

高度好熱菌 *Thermus thermophilus* 由来 Type III-B CRISPR-Cmr 複合体の構造と機能

新海 暁男

1. はじめに

CRISPR (clustered regularly interspaced short palindromic repeat)-Cas (CRISPR-associated) システムは多くの細菌が持つ獲得免疫システムである¹⁾ (図1に概要を示す)。これは細菌に侵入してきたプラスミドやファージのDNAあるいはRNAを異物として認識し分解して除去するシステムであり、CRISPRと呼ばれるDNAの反復配列と複数のCasタンパク質とから構成されている。典型的なCRISPRは約25~40塩基の回文様配列(リピート)が約25~40塩基のスペーサー配列を介して数回から20回程度繰り返している領域である。一方、Casタンパク質はこれまでに約45種類が知られており、それらをコードする遺伝子の多くはCRISPRの近傍に位置している。CRISPRの発見は28年前の1987年に遡る。その18年後の2005年、ファージ遺伝子などの外来遺伝子の部分配列がCRISPRのスペーサー領域に見つかり、CRISPR-Casシステムが獲得免疫に関与している可能性が示唆された。そして、2年後にはそれが実験的に証明された。すなわち、細菌に侵入してきたDNAの一部と同じ配列をスペーサーとして細菌のゲノムが持っている場合、その侵入DNAはCRISPR-Casシステムによって分解されることが示された。

CRISPR-Casシステムは、Adaptation, Expression, Interferenceの三つのフェーズから成り立っている¹⁾ (図1)。Adaptationフェーズでは、Casタンパク質が侵入DNAの一部を切り出し、スペーサー配列としてCRISPRに組み込む。これによって細菌は侵入してきたDNAを“記憶”する。Expressionフェーズでは、CRISPRが転写されて生じたpre-CRISPR RNA (pre-crRNA) がリピート部分で切断されスペーサー単位のcrRNAが生成される。そして、crRNAが他のCasタンパク質(単体あるいは複合体)に結合す

る。このCasタンパク質-crRNA複合体がcrRNAの部分で侵入DNAあるいはRNAに相補的に結合しそれを分解するステップがInterferenceフェーズである。CRISPR-Casシステムは、主としてInterferenceフェーズに関与しているCasタンパク質の種類や作用機作の違いによってType I, -II, -IIIの3種類に分類され、それぞれのタイプはさらにいくつかのサブタイプに細分類されている²⁾。Type IシステムではCascade (CRISPR-associated complex for antiviral defense) と呼ばれるタンパク質複合体およびCas3タンパク質が、Type IIシステムではCas9タンパク質およびtracr (trans-activating CRISPR) RNAと呼ばれる低分子RNAが、Type III-AシステムではCsmタンパク質複合体が、Type III-BシステムではCmrタンパク質複合体がそれぞれcrRNAと結合し、それらのCasタンパク質-crRNA複合体が標的DNAあるいはRNAの分解をつかさどっている。これらのうち、Cas9を用いたシステムはさまざまな生物種において利用可能なゲノム編集のための優れたツールとして注目を集めている。

CRISPRの数、リピートの数と塩基配列、cas遺伝子の種類と数は細菌ごとに異なっている。高度好熱菌 *Thermus thermophilus* HB8株はゲノムサイズが約2 Mbpと比較的小さいにも関わらず比較的多くのCRISPRと多種類のCasタンパク質を持っている³⁾ (図2)。CRISPRはプラスミドpTT27上に9か所(CRISPR-1~7, -9, -10)と染色体上に2か所(CRISPR-11, -12)ある。これらのCRISPRはリピート部分の塩基配列の違いによって3種類(Category-I, -II, -III)に分類されている。一方、本菌株が有するCasタンパク質は、Type I-E (Cascade), Type III-A (Csm), Type III-B (Cmr) 複合体を含む約30種類に及ぶ。このため本菌株は一つの細菌が持つCRISPR-Casシステムを体系的に理解するための格好の研究対象である。本稿では、最近明らかとなった、Type III-Bシステムを構成しているCmr複合体の構造と機能を概説する。

2. *T. thermophilus* HB8株のType III-Bシステムを構成する遺伝子群

Type III-Bを構成しているcmr遺伝子 (cmr1~6) はCRISPR-5と-6の間に位置しオペロンを形成している

理化学研究所横山構造生物学研究室 (〒230-0045 神奈川県横浜市鶴見区末広町1-7-22)

Structure and function of Type III-B CRISPR-Cmr complex from *Thermus thermophilus*

Akeo Shinkai (RIKEN Structural Biology Laboratory, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan)

DOI: 10.14952/SEIKAGAKU.2015.870454

© 2015 公益社団法人日本生化学会

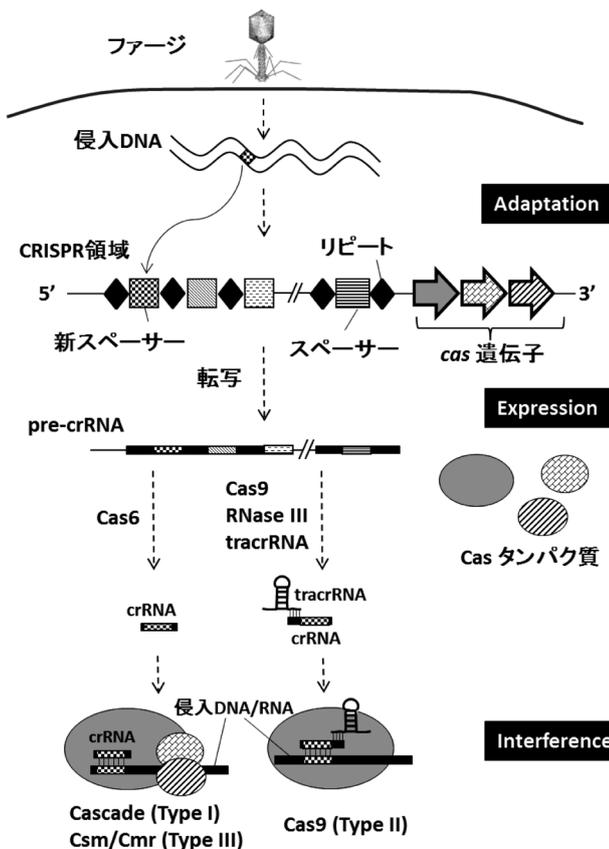


図1 CRISPR-Casシステムの概要

本システムはCRISPRと呼ばれるDNA領域とcas遺伝子群から構成されている。CRISPRは約25~40塩基の回文様配列(リピート)が約25~40塩基のスペーサー配列を介して数回から20回程度繰り返している領域である。このシステムはAdaptation, Expression, Interferenceの三つのフェーズからなる。Adaptationフェーズでは細菌に侵入したDNAの一部を切り取りスペーサーとしてCRISPRに組み込む。ExpressionフェーズではCRISPRが転写されて生じたpre-crRNAが切断され、スペーサー単位のcrRNAが生じる。InterferenceフェーズではcrRNAと結合したある種のCasタンパク質(単体あるいは複合体)が侵入DNAまたはRNAを分解する。本システムは、InterferenceフェーズにCascade複合体が関与しているType I, Cas9/tracrRNAが関与しているType II, Csm複合体あるいはCmr複合体が関与しているType IIIに分類できる。

(図2)。ファージが感染するとcmrオペロンの転写量が上昇するが、その発現を制御している因子はまだ同定されていない³⁾。Type III-BシステムではCas6タンパク質がpre-crRNAの切断を担っている(図1)。Cas6は各リピート部分の3'末端から8塩基上流を切断するので、crRNAは5'末端側にリピート由来の8塩基(5'タグ)を持っている(後述)。T. thermophilus HB8株は二つのcas6遺伝子(染色体上のcas6A, および、pTT27上のcas6B)を持っており(図2)、Cas6A, -BともにCategory-IあるいはCategory-IIIに分類されるリピート配列を特異的に切断する⁴⁾。

プラスミドpTT27

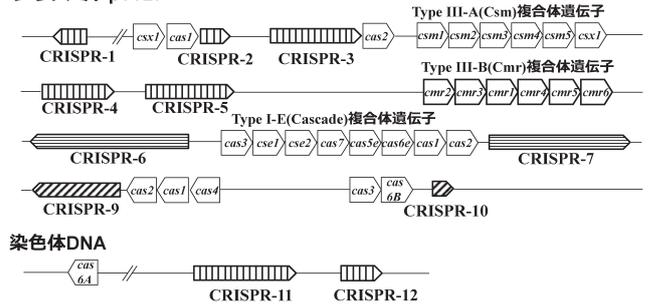


図2 T. thermophilus HB8株のCRISPR-cas遺伝子群
CRISPRはリピートの塩基配列によってCategory-I(縦線), -II(横線), および, -III(斜線)に分類できる。

3. Cmr複合体の構造と機能

T. thermophilus HB8株のCmr複合体(TtCmr)は6種類のCmrタンパク質(Cmr1~6)とcrRNAから構成されている約360kDaの分子で、サブユニットの構成比はCmr₁₂₃₄₅₆:crRNA₁(下付きの数字は分子数)である^{5,6)}。電子顕微鏡を用いた単粒子解析によると、TtCmrの立体構造は約90×100×200Åの大きさの右巻きのらせん形をしている(図3)。TtCmr分子の中央部には4分子のCmr4と3分子のCmr5がらせん状に位置しており、片側の端にはCmr2-Cmr3複合体が結合し、もう片方の端にはCmr6とCmr1が結合している。crRNAは、5'タグ配列(上述)でCmr2-Cmr3複合体に結合し、スペーサー由来の部分でCmr4に沿って中央部に巻きついている。Cmr1が欠失したと考えられる分子や、Cmr4とCmr5がそれぞれ1分子欠失したと考えられる分子も見つかっている^{5,6)}。

T. thermophilus HB8株のCRISPRには合計112種類のスペーサー配列があるが、TtCmrに結合しているcrRNAにはこれらの配列が一様に見いだされるのではなく、Category-Iに分類されるCRISPR由来のスペーサーに偏っている⁵⁾。特に、CRISPR-1のスペーサー3, CRISPR-2のスペーサー1, 2, CRISPR-4のスペーサー2, 3, 5, CRISPR-11のスペーサー1, 3, 8, 9, 12が多く見いだされている。これらのスペーサーの塩基配列や予想二次構造に共通性はない。TtCmrがcrRNAを選別するメカニズムの解明は今後の研究課題の一つである。TtCmrに結合しているcrRNAの5'末端にはTtCmrへ結合するために必要な8塩基の5'タグ配列がある。これはpre-crRNAがCas6によって切断されるためであることは上述したとおりである。crRNAの長さは不ぞろいで、およそ30~50塩基であり、特に、34, 40, 46, 49塩基のものが多く見いだされている⁵⁾。pre-crRNAの3'末端側がプロセシングされるメカニズムは不明である。TtCmrは、結合しているcrRNAと相補的な配列を含む一本鎖RNAを、in vitroでcrRNAの5'側から順に3'側へ向かって6塩基ご

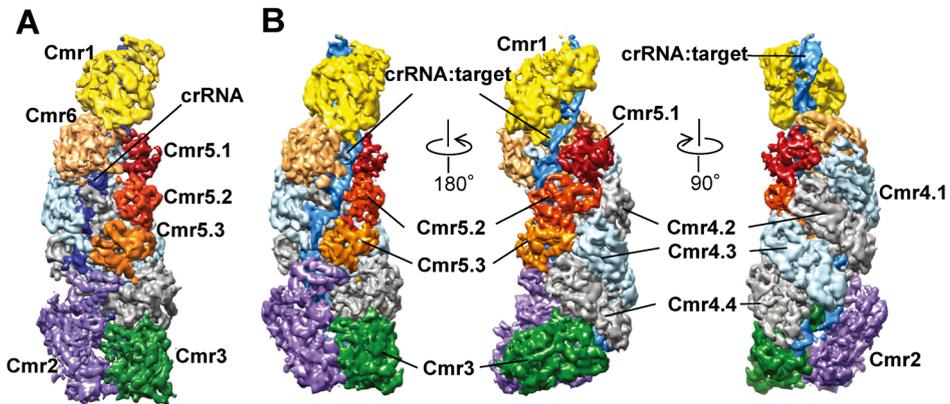


図3 *Thermus thermophilus* Cmr 複合体 (TtCmr) の立体構造

(A) TtCmr のクライオ電子顕微鏡構造 (4.1 Å). (B) 標的 RNA (target) が結合した TtCm のクライオ電子顕微鏡構造 (4.4 Å). 図は文献6から AAAS の許可を得て転載した (カラー図は電子版参照).

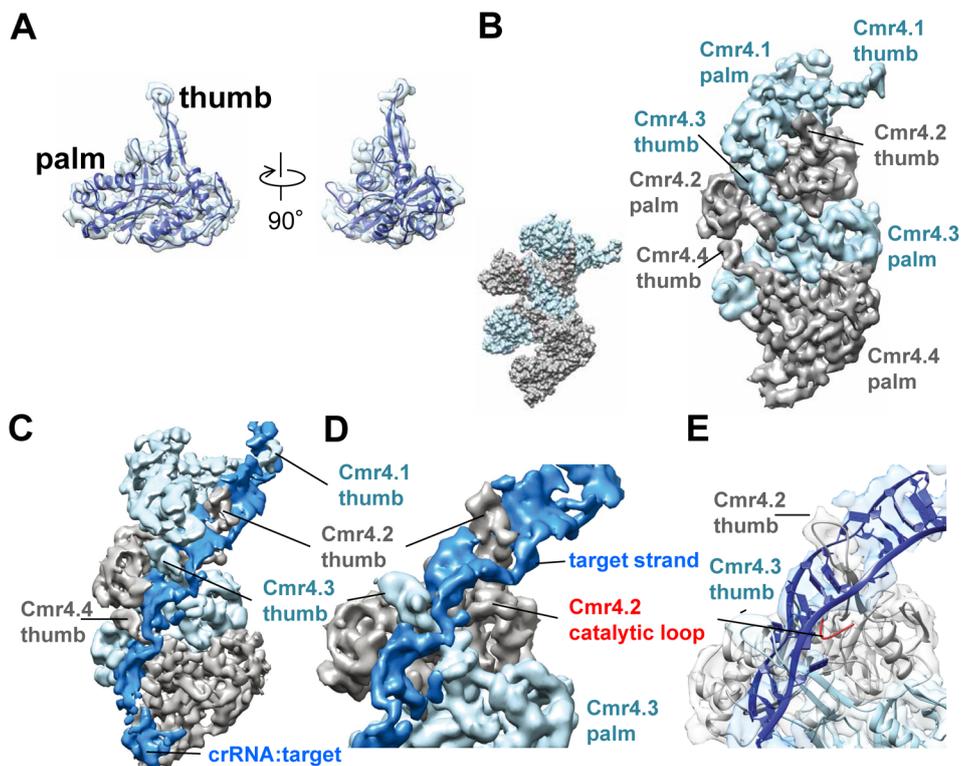


図4 Cmr4サブユニットによる標的RNAの切断メカニズム

(A) Cmr4サブユニットの立体構造. *P. furiosus* 由来の Cmr4⁷⁾ および大腸菌 Cas7 の結晶構造⁹⁾ を基に作製した *T. thermophilus* Cmr4 のホモロジーモデル (リボン図) を TtCmr のクライオ電子顕微鏡構造中の Cmr4サブユニットに当てはめたもの. (B) thumb と palm の相互作用による Cmr4サブユニットの連結. この連結様式は大腸菌 Cascade 複合体における Cas7サブユニットの場合 (左下)⁹⁾ と類似している. (C) 4分子の Cmr4サブユニットの各 thumb ドメインは6塩基ごとに4か所で, crRNA と標的 RNA からなる二本鎖 RNA にインターカレーションしている. (D) (C) の拡大図. thumb ドメインがインターカレーションしている二重鎖 RNA の近傍に推定活性残基 (His16/Asp27) が存在するループ (catalytic loop) が位置していることがわかる. (E) (D) にホモロジーモデル (リボン図) を当てはめたもの. crRNA:target のモデルは大腸菌 Cascade 複合体の crRNA:target ssDNA の構造¹⁰⁾ を基にして作製した. 図は文献6から AAAS の許可を得て転載した (カラー図は電子版参照).

とに5か所で切断する(5'ルーラーメカニズム)⁵⁾。標的RNAを分解する活性残基はCmr4サブユニットに存在することが*Pyrococcus furiosus*由来のCmrで示されているので^{7,8)}、TtCmrの場合も5か所のうちの4か所はCmr4が切断すると考えられる。残りの1か所を切断しているサブユニットと活性残基の同定が急務である。

Cmr複合体の構造をクライオ電子顕微鏡法で4.1Åの分解能で解析した結果、複合体分子の中央部に位置している4分子のCmr4は、突出したβ-ヘアピン構造(thumbドメイン)がそれぞれ隣の分子のα-ヘリックス構造(palmドメイン)と相互作用することによって連結していることがわかった(図3, 図4A, B)。次に、標的RNAが結合したTtCmrの構造を4.4Åの分解能で解析した結果、この複合体分子では、標的RNAが結合しやすくなるようにCmr1, Cmr2, Cmr5分子がコンホメーション変化を起こすことで分子の中央部が開いていた(図3B)⁶⁾。さらに、Cmr4の四つのthumbドメインは、6塩基ごとに4か所で、crRNAと標的RNAからなる二本鎖RNAにインターカレーションし、それによって二本鎖RNAが部分的に解かれていた(図4C~E)。そして、その近傍にはCmr4の予想活性残基(His16とAsp27)があるループ構造が位置していた(図4D, E)。これが、Cmr複合体が一定の間隔で標的RNAを分解する仕組みである。また、Cmr6のthumbドメインも二本鎖RNAにインターカレーションしていたので、もう一つの活性残基はこのサブユニット上にあると考えられる⁶⁾。

一方、クライオ電子顕微鏡法を用いた解析で、完全長のものよりも25Å短い、Cmr4とCmr5がそれぞれ1分子欠失したと考えられるCmrサブ複合体が見つかった⁶⁾。TtCmrに結合しているcrRNAは、長さが46塩基のものと40塩基のもので大半を占めている⁵⁾。長さ25Åは6塩基分のRNAに相当するので、この短いCmrサブ複合体には40塩基のcrRNAが結合しており、完全長の複合体には46塩基のcrRNAが結合していると考えられる。しかし、長さの異なる2種類の複合体が存在する意義は不明である。

興味深いことに、palmドメインとthumbドメインによるサブユニットの連結様式は、二本鎖DNAを分解するシステムである大腸菌Type I-E CRISPRシステムを構成しているCascade複合体のCas7サブユニットにもみられる^{9,10)}(図4B)。この複合体は、5種類のタンパク質とcrRNAから構成されている約405kDaの分子で、サブユニットの構成比はCse1₁Cse2₂Cas5e₁Cas6e₁Cas7₆:crRNA₁(下付きの数字は分子数)である。この分子も右巻きのらせん形であり、分子の中央部にはpalm-thumbドメインの相互作用によって結合している6分子のCas7タンパク質が位置し、そこにcrRNAが巻きついている。Cascade複合体はcrRNA部分で二本鎖DNAの片側の鎖に相補的に結合するが、Cas-

cade複合体自身にはDNAを分解する活性はなく、Cascade-crRNA-標的DNA複合体を認識して結合するCas3タンパク質がそれを担っている¹¹⁾。Cascade-crRNA-一本鎖DNA複合体の構造を解析すると、crRNA:DNAハイブリッド分子は二重らせん構造をしておらず、解けた二重鎖構造を保ち、6塩基ごとに折れ曲がりながら分子の中央部にらせん状に巻きついている。crRNA:DNAの折れ曲がりを誘導し、その構造を維持しているのがCas7のthumbドメインであり、六つのthumbドメインが、Cmr複合体のCmr4_thumbドメインと同様に、crRNA:DNAに6塩基ごとにインターカレーションしている¹⁰⁾。

4. まとめと今後の展望

TtCmrは*in vitro*で一本鎖RNAを6塩基ごとに5か所で切断する。*Sulfolobus islandicus*株由来のCmrは*in vivo*でDNAとRNAの両方を分解することが示されている^{12,13)}。*T. thermophilus*株におけるCmrの役割を解明するために、TtCmrの*in vivo*での活性、基質特異性を解析することは今後の重要な研究課題の一つである。

TtCmr分子の中央部にらせん状に位置している4分子のCmr4サブユニットは、それぞれのthumbドメインが隣の分子のpalmドメインと相互作用することによって連結している。そのCmr4分子に沿ってcrRNAが巻きついており、相補的な結合によって標的RNAがcrRNAに結合する。四つのCmr4_thumbドメインはその二本鎖RNAに6塩基ごとにインターカレーションすることによって部分的に二重鎖を破壊しており、近傍には標的RNAを分解する活性残基が位置している。TtCmrが6塩基ごとに標的RNAを分解するのはこのためである。これらの特徴は古細菌由来のCmr複合体にもみられる¹⁴⁾。しかもCmr複合体にみられるこれらの構造はType I-Eシステムを構成しているCascade複合体にも保存されている。Cascade複合体では、Cas7サブユニットがpalm-thumbドメインの相互作用によって分子の中央部にらせん状に連結しており、それぞれのthumbドメインが6塩基ごとに二本鎖の核酸(crRNA:標的DNA)にインターカレーションしている。さらに興味深いことに、*T. thermophilus*株のType III-Aシステムを構成しているCsm複合体も同様の右巻きのらせん構造をしており、6分子のCsm3が分子の中央部にらせん状に位置していると考えられている¹⁵⁾。そして、CsmもCmrと同様に5'ルーラーメカニズムで6塩基ごとに一本鎖RNAを切断する¹⁵⁾。したがって、Csm複合体ではCsm3がRNAを分解する活性を担っており、この場合もthumbドメインが複合体の構造と機能の維持に重要な役割を果たしていると思像できる。Cascade複合体は、それ自身には標的DNAを分解する活性がない点でCmr複合体やCsm複合体と異なるが、これら

複合体分子の中央部にらせん状に連結しているサブユニット (Cmr4, Cas7, Csm3) は共通の祖先から派生したものと考えられる。 *T. thermophilus* 株の CRISPR-Cas の研究から、細菌における獲得免疫システムそのものや、それを構成している分子群の進化も紐解かれていくことを期待している。

謝辞

この研究は、JSPS 科研費 25440013 の助成を受け、オランダ・ワーゲニンゲン大学の John van der Oost 教授の研究室、アメリカ・カリフォルニア大学バークレー校の Jennifer Doudna 教授の研究室をはじめとする多くの研究室との共同で行われたものです。本研究に携わっていただいた多くの方々に深く感謝致します。

文 献

- 1) Karginov, F.V. & Hannon, G.J. (2010) *Mol. Cell*, **37**, 7–19.
- 2) Makarova, K.S., Haft, D.H., Barrangou, R., Brouns, S.J., Charpentier, E., Horvath, P., Moineau, S., Mojica, F.J., Wolf, Y.I., Yakunin, A.F., van der Oost, J., & Koonin, E.V. (2011) *Nat. Rev. Microbiol.*, **9**, 467–477.
- 3) Agari, Y., Sakamoto, K., Tamakoshi, M., Oshima, T., Kuramitsu, S., & Shinkai, A. (2010) *J. Mol. Biol.*, **395**, 270–281.
- 4) Niewoehner, O., Jinek, M., & Doudna, J.A. (2014) *Nucleic Acids Res.*, **42**, 1341–1353.
- 5) Staals, R.H., Agari, Y., Maki-Yonekura, S., Zhu, Y., Taylor, D.W., van Duijn, E., Barendregt, A., Vlot, M., Koehorst, J.J., Sakamoto, K., Masuda, A., Dohmae, N., Schaap, P.J., Doudna, J.A., Heck, A.J., Yonekura, K., van der Oost, J., & Shinkai, A. (2013) *Mol. Cell*, **52**, 135–145.
- 6) Taylor, D.W., Zhu, Y., Staals, R.H., Kornfield, J., Shinkai, A., van der Oost, J., Nogales, E., & Doudna, J.A. (2015) *Science*, **348**, 581–585.
- 7) Benda, C., Ebert, J., Scheltema, R.A., Schiller, H.B., Baumgartner, M., Bonneau, F., Mann, M., & Conti, E. (2014) *Mol. Cell*, **56**, 43–54.
- 8) Ramia, N.F., Spilman, M., Tang, L., Shao, Y., Elmore, J., Hale, C., Cocozaki, A., Bhattacharya, N., Terns, R.M., Terns, M.P., Li, H., & Stagg, S.M. (2014) *Cell Reports*, **9**, 1610–1617.
- 9) Jackson, R.N., Golden, S.M., van Erp, P.B., Carter, J., Westra, E.R., Brouns, S.J., van der Oost, J., Terwilliger, T.C., Read, R.J., & Wiedenheft, B. (2014) *Science*, **345**, 1473–1479.
- 10) Mulepati, S., Heroux, A., & Bailey, S. (2014) *Science*, **345**, 1479–1484.
- 11) Westra, E.R., van Erp, P.B., Kunne, T., Wong, S.P., Staals, R.H., Seegers, C.L., Bollen, S., Jore, M.M., Semenova, E., Severinov, K., de Vos, W.M., Dame, R.T., de Vries, R., Brouns, S.J., & van der Oost, J. (2012) *Mol. Cell*, **46**, 595–605.
- 12) Deng, L., Garrett, R.A., Shah, S.A., Peng, X., & She, Q. (2013) *Mol. Microbiol.*, **87**, 1088–1099.
- 13) Peng, W., Feng, M., Feng, X., Liang, Y.X., & She, Q. (2014) *Nucleic Acids Res.*, **43**, 406–417.
- 14) Osawa, T., Inanaga, H., Sato, C., & Numata, T. (2015) *Mol. Cell*, **58**, 418–430.
- 15) Staals, R.H., Zhu, Y., Taylor, D.W., Kornfeld, J.E., Sharma, K., Barendregt, A., Koehorst, J.J., Vlot, M., Neupane, N., Varrassieu, K., Sakamoto, K., Suzuki, T., Dohmae, N., Yokoyama, S., Schaap, P.J., Urlaub, H., Heck, A.J., Nogales, E., Doudna, J.A., Shinkai, A., & van der Oost, J. (2014) *Mol. Cell*, **56**, 518–530.

著者寸描

●新海 暁男 (しんかい あけお)



理化学研究所横山構造生物学研究室前任研究員。農学博士。

■略歴 1991年東京大学大学院農学系研究科博士後期課程修了・農学博士。91～2002年協和醗酵工業東京研究所研究員。99～01年ワシントン大学客員研究員。02年より現職。

■研究テーマと抱負 細菌が持つ環境応答の仕組み、他生物とのコミュニケーション

や相互作用の仕組みを分子のレベルで解明し、「細菌社会」に関する新たな知見を得る。

■趣味 家庭菜園、散歩。