PI ドメインの 7 つのドメインから構成されていた。FnCas9 の RuvC ドメインおよび HNH ドメインは SpCas9 や SaCas9 と同様の構造をとっていた。一方、REC ドメインおよび WED ドメインは新規フォールドをもっていた。SpCas9、SaCas9、FnCas9 の構造比較から、これらの 3 種の Cas9 の間の配列相同性は低いにもかかわらず、DNA 切断機構は保存されていることが明らかになった。一方、3 種の Cas9 の間で sgRNA の構造は大きく異なっており、それぞれ特徴的な構造をもつ REC ドメインと WED ドメインによ

り特異的に認識されていた。結晶構造から、FnCas9 が NGG という配列を PAM として認識する分子基盤も明らかになった。T1 は FnCas9 と相互作用していなかった一方、G2 と G3 は Arg1585 と Arg1556 とそれぞれ水素結合していた。さらに、得られた構造情報を基に変異を導入し、NGG ではなく YG という配列を PAM として認識する FnCas9 改変体の作製に成功した。

4. CjCas9 の結晶構造

CjCas9 は 984 残基からなる最小の Cas9 であるため、小型のゲノム編集ツールとして注目されている^[13]。CjCas9 の作動機構の理解を目指し、CjCas9-sgRNA-DNA 複合体を結晶化し、BL41XU にて X線回折データを収集し、SeMet 置換体を用いた SAD 法により結晶構造を決定した^[17] (図 4)。結晶構造から、CjCas9 は他の Cas9 と同様に、2 つのローブからなることが確かめられた。REC ローブは REC1 ドメインと REC2 ドメインから構成される一方、NUC ローブは RuvC

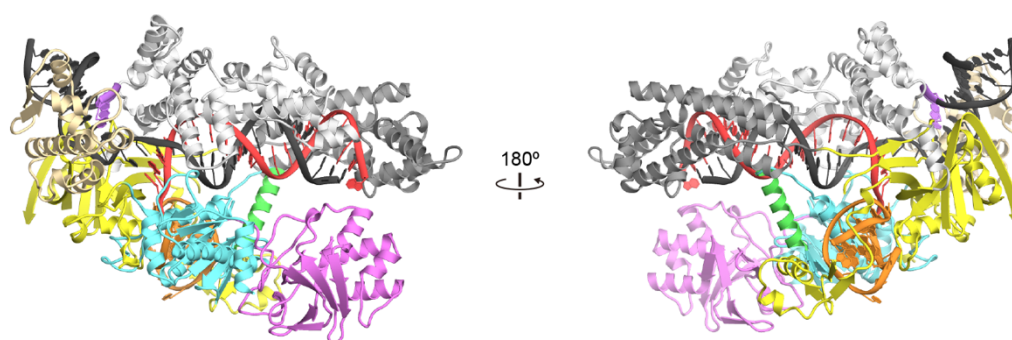


図5 AsCpf1-crRNA-DNA 複合体の結晶構造

ドメイン、WEDドメイン、および、PIドメインから構成されていた。CjCas9はSpCas9よりも小さなREC1ドメインとPIドメインをもち、さらに、SaCas9よりも小さなWEDドメインをもっていた。これらの構造比較からCjCas9の小型化の分子基盤が明らかになった。結晶構造から、CjCas9のsgRNAは他のCas9のsgRNAと異なり、予想外の3重らせん構造をもつことが明らかになった。さらに、他のCas9は非相補鎖の塩基をPAMとして認識するのに対し、CjCas9は相補鎖と非相補鎖の両方の塩基と水素結合を形成することによりPAM (NNNVRYAC) を認識していることが明らかになった。

5. Cpf1の結晶構造

Cpf1の作動機構の理解を目指し、AsCpf1-crRNA-DNA複合体を結晶化し、BL41XUにおいてX線回折データを測定し、SeMet置換体を用いたSAD法により結晶構造を決定した^[18](図5)。結晶構造から、Cpf1は2つのローブ(RECとNUC)からなることが明らかになった。RECローブはREC1ドメインとREC2ドメインから構成され、NUCローブはRuvCドメイン、WEDドメイン、PIドメイン、Nucドメインから構成されていた。RuvCドメインを除く5つのドメインは新規フォールドをもっていた。crRNAのガイド配列は標的DNAと20塩基のRNA:DNAヘテロ2本鎖を形成し、RECローブとNUCローブの間に結合していた。一方、crRNAの他の領域は予想外のシュードノット構造を形成し、WEDドメインとRuvCドメインによって認識されていた。PAM(TTTA)は歪んだ2重らせん構造をとり、WEDドメイン、REC1ドメイン、PIドメインと相互作用していた。PAMの連

続したT塩基はPIドメインのLys607と水素結合していた。結晶構造と変異体解析の結果、NucドメインはDNA切断に関与することが示唆された。Cas9とCpf1の構造比較から2つのCRISPR-Cas酵素の間の機能的な収斂が明らかになった。さらに、LbCpf1-crRNA-DNA複合体^[19]およびAsCpf1改変体-crRNA-DNA複合体^[20]の結晶構造を決定し、Cpf1のPAM認識機構の詳細を明らかにした。

謝辞

X線回折実験は、SPring-8のBL41XU(課題番号:2015A0119)において行った。本研究は、JST戦略的創造研究推進事業(さきがけ)、JSPS、文部科学省および国立研究開発法人日本医療研究開発機構、内閣府戦略的イノベーション創造プログラム(SIP)の支援を受けて行った。

参考文献

- [1] R. Barrangou *et al.*: *Science* **315** (2007) 1709-1712.
- [2] K. S. Makarova *et al.*: *Nat Rev Microbiol* **13** (2015) 722-736.
- [3] H. Nishimasu and O. Nureki: *Curr Opin Struct Biol* **43** (2017) 68-78.
- [4] G. Gasiunas, R. Barrangou, P. Horvath and V. Siksnys: *Proc Natl Acad Sci USA* **109** (2012) E2579-2586.
- [5] M. Jinek *et al.*: *Science* **337** (2012) 816-821.
- [6] L. Cong *et al.*: *Science* **339** (2013) 819-823.
- [7] H. Nishimasu *et al.*: *Cell* **156** (2014) 935-949.
- [8] M. Jinek *et al.*: *Science* **343** (2014) 1247-997.
- [9] F. Jiang, K. Zhou, L. Ma, S. Gressel and J. A. Doudna: *Science* **348** (2015) 1477-1481.
- [10] C. Anders, O. Niewoehner, A. Duerst and M. Jinek: *Nature* **513** (2014) 569-573.

- [11] F. Jiang *et al.*: *Science* **351** (2016) 867-871.
- [12] F. A. Ran *et al.*: *Nature* **520** (2015) 186-191.
- [13] E. Kim *et al.*: *Nature Commun* **8** (2017) 14500.
- [14] B. Zetsche *et al.*: *Cell* **163** (2015) 759-771.
- [15] H. Nishimasu *et al.*: *Cell* **162** (2015) 1113-1126.
- [16] H. Hirano *et al.*: *Cell* **164** (2016) 950-961.
- [17] M. Yamada *et al.*: *Mol Cell* **65** (2017) 1109-1121.
- [18] T. Yamano *et al.*: *Cell* **165** (2016) 949-962.
- [19] T. Yamano *et al.*: *Mol Cell* **67** (2017) 633-645.
- [20] H. Nishimasu *et al.*: *Mol Cell* **67** (2017) 139-147.

西増 弘志 NISHIMASU Hiroshi

東京大学大学院 理学系研究科
〒113-0033 東京都文京区本郷 7-3-1
TEL : 03-5841-4391
e-mail : nisimasu@bs.s.u-tokyo.ac.jp

滯木 理 NUREKI Osamu

東京大学大学院 理学系研究科
〒113-0033 東京都文京区本郷 7-3-1
TEL : 03-5841-4391
e-mail : nureki@bs.s.u-tokyo.ac.jp