

長期利用課題報告 1

ゲノム編集ツール CRISPR-Cas ヌクレアーゼの構造解析

東京大学大学院 理学系研究科
濡木 理

Abstract

原核生物のもつ CRISPR-Cas 機構に由来する様々な Cas タンパク質は、ガイド RNA と相補的な核酸を特異的に切断する。この性質を利用したものとして Cas9 の 2 本鎖 DNA 切断によるゲノム編集が挙げられるが、近年では Cas13 によるコロナウイルスの高感度・高速検出技術が開発されるなど^[1]、その応用範囲を拡大している。本長期課題では応用的に重要視されている CRISPR-Cas 酵素の結晶構造を決定し、その RNA 依存性核酸切断機構を明らかにすることに成功した。

1. CRISPR-Cas 機構

CRISPR (clustered regularly interspaced short palindromic repeat) -Cas (CRISPR associated) 機構は原核生物のもつ獲得免疫機構であり、ファージやプラスミドなどに由来する外来核酸に対する防御を行う^[2]。CRISPR-Cas 機構は Cas タンパク質と crRNA (CRISPR RNA) から構成される。Cas タンパク質は crRNA と複合体を形成し、crRNA のガイド配列と相補的な外来核酸を認識し、切断する。CRISPR-Cas 機構は哺乳類細胞内でも機能することから、任意の配列の核酸を標的可能なゲノム編集ツールとして注目されている。CRISPR-Cas 機構はそれを構成する遺伝子の種類などによりさまざまなタイプに分類されるが、その中でも特に単一のタンパク質が標的核酸の切断を担うことから、クラス 2 に分類される II 型、V 型、VI 型の CRISPR-Cas 機構がゲノム編集ツールとして利用されている^[3,4]。

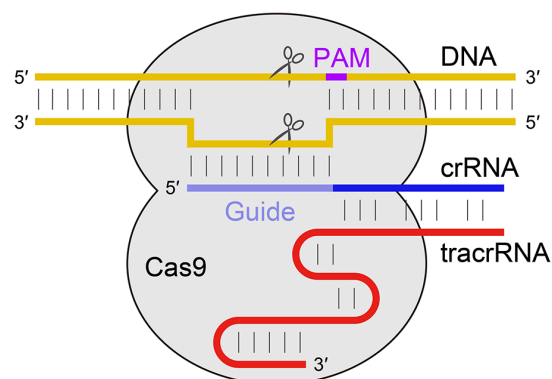


図1 Cas9 による標的 DNA 切断機構

ゲノム編集に広く利用されているのは II 型の CRISPR-Cas 機構で機能する、Cas9 とよばれている Cas タンパク質である。Cas9 は crRNA に加えて tracrRNA とよばれる RNA 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^[5,6] (図 1)。crRNA と tracrRNA を人工的に接続した sgRNA も同様の機能をもつことが知られている。Cas9 による DNA の認識にはガイド配列と標的 DNA の相補性に加え、その近傍に PAM (protospacer adjacent motif) とよばれる特定の配列が必要である。また、Cas9 は分子系統的に II-A 型、II-B 型、II-C 型に分類され、さらにその中でも配列相動性が低く、由来する生物種によって複合体を形成する RNA の配列、認識する PAM 配列、分子サイズなどが多様であることが知られている^[7]。本研究グループはこれまでに様々な生物種に由来する Cas9 の結晶構造解析を行うことにより、その基礎的な分子メカニズムを解明してきた。まず、ゲノム編集ツールとして最もよく利用されている II-A 型の *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た^[8]。次いで、II-A 型の *Staphylococcus aureus* 由来 Cas9 (SaCas9)、II-B 型の *Francisella novicida* 由来 Cas9 (FnCas9)、II-C 型の *Campylobacter jejuni* 由来 Cas9 (CjCas9) の sgRNA および DNA との複合体の結晶構造解析に成功した^[9-11]。これら一連の構造解析により、Cas9 による RNA 依存性 DNA 切断機構のみな

らず、多様な PAM 認識機構および RNA 認識機構などを明らかにしてきた。

本長期課題では、新規ゲノム編集ツールとして期待されている Cas タンパク質に焦点を当てて研究を推進した。CRISPR-Cas 機構を利用したゲノム編集の問題点として、PAM 配列の必要性による標的可能な配列の制限と、分子サイズが大きいことによる細胞内への導入効率の低さが挙げられる。例として、ゲノム編集に広く利用されている SpCas9 は 5'-NGG-3' (N は任意の塩基) という配列のみを認識し、1368 アミノ酸からなる比較的大きいタンパク質である。本研究ではそれらの課題を部分的に解決するものとして期待されている SpCas9 改変体 (SpCas9-NG)、*Brevibacillus laterosporus* 由来 Cas9 (BICas9) および *Corynebacterium diphtheriae* 由来 Cas9 (CdCas9)、加えて、Cas9 と異なり配列依存的に 1 本鎖 RNA と結合する VI 型の Cas タンパク質である Cas13b の結晶構造解析を行った。

2. SpCas9-NG の結晶構造解析

SpCas9 の結晶構造から 5'-NGG-3' の 2 塩基目および 3 塩基目の G はそれぞれ Arg1333 および Arg1335 により認識され、PAM を含む 2 本鎖 DNA は PI (PAM-interacting) ドメインによって形成される正電荷を帯びた溝に結合することが明らかにされていた^[12]。そこで、我々はこの Arg1335 を変異させ G 塩基との塩基特異的な相互作用を無くすと同時に、2 本鎖 DNA の糖-リン酸骨格へ相互作用する変異を PI ドメインに導入することにより、5'-NG-3' を PAM として認識する SpCas9 が得られると考えた。この仮説に基づき様々な変異体を設計し、DNA 切断活性を測定したところ、

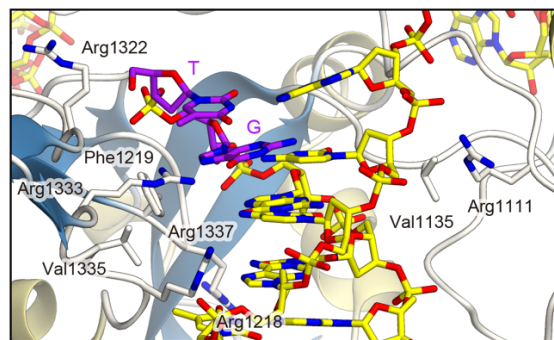


図2 SpCas9-NG による PAM を含む 2 本鎖 DNA 認識機構

Arg1335Val 置換、Leu1111Arg 置換、Asp1135Val 置換、Gly1218Arg 置換、Glu1219Phe 置換、Ala1322Arg 置換、Thr1337Arg 置換の 7 つの変異をもつ SpCas9 は効率よく 5'-TGN-3' PAM をもつ DNA を切断することが明らかとなった^[13]。さらに、この 7 つの変異をもつ SpCas9 を用いて、65,536 (4 の 8 乗) 通りの配列をもつ DNA ライブラリーを *in vitro* において切断し次世代シーケンサーを用いて解析した結果、僅かな嗜好性があるものの 5'-NG-3' PAM を認識することが確かめられ、これを SpCas9-NG と命名した。SpCas9-NG による PAM の認識機構を解明するために、SpCas9-NG、sgRNA および標的 DNA からなる複合体を結晶化し、BL41XU において X 線回折データを収集し、SpCas9-sgRNA-標的 DNA 複合体の結晶構造をサーチモデルとした分子置換法により結晶構造を決定した。結晶構造から、予想通り Arg1335Val の置換により PAM の 3 番目の G 塩基との塩基特異的な相互作用は失われ、Val1135 および Phe1219 は糖骨格と疎水性相互作用を、Arg1111 はリン酸骨格と静電相互作用をしていることが明らかになった (図2)。その一方で Arg1218 と Arg1337 は予想とは異なり、Arg1218 は周辺のアミノ酸の主鎖と水素結合を形成していた一方で、Arg1337 は結晶化に用いた DNA の塩基に特異的な水素結合を形成していた。このことから、Arg1218 は PI ドメインの構造安定化に寄与し、Arg1337 は SpCas9-NG の PAM 配列の嗜好性に関与していることが示唆された。

3. 小型 Cas9 の結晶構造解析

B.lacterosporus 由来 Cas9 (BICas9) および *C.diphtheriae* 由来 Cas9 は II-C 型に属する Cas9 で小型 (約 1000 アミノ酸残基) である^[14,15]。BICas9 は 5'-NNNNC-3' と C 塩基 1 文字を認識するのに対し、CdCas9 は 5'-NNRHHHY-3' (R=A/G, H=A/C/T, Y=C/T) と緩い特異性で長い PAM を認識する。これらの Cas9 による RNA 依存性 DNA 切断機構を解明するために、BICas9 および CdCas9 の sgRNA および標的 DNA を含む複合体をそれぞれ結晶化し、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)。BICas9 および CdCas9 は既存の小型の Cas9

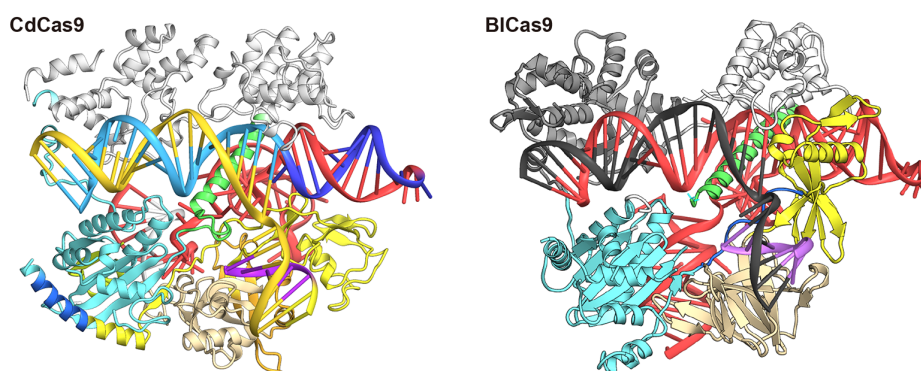


図3 CdCas9 および BICas9 の全体構造

と全体構造は類似していたものの、異なる配列をもつ RNA を特異的に認識し複合体を形成していることが明らかになった。また、構造から各々の特異的な PAM 認識機構が明らかとなった。BICas9 は Asp1023 と Lys1041 がそれぞれ C 塩基と G 塩基と水素結合を形成することにより強固に G:C 塩基対を認識しており、これにより 1 塩基のみの PAM 認識を可能にしていることが示唆された。一方で、CdCas9 の PI ドメインには Phe1011、Lys1015、Pro1043、Leu1046 などから形成される疎水性パッチが存在し、これが PAM を含む 2 本鎖 DNA の主溝に接近していた。この疎水性パッチによるファンデルワールス相互作用によって、緩い特異性の PAM 認識が可能になっていることが示唆された。さらに、SpCas9-NG と同様の戦略により BICas9 の結晶構造を基にして PI ドメインに変異を入れ、異なる塩基配列を PAM として認識する改変型 BICas9 の開発に成功した。

4. Cas13bt の結晶構造

近年、ゲノムデータベースの拡大によって様々な新規 Cas タンパク質が発見されており、中でも Cas13b (Cas13bt3) (775 残基) は小型で crRNA 依存的に 1 本鎖 RNA を切断する機能があることで注目されていた^[17]。本研究では、Cas13b と crRNA からなる複合体を結晶化し、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 1.9 Å 分解能で結晶構造を決定した (図 4)。得られた構造から、Cas13bt3 は REC (recognition) ロープと NUC (nuclease) ロープからなり、crRNA が REC ロープにアンカーされていることが明らかとなった。

REC ロープは Helical-1 ドメイン、Lid ドメイン、Helical-2 ドメインから構成され、NUC ロープは HEPN1 ドメイン、および HEPN2 ドメインから構成されていた。crRNA はガイド領域、2つのステム領域 (Stem1、Stem2)、および2つのループ領域 (internal loop, hairpin loop) から構成されており、Cas13bt3 の REC ロープによって認識されていた。特に、crRNA からフリップした3つの塩基、C (-8)、G (-28)、A (-32)は Helical-1 ドメイン、Lid ドメイン、Helical-2 ドメインによって塩基特異的に認識されていた。Cas13bt3 はすでに構造が報告されている PbuCas13b と比較して 350 塩基小さく、HEPN1 ドメインを除いて配列相同性をもたないにも関わらず、全体構造は PbuCas13b と類似していた (RMSD: 4.31 Å)。ドメインごとの比較から、Cas13bt3 はすべてのドメインが PbuCas13b よりもそれぞれ 30-100 塩基ほど小さくなっていた。したがって、これらの違いが Cas13bt3 の小型化に貢献していることが明らかとな

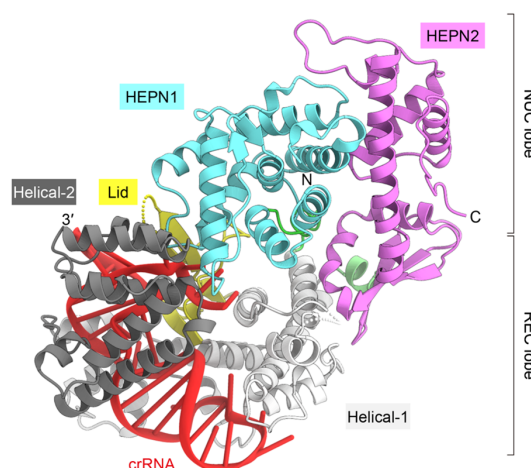


図4 Cas13bt3 の全体構造

った。今回の結果は、CRISPR-Cas システムにおける共通性および多様性を明らかにするとともに、新たな革新的技術の開発基盤となることが期待される。本研究成果は *Molecular Cell* 誌に掲載予定である。

参考文献

- [1] H. Shinoda *et al.*: *Commun. Biol.* **4** (2021) 476.
- [2] R. Barrangou *et al.*: *Science* **315** (2007) 1709-1712.
- [3] K. S. Makarova *et al.*: *Nat. Rev. Microbiol.* **13** (2015) 722-736.
- [4] H. Nishimasu and O. Nureki: *Curr. Opin. Struct. Biol.* **43** (2017) 68-78.
- [5] G. Gasiunas, R. Barrangou, P. Horvath and V. Siksnys: *Proc. Natl. Acad. Sci. U.S.A.* **109** (2012) E2579-2586.
- [6] M. Jinek *et al.*: *Science* **337** (2012) 816-821.
- [7] K. Chylinski *et al.*: *Nucleic Acids Res.* **42** (10) (2014) 6091-105.
- [8] H. Nishimasu *et al.*: *Cell* **156** (2014) 935-949.
- [9] H. Nishimasu *et al.*: *Cell* **162** (2015) 1113-1126.
- [10] H. Hirano *et al.*: *Cell* **164** (2016) 950-961.
- [11] M. Yamada *et al.*: *Mol. Cell* **65** (2017) 1109-1121.
- [12] M. Jinek *et al.*: *Nature* **513** (2014) 569-573.
- [13] H. Nishimasu *et al.*: *Science* **361** (2018) 1259-1262.
- [14] T. Karvelis *et al.*: *Genome Biol.* **16** (2015) 253.
- [15] F. A. Ran *et al.*: *Nature* **520** (2015) 186-191.
- [16] S. Hirano *et al.*: *Nature Commun* **10** (2019) 1968.
- [17] S. Kannan *et al.*: *Nat Biotechnol* **40** (2) (2021) 194-197.

瀨木 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