6074–6082 *Nucleic Acids Research, 2006, Vol. 34, No. 21* *Published online 11 November 2006*
doi:10.1093/nar/gkl729

Structural basis for sequence-dependent recognition of colicin E5 tRNase by mimicking the mRNA–tRNA interaction

Shunsuke Yajima*, Sakura Inoue[1], Tetsuhiro Ogawa[1], Takamasa Nonaka[2], Kanju Ohsawa and Haruhiko Masaki[1,*]

Department of Bioscience, Tokyo University of Agriculture, Sakuragaoka 1-1-1, Setagaya-ku, Tokyo 156-8502, Japan, [1]Department of Biotechnology, Graduate School of Agricultural and Life Sciences, The University of Tokyo, Yayoi 1-1-1, Bunkyo-ku, Tokyo 113-8657, Japan and [2]Department of Bioengineering, Nagaoka University of Technology, Nagaoka, Niigata 940-2188, Japan

Received May 15, 2006; Revised September 2, 2006; Accepted September 15, 2006

ABSTRACT

Colicin E5—a tRNase toxin—specifically cleaves QUN (Q: queuosine) anticodons of the *Escherichia coli* tRNAs for Tyr, His, Asn and Asp. Here, we report the crystal structure of the C-terminal ribonuclease domain (CRD) of E5 complexed with a substrate analog, namely, dGpdUp, at a resolution of 1.9 Å. This structure is the first to reveal the substrate recognition mechanism of sequence-specific ribonucleases. E5-CRD realized the strict recognition for both the guanine and uracil bases of dGpdUp forming Watson–Crick-type hydrogen bonds and ring stacking interactions, thus mimicking the codons of mRNAs to bind to tRNA anticodons. The docking model of E5-CRD with tRNA also suggests its substrate preference for tRNA over ssRNA. In addition, the structure of E5-CRD/dGpdUp along with the mutational analysis suggests that Arg33 may play an important role in the catalytic activity, and Lys25/Lys60 may also be involved without His in E5-CRD. Finally, the comparison of the structures of E5-CRD/dGpdUp and E5-CRD/ImmE5 (an inhibitor protein) complexes suggests that the binding mode of E5-CRD and ImmE5 mimics that of mRNA and tRNA; this may represent the evolutionary pathway of these proteins from the RNA–RNA interaction through the RNA–protein interaction of tRNA/E5-CRD.

INTRODUCTION

Colicins, which are produced by bacteria carrying the corresponding Col plasmids, kill sensitive *Escherichia coli* cells through activities involving the ion channels in the inner membranes, DNases or RNases (1,2). The DNase-type colicins nonspecifically degrade the genomic DNAs of the sensitive cells (3–5), whereas those of the RNase type cleave 16S rRNA at the 49th phosphodiester bond from the 3′ end (6–8). These colicins contain three domains—a membrane-translocating domain, a receptor-binding domain, and a catalytic domain—in their primary sequences. Colicinogenic cells also produce inhibitor proteins (Imm), the genes for which are located downstream of the *col* genes in the colicin operons that are under the control of the SOS-dependent promoters. Imms bind specifically to cognate colicins in order to protect their host cells.

In addition to the known RNase-type colicins, colicin E5 that has been recently characterized as a tRNase has a unique target for its toxicity. It specifically cleaves the anticodons of *E.coli* tRNAs for Tyr, His, Asn and Asp. These tRNAs decode NAY (N: any nucleotide, A: adenosine and Y: a pyrimidine nucleotide) codons by means of QUN anticodons, where queuosine (Q) is a 7-deazaguanosine with a cyclopentenediol side chain (9). Initially, colicin E5 was observed to cleave only the QU sequence of these tRNAs between positions 34 and 35, leaving a 2′,3′-cyclic phosphate on Q and a 5′-OH on U. In tRNA-guanine transglycosylase-deficient strains the inherent G is not replaced with a precursor base of Q; this gives rise to tRNAs with GUN anticodons. These strains were demonstrated to be sensitive to colicin E5; the tRNAs were still subjected to cleavage (10). It is assumed that colicin E5 consists of 556 amino acids (11). The ribonuclease activity of E5 resides solely within its C-terminal ribonuclease domain (E5-CRD; 115 amino acids), and this domain is responsible for its substrate specificity. Furthermore, based on the results obtained from our assay using synthetic minihelices or linear oligoribonucleotides, E5-CRD can be referred to as an RNA restriction enzyme that

*To whom correspondence should be addressed. Tel: +81 3 5477 2768; Fax: +81 3 5477 2668; Email: yshun@nodai.ac.jp
*Correspondence may also be addressed to Haruhiko Masaki. Tel: +81 3 5841 8248; Fax: +81 3 5841 8248; Email: hmasaki@mcb.bt.a.u-tokyo.ac.jp

specifically recognizes and cleaves single-stranded GU sequences (12).

In addition to colicin E5, colicin D (13), PrrC (14) and zymocin (15) are also known to be tRNases. However, these enzymes are nonhomologous with each other and target tRNAs and cleavage sites that are different from those targeted by colicin E5. Colicin D, which is produced by *E.coli* harboring a ColD plasmid, cleaves only four isoacceptors of $tRNA^{Arg}$ between positions 38 and 39 at the 3′ junction of the anticodon stem and loop. PrrC is induced in some *E.coli* isolates by T4 phage infection, and it only cleaves $tRNA^{Lys}$ between positions 33 and 34 at the 5′ side of the anticodon. Zymocin, which is produced by *Kluyveromyces lactis*, was recently reported to cleave some yeast tRNAs at the 3′ side of the modified nucleotide U in the anticodons of 5-methoxycarbonylmethyl-2-thiouridine uridine cytosine (mcm^5s^2UUC), mcm^5s^2UUU and mcm^5s^2UUG. Among these tRNases, the crystal structure of the C-terminal catalytic domain of colicin D (D-CRD) has been solved (16,17); the structure revealed a curved row of basic residues on the molecular surface that could confer the tRNA recognition specificity.

ImmE5—a specific inhibitor protein of colicin E5—is expressed in host cells and binds to E5-CRD to prevent cell death. In addition to the nuclease-type colicin family, this type of proteinaceous toxin–antitoxin system (comprising addiction molecules), including MazE/F (18) and RelB/E (19) whose targets are considered to be the ACA sequences in mRNAs and stop codons of mRNAs in the ribosomal A site, respectively, has also been well studied. The study of these molecules reveals the survival mechanism in bacteria; additionally, they serve as very good examples of protein–protein interactions. Furthermore, from the structural viewpoint, it is of special interest to determine whether ribonuclease inhibitors, i.e. inhibitors of toxins that target DNA or RNA, bind to the corresponding enzymes by mimicking the substrate RNAs/DNAs.

Further, although E5-CRD exhibits ribonuclease activity, it lacks His that usually serves as a catalytic residue in all known ribonucleases. Thus, E5-CRD must have an alternative mechanism that is responsible for its ribonuclease activity.

Since E5-CRD shows no homology with any other proteins, we performed the structural analysis of E5-CRD in order to explore the novel features of colicin E5. Recently, Huang's group has reported the structures of E5-CRD (20) and E5-CRD complexed with ImmE5 (21) at a resolution of 1.5 and 1.15 Å, respectively; they also proposed a putative model of the interaction between E5-CRD and tRNA. However, the precise structure of the active site in the pocket with catalytic residues remains unknown. Here, we report the crystal structures of E5-CRD complexed with its substrate analog and E5-CRD complexed with ImmE5. These structures clearly demonstrate the mechanism by which specific substrate recognition is achieved and reveal the residues that could play important roles for the catalytic activity in the absence of His. Based on these structures, we propose the double mimicry model to demonstrate the mechanism by which the protein–protein interaction replaces the RNA–RNA interaction.

MATERIALS AND METHODS

Plasmids

The pTO502 E5-CRD/ImmE5 expression plasmid was used for the crystallization; this plasmid was a derivative of pTO501, which has been described previously (10). The pTO501 plasmid contained a single mutation of I62M that was attributed to a PCR-cloning artifact in ImmE5; thus, this plasmid produces the wild-type ImmE5.

The pKF601 expression plasmid was used for the cell-killing assay. This plasmid contained an operon coding for colicin E5 and ImmE5 that was under the control of the colicin promoter. The pKF601 plasmid was constructed using the pTO502 plasmid, in which the *colE5-immE5* fragment from pColE5-099 was replaced with the *E5-CRD*/*immE5* fragment. The K25Q, R33Q, K60Q and K25Q/K60Q mutants of colicin E5 were constructed based on the pKF601 plasmid using the Quickchange mutagenesis kit (Stratagene, La Jolla, CA).

Protein expression and purification

The protein expressions from both plasmids pTO502 and pKF601 including its derivatives were induced by the addition of 0.2 mg/l mitomycin C, which activates the colicin promoters.

For crystallization, the E5-CRD/ImmE5 complex was purified from *E.coli* K12 RR1 [pTO502] using a Phenyl-Toyopearl 650 M column (Tosoh Co., Tokyo, Japan) with a decreasing gradient starting at 1 M ammonium sulfate in 20 mM potassium phosphate buffer (pH 7.0); this was followed by purification using a DEAE-Toyopearl 650S column (Tosoh Co.) with a KCl gradient elution in 20 mM Tris–HCl (pH 8.5). In order to obtain the E5-CRD protein, the E5-CRD/ImmE5 complex was dissociated using 6 M urea in 20 mM Tris–HCl (pH 8.5), and E5-CRD was separated on a MonoQ column and then renatured by dialysis in 20 mM Tris–HCl (pH 8.5).

For the cell-killing assay, the *E.coli* K12 RR1 strain was also used as a host for the expression of the wild-type (pKF601) or the mutant (pKF601 derivatives) colicin E5/ImmE5 complexes. The purification of the complex proteins was performed using a DEAE-Toyopearl 650S column in 20 mM Tris–HCl (pH 8.0) with a 0–0.5 M KCl gradient, followed by purification using a MonoS column in 20 mM potassium phosphate (pH 6.0) with a 0–0.5 M KCl gradient.

Crystallization and data collection

An *I*222 crystal of the E5-CRD/ImmE5 complex was obtained by the hanging-drop vapor diffusion method using 2 M ammonium sulfate and 0.1 M Tris–HCl (pH 8.0) at 20°C. A $P2_12_12_1$ crystal of the E5-CRD/dGpdUp complex was also obtained under the same conditions as those for the E5-CRD/ImmE5 complex. For the E5-CRD/dGpdUp complex, synthesized and HPLC-purified dGpdUp was mixed with the E5-CRD protein in a 2:1 molar ratio. The X-ray data of the E5-CRD/ImmE5 crystal were collected using a DIP2030 (MacScience Inc., Yokohama, Japan) with an image plate area detector system. The E5-CRD/dGpdUp crystals were flash-frozen in a 100 K nitrogen stream with

30% glycerol that was used as a cryoprotectant in the reservoir solution. The data were collected on beamline BL41XU at SPring-8. Both the data were processed using the DENZO and SCALEPACK programs (22).

Structure determinations and refinements

The E5-CRD/ImmE5 complex structure was solved by the multiple isomorphous replacement (MIR) method, and it was refined at a 1.9 Å resolution to an *R*-factor of 17.9% (R_{free} = 19.8%). MIR phasing and refinement of the heavy atom positions were carried out using the program MLPHARE, and density modification was performed using the program DM. Iterative rounds of refinement and rebuilding of the native model were performed by using the programs CNS (23) and QUANTA (Accelrys Inc., San Diego, CA), respectively. The structure of E5-CRD complexed with dGpdUp was determined by the molecular replacement method using the program AMoRe. The E5-CRD coordinates obtained by MIR were used as a target. The programs MLPHARE, DM and AMoRe are included in the CCP4 suite (24). The structure was refined at a 1.9 Å resolution to an *R*-factor of 19.7% (R_{free} = 21.7%). Figures 1, 3, 5, 6 and 7 were created using PyMOL (25), and Figure 2 was created using Molscript (26). The atomic coordinates and structure factors of the E5-CRD/ImmE5 and E5-CRD/dGpdUp complexes have been deposited in the PDB under the accession codes 2DFX and 2DJH, respectively.

Cell-killing assays

Cell-killing assays were performed by a spot test. Wild-type, K25Q, R33Q, K60Q and K25Q/K60Q colicin E5s were serially diluted from a starting concentration of 65 μg/ml by a factor of two using L-broth medium, and 3 μl of each dilution was applied onto a soft agar lawn of a sensitive indicator strain (*E.coli* K12 RR1). After overnight incubation at 37°C, the growth inhibition zones were compared.

RESULTS

Crystallizations and structure determinations

The crystals of the E5-CRD/ImmE5 complex grew to a typical size of 0.2 × 0.2 × 0.4 mm within a week. The data were collected in the absence of a cryocooling condition; this resulted in a very low mosaicity of 0.15 as calculated by the software SCALEPACK (22). In order to solve the phase problem, the platinum and mercury derivative crystals were used for the MIR methods.

The deoxy derivative dGpdUp was used as a substrate analog for the crystallization of E5-CRD because E5-CRD very efficiently cleaves between the G and U of the target tRNAs, and the diribonucleotide GpUp was also specifically cleaved by E5-CRD *in vitro* (12). E5-CRD and dGpdUp were co-crystallized to obtain the complex; the crystals appeared in a few days and grew to a maximum size of 0.3 × 0.3 × 0.7 mm within a week. The diffraction data were collected at 100 K under a liquid nitrogen stream. The molecular replacement method was applied to solve the structure by using the coordinates of E5-CRD in the ImmE5 complex as the search model. The data were collected at 1.8 Å and refined at 1.9 Å because the refinement at 1.8 Å did not improve the statistics. The Ramachandran plot analysis of the refined structure showed that 92% of the residues were in the most favored regions, and the remaining residues were in additional allowed regions.

The data collection and the refinement statistics for both crystals are summarized in Table 1.

Table 1. Data collection, phase determination and refinement statistics of E5-CRD/dGpdUp and E5-CRD/ImmE5

	E5-CRD/dGpdUp Native	E5-CRD/ImmE5 Native	K_2PtCl_4	Mersalyl acid
Data collection				
Space group	$P2_12_12_1$	$I222$		
Cell dimensions				
a, *b*, *c* (Å)	*a* = 29.7, *b* = 49.4, *c* = 63.9	*a* = 63.5, *b* = 74.4, *c* = 111.2		
Resolution (Å)	1.8 (1.86–1.80)	1.86 (1.93–1.86)	2.4 (2.49–2.40)	1.86 (1.93–1.86)
R_{merge}[a]	0.046 (0.131)	0.040 (0.197)	0.059 (0.145)	0.076 (0.405)
$I/\sigma I$	20.8 (8.8)	11.1 (5.7)	12.3 (7.0)	7.4 (2.9)
Completeness (%)	99.2 (100.0)	92.7 (73.0)	93.5 (62.8)	97.9 (82.0)
Total reflections	116 803	124 539	77 483	159 223
Unique reflections	8694	20 556	9087	21 662
Number of sites			3	1
R_{cullis}[b]			0.51	0.68
Phasing power[c]			1.4	0.9
Refinement				
Resolution (Å)	26.5–1.9	19.1–1.9		
Number of reflections	7399	19 713		
R_{work}/R_{free}[d]	0.197/0.217	0.179/0.198		
R.m.s. deviations				
Bond lengths (Å)	0.005	0.005		
Bond angles (°)	1.3	1.3		

[a] $R_{merge} = \Sigma_h\Sigma_i |I_i(h) - < I(h) > |/\Sigma_h\Sigma_i I_i(h)$, where $I_i(h)$ is the *i*th measurement.
[b] $R_{cullis} = \Sigma[|F_H| - (|F_{PH}| - |F_P|)]/\Sigma(|F_{PH}| - |F_P|)$, where $|F_H|$ represents the calculated heavy atom structure factor.
[c] Phasing power = f_{rms}/E_{rms}, where $f_{rms} = [(\Sigma f^2{}_H)/n]^{1/2}$ and $E_{rms} = [\Sigma(F_{PH} - |F_P + f_H|)^2/n]^{1/2}$.
[d] $R = \Sigma|F_O - F_C|/\Sigma F_O$ for all reflections and $R_{free} = \Sigma|F_O - F_C|/\Sigma F_O$, calculated with the 5% of the data excluded from refinement.

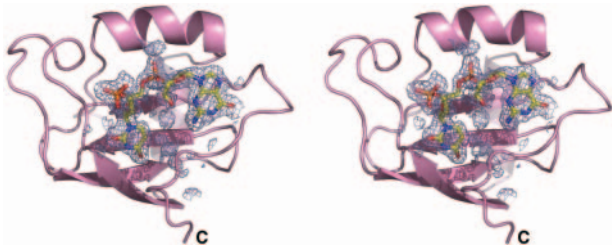

Figure 1. Stereoview of the E5-CRD/dGpdUp complex structure. E5-CRD is represented as a ribbon model and dGpdUp, as a ball-and-stick model. The *Fo*–*Fc* omit map of dGpdUp ($\sigma = 1.0$) is also presented.

Structure of the E5-CRD/dGpdUp complex

In the crystal structure of E5-CRD complexed with dGpdUp, the dinucleotide is tightly trapped in the pocket of E5-CRD. As shown in Figure 1, the pocket is formed by a long partially coiled loop on the left, the parallel α-helix at the back, and the C-terminal loop on the right. The nucleotide conformations are *syn* and *anti* for dG and dU, respectively. Although the *anti* conformation is known to be predominant for dG in solution, the *syn* conformation for dG in our structure is stabilized by interactions with E5-CRD, as described below. The sugar puckers of the deoxyribose rings are C4′-*exo* and C2′-*endo* for dG and dU, respectively.

The superposition of three E5-CRD structures—one complexed with dGpdUp, another with ImmE5, and the apo form (PDB code 2A8K)—showed that they fit well in the central region of the β-sheet with the α-helix being parallel to the β-sheet (parallel α-helix); however, they exhibit slight differences at both ends of the β-sheet and the C-terminal region. The structure with the analog had the most closed conformation. This observation suggests the flexibility of E5-CRD to bind both tRNAs and ImmE5 (Figure 2).

The most interesting feature is the specific interaction between the dinucleotide and the protein surface. The guanine base is stacked with the indole ring of Trp102; these two double rings face each other, and the distance between them is almost the same as that of B-DNA. In addition, the guanine base forms three hydrogen bonds—O6 of dG with amide N of Val103, N1 of dG with carbonyl O of Val103, and N2 of dG with carbonyl O of Val103. With regard to the dU moiety of the substrate analog, the uracil base forms three hydrogen bonds—O4 of dU with hydroxyl O of Ser52, N3 of dU with carbonyl O of Phe53 and O2 of dU with amide N of Lys55. The uracil base ring also undergoes a stacking interaction with a pseudo ring formed by the carboxyl group of Asp105 and the guanidino group of Arg107 (Figure 3A). The distance between the uracil base and the pseudo ring is the same as that between the guanine base and the indole ring of Trp102 mentioned above.

While the base rings are restrained by hydrogen bonds and ring interactions, backbone phosphates are also retained by several hydrogen bonds. The phosphate between dG and dU forms hydrogen bonds with NH2 of Arg33, NE2 of Gln29 and NZ of Lys25, whereas the 3′-terminal phosphates form hydrogen bonds with hydroxyl O and amide N of Ser57 and NZ of Lys60 (Figure 3B).

The inspection of the structure surrounding the substrate analog revealed a positively charged cluster that was composed of two Lys and one Arg residues (Figure 3B). These residues are positioned in proximity to the phosphodiester bond between dG and dU, and no other charged side chains were present close to this bond. As shown in Figure 3B, Arg33 extends its side chain to the phosphate group from the guanosine side, while Lys25 and Lys60 do so from the uridine side. The distances between NH1 of Arg33 and O1P and that between NH1 of Arg33 and the expected position of 2′-O of guanine are 2.8 and 3.5 Å, respectively. In addition, the distances between NZ of Lys25 and 5′-O and that between NZ of Lys60 and 5′-O of uridine are 4 and 3.5 Å, respectively.

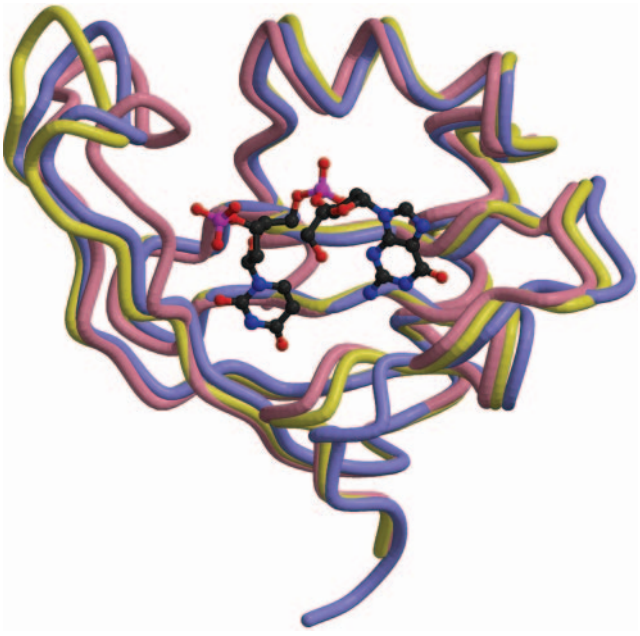

Figure 2. The superposition of three E5-CRD structures. The E5-CRD structure complexed with dGpdUp, another complex with ImmE5, and the apo form (PDB code 2A8K) are shown in pink, yellow and purple, respectively. The dGpdUp structure is also shown as a ball-and-stick model.

Cell-killing assay

Based on the structure of the E5-CRD/dGpdUp complex, three basic residues were implicated to be important for the tRNase activity. Hence, we introduced four types of mutations—R33Q, K25Q, K60Q and K25Q/K60Q—in the C-terminal domain of colicin E5 and analyzed their killing activities (Figure 4). Serial dilution of purified colicin E5s clearly showed that mutations R33Q and K25Q/K60Q abolished the killing activity of the protein against sensitive cells; on the other hand, K25Q or K60Q reduced the killing activity.

DISCUSSION

Sequence-specific recognition mechanism of E5-CRD

Among the colicin family proteins, the crystal structure of RNase-type colicin E3 has been reported (27,28). Colicins E3 and E5 are homologous at the receptor-binding and membrane-translocating domains. On the other hand, the

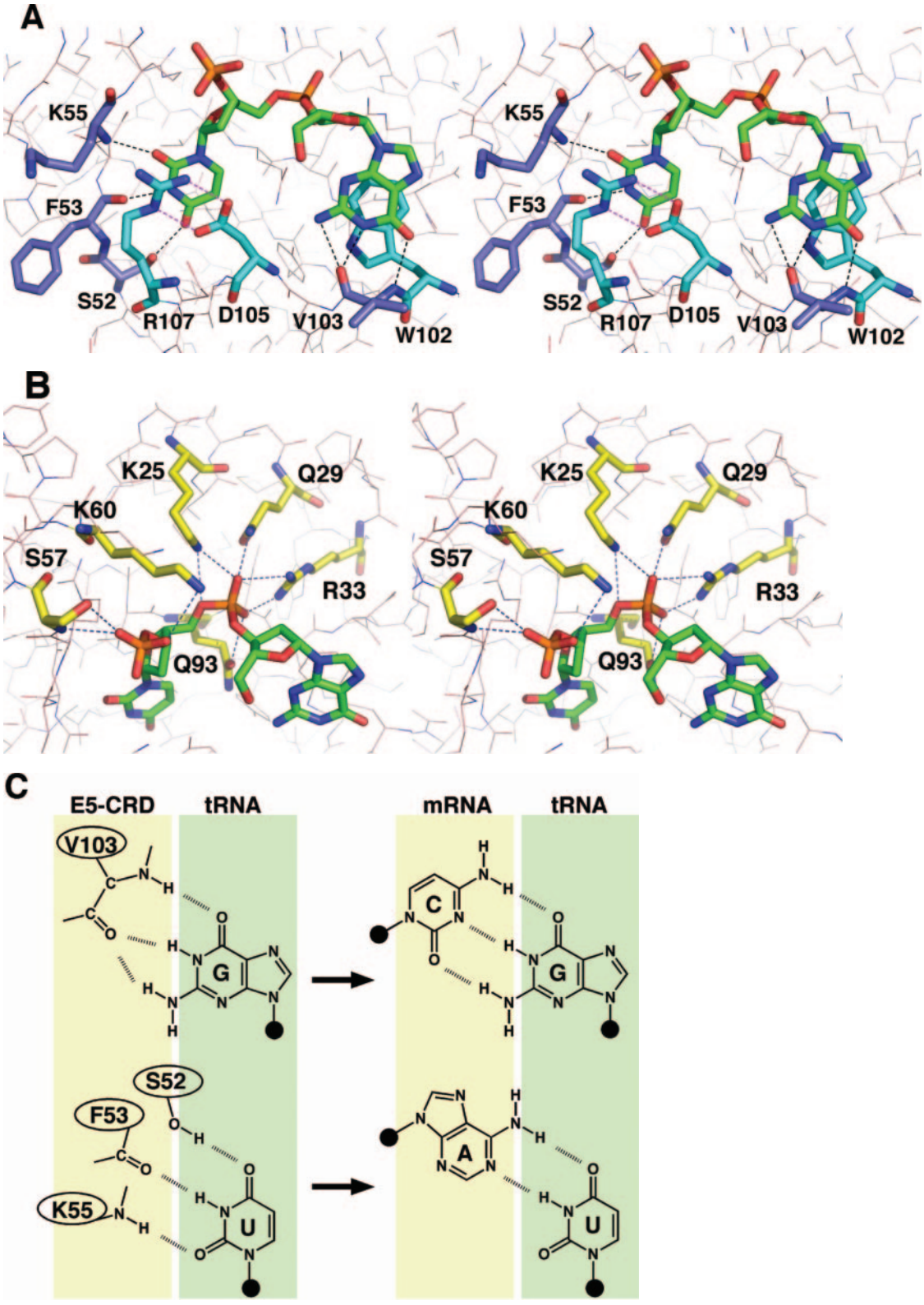

Figure 3. The binding mode of dGpdUp to E5-CRD. (**A**) The Watson–Crick-type interaction between dGpdUp and E5-CRD. The cyan- and purple-colored residues contribute to the interactions with ring–ring interactions and the hydrogen bonds of nucleotide bases, respectively. (**B**) The yellow-colored residues contribute to the interactions with the phosphates of dGpdUp. (**C**) A schematic representation of the Watson–Crick-type hydrogen bonds between E5-CRD and dGpdUp, which mimic the mRNA–tRNA interaction. The yellow- and green-colored column pairs at the left and at the right show E5-CRD/tRNA(dGpdUp) and mRNA/tRNA interactions, respectively.

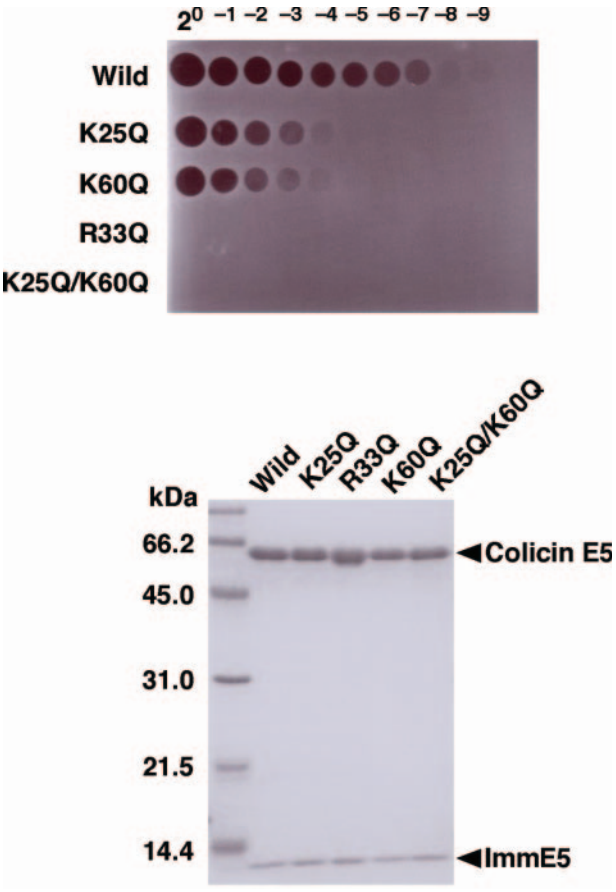

Figure 4. Cell-killing assay. Purified wild-type colicin E5, K25Q, R33Q, K60Q or K25Q/K60Q was applied on a soft agar lawn of sensitive *E.coli* cells. The proteins were serially diluted by a factor of two starting at a concentration of 65 μg/ml.

catalytic domains of both these colicins showed no homology, reflecting completely different targets. Colicin E5 is a tRNase that specifically targets the GU and QU sequences of the anticodons specifically. In order to understand this specific recognition mechanism, we solved the crystal structure of the C-terminal tRNase domain of colicin E5 complexed with its substrate analog, dGpdUp. The tight binding of dGpdUp to E5-CRD was achieved by ring–ring-interactions and the maximum possible hydrogen bonds, suggesting the mimicry of Watson–Crick-type interactions in both dG and dU (Figure 3A and C). If the dG base were to be substituted by dA, none of these hydrogen bonds would be possible. In contrast, if the same dG base was to be substituted by a Q base, all these hydrogen bonds would be retained, while the modification extending from the N7 position would easily extend into the solvent. The above modeling consideration is consistent with the experimental observation that the tRNAs with Q as well as the corresponding molecules without Q are cleaved by E5-CRD.

If the dU base were to be substituted by dC, only one hydrogen bond would be retained. Thus, the nucleotide sequence G (or Q)-U appears to be favored by E5-CRD through the formation of maximum possible hydrogen bonds and aromatic ring (or pseudo ring) stacking interactions. All the hydrogen bonds use the backbone

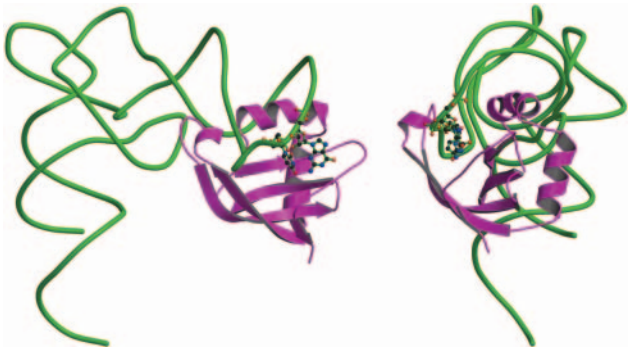

Figure 5. Docking model of E5-CRD and tRNA. The $tRNA^{Asp}$ from 1asy.pdb was manually docked to E5-CRD from the dGpdUp complex structure. The tRNA was placed at the position where G–U in its anticodon, which is shown as a ball-and-stick model, superposed onto dGpdUp of the E5-CRD complex. The right figure is a side view of the left figure.

amide N or carbonyl O of E5-CRD, except hydroxyl O of Ser52.

In addition to the interactions of the bases by the hydrogen bonds and ring interactions, backbone phosphates are also retained by several hydrogen bonds. Thus, this observation also suggests that E5-CRD has less flexibility for substrate recognition; this results in binding to only the specific target.

These findings explain that the specific cleavage of the Q(G)UN anticodon sequence by E5-CRD is based on the GU-specific base recognition by E5-CRD. In addition, this specific recognition between the protein and the RNA is realized in a manner similar to the Watson–Crick-type interaction used in an RNA double strand.

There are several examples of RNAs being processed in a site-specific manner. MazF and ChpBK cleave free RNAs at the ACA and ACY (Y: U, A or G) sequences, respectively (29,30). Both proteins are *E.coli* chromosomal toxins and inhibit translation by cleaving mRNA. The crystal structure of MazF has been recently reported (31); however, no structural information on its substrate binding is available.

Although there should be a large number of GU sequences in cellular RNAs, E5 preferentially acts on tRNA anticodons *in vivo* (10). When we docked E5-CRD and a tRNA molecule ($tRNA^{Asp}$, PDB code 1ASY) manually by superimposing dGpdUp and GU of the tRNA, E5-CRD made contact only with the anticodon arm of tRNA (Figure 5). Interestingly, the parallel helix of E5-CRD fits in the helix of the anticodon arm. This may explain the substrate specificity of tRNA, where E5-CRD tightly binds to tRNA thereby allowing the capture of G–U bases at the bottom of the binding pocket. Therefore, E5-CRD may prefer GU sequences embedded in specific secondary structures such as anticodon loops, and in fact, the enzyme kinetic study also demonstrated this GU preference by E5-CRD (12).

Which amino acids serve as the catalytic residues?

If E5-CRD is a true ribonuclease that lacks His, which residues serve as the catalysts? Based on the crystal structure of the E5-CRD/dGpdUp complex, only Lys25, Lys60 and Arg33 were observed as charged residues with the proximity of the phosphates in the substrate (Figure 3B). With reference

to the well-studied case of pancreatic RNase A in which His12 is a general base located close to 2′-OH and His119 is a general acid located close to 5′-O in the 2′,3′-cyclizing step (32), Arg33 may position itself corresponding to that of His12, and Lys25 or Lys60 may position itself corresponding to that of His119. The mutation analysis revealed that both R33Q and K25Q/K60Q mutants lost their killing activity, and either of the K25Q or the K60Q mutant exhibited a decreased activity. Thus, this result suggests that these three basic residues are important for the catalytic activity. Interestingly, E5-CRD preferred a very high pH to cleave a dinucleotide substrate (12). The normally accepted pK_a values of the Arg and Lys side chains might explain this unusual pH preference. In addition, the carbonyl group of Ile94 is close to the guanidino group of Arg33, thereby forming a hydrogen bond; this possibly increases the basicity of the protonated guanidino group of Arg33. However, we could not thoroughly measure the activity at higher pH values due to RNA instability. On the other hand, in *E.coli* cells, in which it is difficult to achieve high pH values under physiological conditions, both Arg and Lys are assumed to be protonated. Thus, it is unlikely for these basic residues to function as general acid–base catalysts.

Lin *et al*. have reported the crystal structure of the apo form of E5-CRD; they showed that Asp54, Arg56, Lys59, Asp105 and Arg107 (corresponding to their residue numbers Asp46, Arg48, Lys51, Asp97 and Asp99, respectively) affected its ribonuclease activity due to mutations (20). Further, among these residues, Asp54 and Arg56 were suggested to be putative catalytic residues. Based on the structure of our complex, however, the Cα atoms of Asp54 and Arg56 were located approximately 14 and 11 Å, respectively, apart from the 2′-C atom of the deoxyguanosine of dGpdUp; in addition, the residues were blocked by uridine thereby preventing their contact with the 2′-C atom and the phosphate group between dG and dU. Furthermore, instead of playing a role in substrate binding, Asp54 and Arg56 formed a salt bridge, enabling the complex structure to maintain its conformation. In addition, Asp105–Arg107 also formed a salt bridge, thereby contributing to the interaction with the uracil base ring of the substrate. Thus, the loss of the ribonuclease activity due to mutations of these two residues demonstrates their importance with regard to the uridine binding revealed by our structure. Lys59 may also contribute to the substrate binding with the phosphate group.

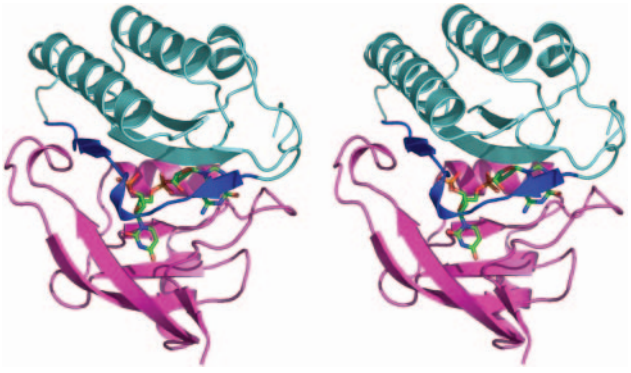

Figure 6. The structure of E5-CRD complexed with ImmE5. E5-CRD and ImmE5 are represented in magenta and cyan, respectively. The loop region (Phe60–Pro73) of ImmE5 is represented in blue. In the E5-CRD/dGpdUp complex structure, only dGpdUp is represented as a ball-and-stick model. The position of dGpdUp is derived by the superposition of E5-CRDs complexed with ImmE5 or dGpdUp.

Therefore, the residues found in this study and by Lin *et al*. are crucial for the catalytic activity; however, none of these appear to be apparent catalytic residues, suggesting the existence of an alternative novel mechanism responsible for the E5 activity.

In deoxyribonucleases (DNases), a water molecule that is activated by divalent metal ions is involved in the catalytic reaction. However, no metal ions were found to be associated with the E5-CRD/dGpdUp complex; in fact, the enzyme reaction was resistant to EDTA (data not shown). Although a water molecule was found to bind to O1P of the phosphate of dGpdUp in the active site, it may function to fix the substrate at the active site.

E5-CRD/ImmE5 mimics the molecular recognition of RNA–RNA interaction

Recently, Luna-Chávez *et al*. have solved the crystal structure of E5-CRD complexed with its cognate inhibitor protein ImmE5 through a multiwavelength anomalous diffraction experiment (21). We have also been studying the same structure by using MIR. As a result, when we superposed our E5-CRD/ImmE5 structure onto the structure obtained by Luna-Chávez *et al*., the r.m.s. difference for the backbone atoms (Cα, C, O and N) was 0.18 Å. The resolutions of our structure and their structures are 1.9 and 1.15 Å, respectively. As shown in the case of barnase-barstar (33), this may indicate that the E5-CRD/ImmE5 complex structure is very rigid, thereby suggesting the tight binding to the target tRNAs. However, the free E5-CRD structure is rather flexible as shown in Figure 2. The interaction of E5-CRD and ImmE5 was realized mainly by the electrostatic interaction between the positive charges on E5-CRD and the negative charges on ImmE5. In addition to the observation on this binding mode by Luna-Chávez *et al*., we found that 12 water molecules interacted with the main chains or side chains of ImmE5 and E5-CRD through the hydrogen bonds resulting in the filling of the gap between the two proteins; this may contribute to their tight binding. As determined in a preliminary modeling study on E5-CRD and the tRNA molecule, an anticodon loop fits in the pocket within which are located a long loop (Phe60–Pro73) comprising two short strands and a helix of ImmE5 (Figure 6). Based on this conformation of the interaction of E5-CRD/ImmE5, we considered that the long loop region in ImmE5 may mimic the anticodon loop of a tRNA. Additionally, the basic residues Lys25, Arg33 and Lys60, which are essential for the catalytic activity of E5-CRD are in contact with the acidic residues Asp51 and Asp94 in ImmE5. When we superposed the E5-CRD/ImmE5 complex structure onto the E5-CRD/dGpdUp complex structure, the positions of Asp51 and Asp94 corresponded with those of the two phosphates of dGpdUp (Figure 7). This suggests that the acidic residues in ImmE5 mimic the phosphates in the anticodon; thus, ImmE5 inhibits the E5-CRD activity by mimicking the binding mode of the tRNA anticodon loop.

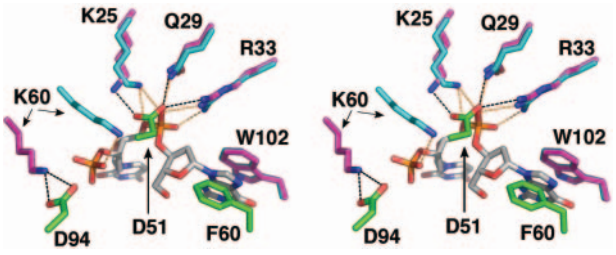

Figure 7. The mimicry of the molecular recognition of mRNA–tRNA by E5-CRD/ImmE5. The residues of ImmE5 are represented in green. The residues of E5-CRD complexed with ImmE5 and with dGpdUp are represented in magenta and cyan, respectively.

As described above, dGpdUp binds to E5-CRD using ring–ring interactions and hydrogen bonds at base rings in a manner similar to the Watson–Crick-type interaction (Figure 3A and C). Thus, it is very likely that the E5-CRD/dGpdUp binding mode mimics an mRNA–tRNA interaction. If E5-CRD mimics mRNA while binding to tRNA and ImmE5 mimics tRNA while binding to E5-CRD, we speculate that the molecular recognition mode of the protein–protein interaction in E5-CRD/ImmE5 may be the culmination of the molecular evolution originating from the mRNA–tRNA interaction.

In the evolutionary process, initially, E5-CRD could replace the mRNA interacting with the tRNA by means of Trp102 and Asp105–Arg107 that mimicked the nucleotide base stacking with the G–U nucleotides of tRNA. Trp102 and Asp105–Arg107 in E5-CRD mimicked the nucleotide bases in the mRNA, suggesting the interaction of Trp102 and Asp105–Arg107 with G–U. Subsequently, additional interactions between the phosphate groups of the substrate and basic residues in E5-CRD were established to stabilize the binding (the interactions between blue-colored residues and dGpdUp in Figure 7). Then, ImmE5 could replace the tRNA to interact with E5-CRD; the guanine base was replaced by Phe60 of ImmE5, and two phosphates were replaced by Asp51 and Asp94. Due to this replacement, the hydrogen bonds and the ring interactions between dGpdUp and the residues in E5-CRD are almost precisely retained by the residues in ImmE5 and E5-CRD at the corresponding positions.

CONCLUSION

The E5-CRD/dGpdUp complex structure reported here clearly demonstrated the mechanism by which sequence-specific recognition is achieved by colicin E5. The crystal structures of RNases such as RNase A (34), T1 (35) and barnase (36) have been solved using their corresponding substrates. When compared with these, the recognition mechanism of E5-CRD was unique and corresponded to the substrate specificity. We also found three basic residues that are important for the E5-CRD activity. However, since the pK_a values of their side chains are very high and no other charged residues are present around the cleavage site, the precise catalytic mechanism without histidines remains to be elucidated.

In addition to the analysis of the E5-CRD functions, a comparison of the crystal structures of the E5-CRD/ImmE5 and E5-CRD/dGpdUp complexes suggested that both E5-CRD and ImmE5 mimic the molecular recognition mechanism of RNA, in which a pair comprising a ribonuclease and the cognate inhibitor protein replaced the RNA–RNA interaction to realize the protein–protein interaction through an RNA–protein interaction. In the evolutionary process of molecules, if RNAs were to be the sole precursors, the roles of RNA molecules could have been fulfilled by proteins following their emergence. Thus, the E5-CRD/ImmE5 complex structure might represent an example of this evolutionary process.

ACKNOWLEDGEMENTS

We are grateful to the SPring-8 facility for providing us access and user support. This work was supported by a grant from the Nissan Science Foundation and the Institute of Applied Bioscience, Tokyo University of Agriculture to S.Y. and also partly by a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology to H.M. Funding to pay the Open Access publication charges for this article was provided by Tokyo University of Agriculture.

Conflict of interest statement. None declared.

REFERENCES

1. James,R., Kleanthous,C. and Moore,G.R. (1996) The biology of E colicins: paradigms and paradoxes. *Microbiology*, **142**, 1569–1580.
2. Braun,V., Patzer,S.I. and Hantke,K. (2002) Ton-dependent colicins and microcins: modular design and evolution. *Biochimie*, **84**, 365–380.
3. Schaller,K. and Nomura,M. (1976) Colicin E2 is DNA endonuclease. *Proc. Natl Acad. Sci. USA*, **73**, 3989–3993.
4. Toba,M., Masaki,H. and Ohta,T. (1988) Colicin E8, a DNase which indicates an evolutionary relationship between colicins E2 and E3. *J. Bacteriol*., **170**, 3237–3242.
5. Pommer,A.J., Wallis,R., Moore,G.R., James,R. and Kleanthous,C. (1998) Enzymological characterization of the nuclease domain from the bacterial toxin colicin E9 from *Escherichia coli*. *Biochem. J*., **334 (Pt 2)**, 387–392.
6. Bowman,C.M., Dahlberg,J.E., Ikemura,T., Konisky,J. and Nomura,M. (1971) Specific inactivation of 16S ribosomal RNA induced by colicin E3 *in vivo*. *Proc. Natl Acad. Sci. USA*, **68**, 964–968.
7. Bowman,C.M., Sidikaro,J. and Nomura,M. (1971) Specific inactivation of ribosomes by colicin E3 *in vitro* and mechanism of immunity in colicinogenic cells. *Nat. New Biol*., **234**, 133–137.
8. Senior,B.W. and Holland,I.B. (1971) Effect of colicin E3 upon the 30S ribosomal subunit of *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **68**, 959–963.
9. Yokoyama,S., Miyazawa,T., Iitaka,Y., Yamaizumi,Z., Kasai,H. and Nishimura,S. (1979) Three-dimensional structure of hyper-modified nucleoside Q located in the wobbling position of tRNA. *Nature*, **282**, 107–109.
10. Ogawa,T., Tomita,K., Ueda,T., Watanabe,K., Uozumi,T. and Masaki,H. (1999) A cytotoxic ribonuclease targeting specific transfer RNA anticodons. *Science*, **283**, 2097–2100.
11. Lau,P.C. and Condie,J.A. (1989) Nucleotide sequences from the colicin E5, E6 and E9 operons: presence of a degenerate transposon-like structure in the ColE9-J plasmid. *Mol. Gen. Genet*., **217**, 269–277.
12. Ogawa,T., Inoue,S., Yajima,S., Hidaka,M. and Masaki,H. (2006) Sequence-specific recognition of colicin E5, a tRNA-targeting ribonuclease. *Nucleic Acids Res*., **34**, 6065–6073.
13. Tomita,K., Ogawa,T., Uozumi,T., Watanabe,K. and Masaki,H. (2000) A cytotoxic ribonuclease which specifically cleaves four isoaccepting

arginine tRNAs at their anticodon loops. *Proc. Natl Acad. Sci. USA*, **97**, 8278–8283.
14. Kaufmann,G. (2000) Anticodon nucleases. *Trends Biochem. Sci.*, **25**, 70–74.
15. Lu,J., Huang,B., Esberg,A., Johansson,M.J. and Bystrom,A.S. (2005) The *Kluyveromyces lactis* gamma-toxin targets tRNA anticodons. *RNA*, **11**, 1648–1654.
16. Graille,M., Mora,L., Buckingham,R.H., van Tilbeurgh,H. and de Zamaroczy,M. (2004) Structural inhibition of the colicin D tRNase by the tRNA-mimicking immunity protein. *EMBO J.*, **23**, 1474–1482.
17. Yajima,S., Nakanishi,K., Takahashi,K., Ogawa,T., Hidaka,M., Kezuka,Y., Nonaka,T., Ohsawa,K. and Masaki,H. (2004) Relation between tRNase activity and the structure of colicin D according to X-ray crystallography. *Biochem. Biophys. Res. Commun.*, **322**, 966–973.
18. Zhang,Y., Zhang,J., Hara,H., Kato,I. and Inouye,M. (2005) Insights into the mRNA cleavage mechanism by MazF, an mRNA interferase. *J. Biol. Chem.*, **280**, 3143–3150.
19. Pedersen,K., Zavialov,A.V., Pavlov,M.Y., Elf,J., Gerdes,K. and Ehrenberg,M. (2003) The bacterial toxin RelE displays codon-specific cleavage of mRNAs in the ribosomal A site. *Cell*, **112**, 131–140.
20. Lin,Y.L., Elias,Y. and Huang,R.H. (2005) Structural and mutational studies of the catalytic domain of colicin E5: a tRNA-specific ribonuclease. *Biochemistry*, **44**, 10494–10500.
21. Luna-Chávez,C., Lin,Y.L. and Huang,R.H. (2006) Molecular basis of inhibition of the ribonuclease activity in colicin e5 by its cognate immunity protein. *J. Mol. Biol.*, **358**, 571–579.
22. Otwinowski,Z. and Minor,W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.*, **276**, 307–326.
23. Brunger,A.T., Adams,P.D., Clore,G.M., DeLano,W.L., Gros,P., Grosse-Kunstleve,R.W., Jiang,J.S., Kuszewski,J., Nilges,M., Pannu,N.S. *et al.* (1998) Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr.*, **D54**, 905–921.
24. Collaborative Computational Project, Number 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr.*, **D50**, 760–763.
25. DeLano,W.L. (2002) The PyMOL Molecular Graphic System 2002.
26. Kraulis,P.J. (1991) MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.*, **24**, 946–950.
27. Carr,S., Walker,D., James,R., Kleanthous,C. and Hemmings,A.M. (2000) Inhibition of a ribosome-inactivating ribonuclease: the crystal structure of the cytotoxic domain of colicin E3 in complex with its immunity protein. *Structure*, **8**, 949–960.
28. Soelaiman,S., Jakes,K., Wu,N., Li,C. and Shoham,M. (2001) Crystal structure of colicin E3: implications for cell entry and ribosome inactivation. *Mol. Cell.*, **8**, 1053–1062.
29. Zhang,Y., Zhang,J., Hoeflich,K.P., Ikura,M., Qing,G. and Inouye,M. (2003) MazF cleaves cellular mRNAs specifically at ACA to block protein synthesis in *Escherichia coli*. *Mol. Cell.*, **12**, 913–923.
30. Zhang,Y., Zhu,L., Zhang,J. and Inouye,M. (2005) Characterization of ChpBK, an mRNA interferase from *Escherichia coli*. *J. Biol. Chem.*, **280**, 26080–26088.
31. Kamada,K., Hanaoka,F. and Burley,S.K. (2003) Crystal structure of the MazE/MazF complex: molecular bases of antidote-toxin recognition. *Mol. Cell.*, **11**, 875–884.
32. Raines,R.T. (1998) Ribonuclease A. *Chem. Rev.*, **98**, 1045–1066.
33. Buckle,A.M., Schreiber,G. and Fersht,A.R. (1994) Protein–protein recognition: crystal structural analysis of a barnase-barstar complex at 2.0-Å resolution. *Biochemistry*, **33**, 8878–8879.
34. Fontecilla-Camps,J.C., de Llorens,R., le Du,M.H. and Cuchillo,C.M. (1994) Crystal structure of ribonuclease A.d(ApTpApApG) complex. Direct evidence for extended substrate recognition. *J. Biol. Chem.*, **269**, 21526–21531.
35. Heinemann,U. and Saenger,W. (1982) Specific protein–nucleic acid recognition in ribonuclease T1-2′-guanylic acid complex: an X-ray study. *Nature*, **299**, 27–31.
36. Baudet,S. and Janin,J. (1991) Crystal structure of a barnase-d(GpC) complex at 1.9 A resolution. *J. Mol. Biol.*, **219**, 123–132.