The crystal structure of the DNase domain of colicin E7 in complex with its inhibitor Im7 protein

Tzu-Ping Ko[1], Chen-Chung Liao[2], Wen-Yen Ku[1,3], Kin-Fu Chak[2]* and Hanna S Yuan[1]*

Background: Colicin E7 (ColE7) is one of the bacterial toxins classified as a DNase-type E-group colicin. The cytotoxic activity of a colicin in a colicin-producing cell can be counteracted by binding of the colicin to a highly specific immunity protein. This biological event is a good model system for the investigation of protein recognition.

Results: The crystal structure of a one-to-one complex between the DNase domain of colicin E7 and its cognate immunity protein Im7 has been determined at 2.3 Å resolution. Im7 in the complex is a varied four-helix bundle that is identical to the structure previously determined for uncomplexed Im7. The structure of the DNase domain of ColE7 displays a novel α/β fold and contains a Zn^{2+} ion bound to three histidine residues and one water molecule in a distorted tetrahedron geometry. Im7 has a V-shaped structure, extending two arms to clamp the DNase domain of ColE7. One arm (α1*–loop 12–α2*; where * represents helices in Im7) is located in the region that displays the greatest sequence variation among members of the immunity proteins in the same subfamily. This arm mainly uses acidic sidechains to interact with the basic sidechains in the DNase domain of ColE7. The other arm (loop 23–α3*–loop 34) is more conserved and it interacts not only with the sidechain but also with the mainchain atoms of the DNase domain of ColE7.

Conclusions: The protein interfaces between the DNase domain of ColE7 and Im7 are charge-complementary and charge interactions contribute significantly to the tight and specific binding between the two proteins. The more variable arm in Im7 dominates the binding specificity of the immunity protein to its cognate colicin. Biological and structural data suggest that the DNase active site for ColE7 is probably near the metal-binding site.

Addresses: [1]Institute of Molecular Biology, Academia Sinica, Taipei, Taiwan 11529, Republic of China, [2]Institute of Biochemistry, National Yang-Ming University, Taipei, Taiwan, Republic of China and [3]National Defense University, Institute of Life Science, Taipei, Taiwan, Republic of China.

*Corresponding authors.
E-mail: mbyuan@ccvax.sinica.edu.tw
chak4813@ms2.hinet.net

Key words: DNase, E-group colicins, protein–protein interaction, protein recognition

Received: **14 September 1998**
Revisions requested: **29 October 1998**
Revisions received: **17 November 1998**
Accepted: **20 November 1998**

Published: **4 January 1999**

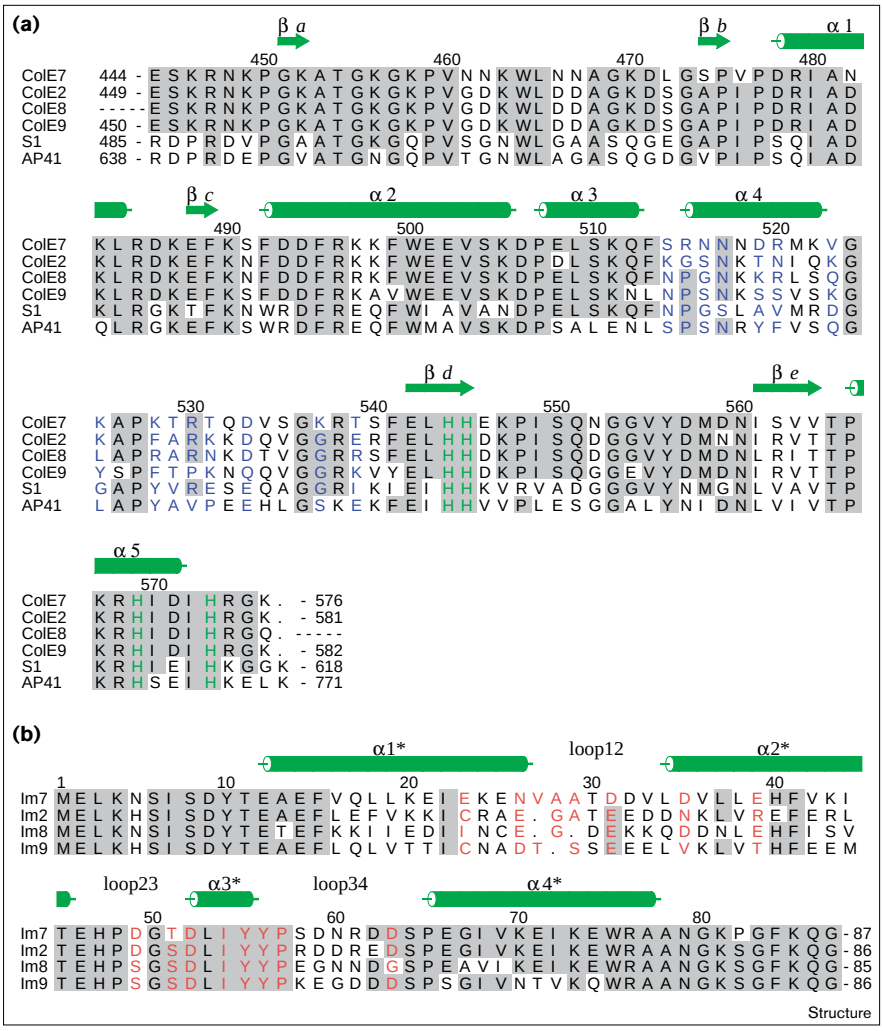
Structure January 1999, **7**:91–102
http://biomednet.com/elecref/0969212600700091

Introduction

Protein–protein interactions are involved in many fundamental processes that occur inside cells, such as cellular signal transduction, gene regulation, muscle contraction, and immunoprotection. Protein recognition is thus one of the most intensely studied areas in structural biology. The specific interactions between colicins and their immunity proteins provide a good system for investigating the principles of protein recognition. Colicins are a group of plasmid-encoded protein toxins produced by *Escherichia coli* to kill other sensitive *E. coli* and closely related coliform bacteria [1]. Immunity proteins coexpressed with colicins are the natural inhibitors of colicins. In colicin-producing cells, the cytotoxic activity of a colicin is inactivated by forming a one-to-one noncovalent complex with its cognate immunity protein.

Colicin plasmids are found in 30% of naturally occurring populations of *E. coli* and more than twenty colicins have been identified [2]. Colicins comprise three functional domains: receptor binding, membrane translocation and cytotoxic domains [3] and are classified into different subfamilies according to their binding to cell-membrane receptors. For example, colicins containing a receptor-binding domain that binds only to the vitamin B12 receptor (BtuB) are classified as E-group colicins. The cytotoxic activities of colicins against their target cells are well documented: as an RNase [4,5], a DNase [6], or an ionophore [7]. The pore-forming colicins have been studied more extensively and were used as model systems for investigating the mechanisms of protein translocation and channel formation. The crystal structures of two pore-forming colicins, Ia [8] and N [9], were reported recently. In contrast, the three-dimensional structures of the enzymatic colicins are less known. Crystal or nuclear magnetic resonance (NMR) spectroscopy solution structures of three immunity proteins, Im3 [10], Im7 [11–13], and Im9 [14], have been reported, but the structure of an enzymatic colicin has not been resolved yet. As a result, the molecular basis for the enzymatic activity of colicins has still not been clearly elucidated.

Figure 1



Sequence alignment of several colicin and pyocin DNase domains, and of several immunity proteins. **(a)** The DNase domain of ColE7 aligns with other DNase-type colicins and related toxins. S1 and AP41 refers to two pyocins identified from *Pseudomonas aeruginosa* [44,45]. **(b)** Im7 aligns with other immunity proteins. The conserved residues are shaded in gray, and the residues involved in protein–protein interactions are in blue in (a) and in red in (b). The histidine residues directly bound to the Zn^{2+} ion in (a) are in green. The secondary structure elements indicated are those defined by the present work using the program PROCHECK [46]; α helices are represented by green cylinders, and β sheets by green arrows. The figure was prepared using the program ALSCRIPT [47].

Colicin E7 (ColE7) is a nonspecific DNase of the E-group colicins [15]. Three other E-group colicins, ColE2 [16], ColE8 [17] and ColE9 [18], have been identified as DNase-type colicins and their specific inhibitors are designated as Im2, Im8, and Im9, respectively. The sequences of the C-terminal DNase domains of these colicins are approximately 80% identical (see Figure 1) and the sequences of their cognate immunity proteins are approximately 50% identical. However, these immunity proteins only bind with high affinity to the DNase domain of their cognate colicins, with dissociation constants ranging from 10^{-14} to 10^{-17} M [19]. The binding affinities of colicins with their cognate immunity proteins are amongst the strongest recorded for protein–protein interactions [20]. We have previously determined the crystal structure of Im7 [11,12]. In this paper, we report the crystal structure of a complex between Im7 and the DNase domain of ColE7 at 2.3 Å resolution. The structure reveals the detailed interactions and suggests a molecular basis for the extremely tight and specific binding. It also provides a structural view for the novel folding of the DNase domain of ColE7. Possible locations for the DNA-binding sites and the nuclease active site in the DNase domain of ColE7 are discussed.

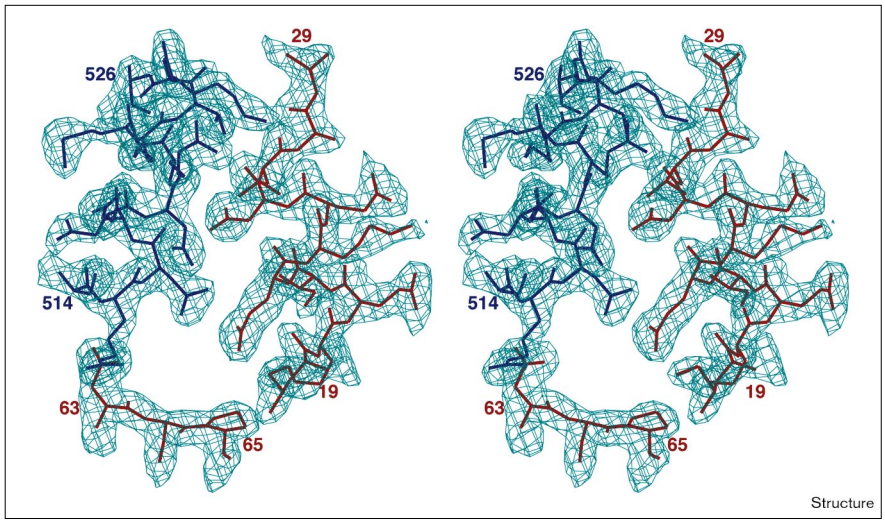
Results

Overall structure

The crystal structure of the protein complex of Im7 and the DNase domain of ColE7 was determined at 2.3 Å resolution by molecular replacement methods in combination with electron-density modification. The electron-density map in a region of the protein interface is shown in Figure 2. The Cα-carbon backbone and the ribbon model of the complex are shown in Figure 3. The final model includes the full length (residues 1–87) of Im7 and the C-terminal DNase domain of colicin E7 from residue 447 to 573. The first 15 residues located at the N-terminal end of the DNase domain of ColE7, including the histidine tag (MRGSHHHHHHGS; single-letter amino acid code)

Figure 2

A stereoview of the electron-density map of the DNase domain of ColE7 complexed with Im7, superimposed onto the final model in a region of the protein–protein interface. The Im7 model is in red and the ColE7 model is in blue. The Fourier coefficients are $(2F_o–F_c)$, with the model phases using data in the 40–2.3 Å range. The contour level is 1σ.



constructed for the purpose of protein purification, and the three C-terminal residues (Arg574, Gly575 and Lys576), were not observed in the structure due to disorder. Im7 in the complex is composed of four α helices (α1*–α4*) folded in a varied four-helix-bundle structure that is identical to the previously determined crystal structure of the free monomeric Im7 [11]. The α helices in Im7 are written with an * to distinguish them from the helices in the DNase domain of ColE7, and the loops in Im7 are renamed as loop12, loop23 and loop34, as shown in Figure 1. The root mean square (rms) deviation between Im7 in the complex and the free Im7 is 0.58 Å for all Cα atoms. Backbones for the N terminus and the C terminus (residues 1 and 2, and 86 and 87) and for loop 12 (residues 30 and 31) have the largest differences. If these regions are disregarded, the rms difference is only 0.47 Å. The sidechains of Im7 in the complex also retain their conformation, except for the sidechain of Trp75, which is different from the one in free Im7. This is probably because Trp75 is involved in the crystal packing in the orthorhombic unit cell of free Im7. Therefore, the formation of the protein complex essentially does not induce any prominent conformational change in Im7.

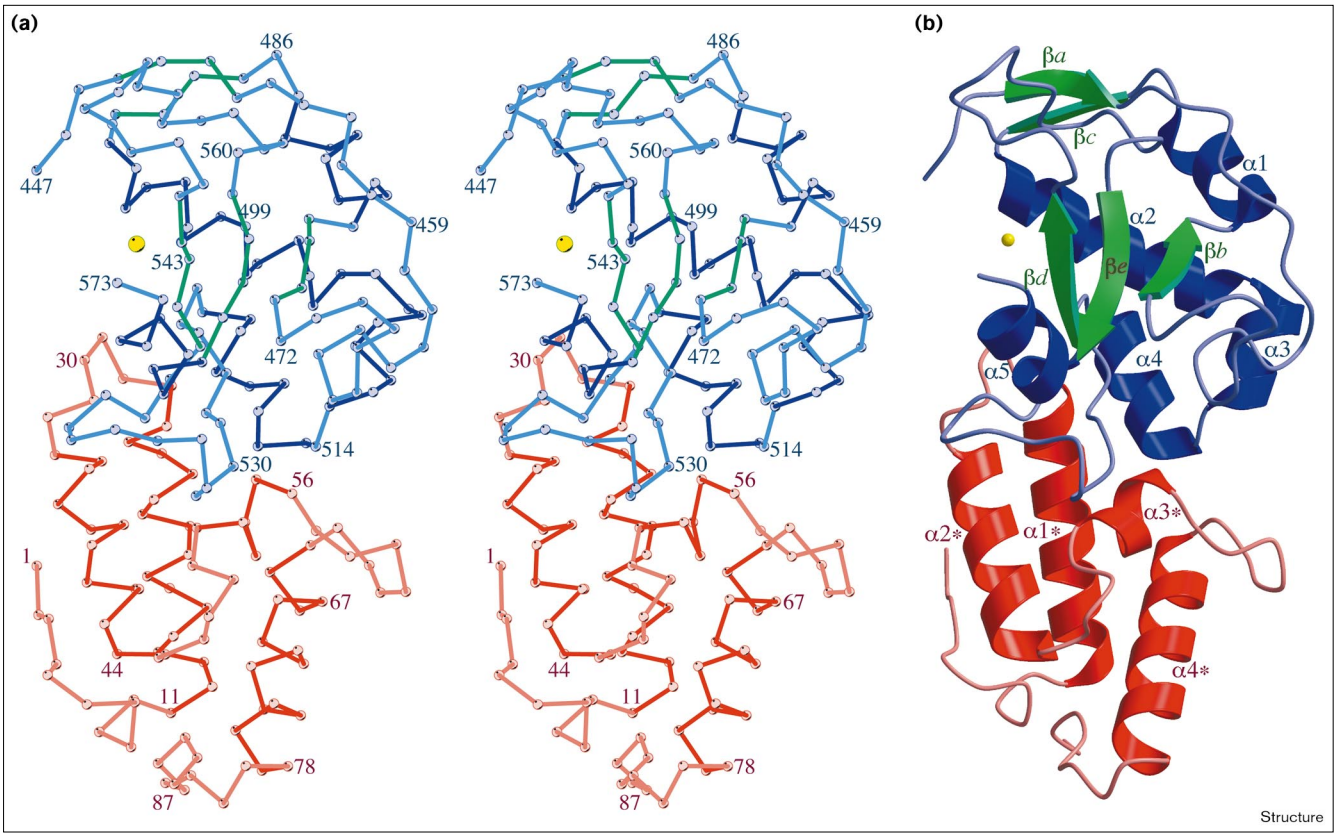
The structure of the DNase domain of ColE7 displays a novel α/β fold. A structural-homology search program using 1389 protein structures in the DALI [21] database revealed no similar fold. The structure contains a three-stranded antiparallel β sheet (βb–βe–βd) that is packed against four α helices (α1–α4). A two-stranded antiparallel β sheet (βa–βc) caps on the top and a small α helix (α5) flanks the other side of the central β sheet. A hydrophobic core is formed between the central β sheet and the α helices that are layered next to the sheet. This core is comprised of 22 residues, including all the methionine, phenylalanine, and tryptophan residues. A Zn^{2+} ion is located at the edge of the central β sheet between βd and α5. This metal atom is coordinated to a water molecule and three histidine residues, His544, His569 and His573. The topology of the two antiparallel β strands (βd and βe) connected to a C-terminal α helix (α5) and the linking of the β strands to the α helix via a Zn^{2+} ion resembles a zinc-finger structure [22]. Thus the metal ion is likely to be essential for the structure to fold properly.

The colicin–immunity-protein Interface

Im7 has a V-shaped structure that clamps around the DNase domain of ColE7, with the α1*–loop12–α2* in one arm and loop23–α3*–loop34 in the other arm. The α4 of ColE7 is aligned in a parallel arrangement with the α1* of Im7. The long and extended loop between α4 and βd of ColE7 interacts extensively with α3* and the neighboring loops of Im7. The solvent-accessible surface areas calculated individually for each of the proteins are 5397 $Å^2$ for Im7 and 7750 $Å^2$ for the DNase domain of ColE7. After complex formation, the buried solvent-accessible surface areas are 1473 $Å^2$ in total; 717 $Å^2$ for Im7 and 756 $Å^2$ for the DNase domain of ColE7. The average buried surface area of 785 $Å^2$ was calculated for each subunit on the basis of ten different proteinase–inhibitor complexes in the Brookhaven Protein Data Bank (PDB) [23]. Therefore the buried surface areas for the complex of the DNase domain of ColE7 with Im7 are comparable with the previously reported values.

Figure 4 shows the molecular surface at the interface of the complex color coded by electrostatic potential. The interacting surfaces are not only highly charged, but are also charge-complementary. The most acidic surface areas in Im7, located around residue Asp31, interact with the most

Figure 3



The structure of the DNase domain of ColE7 in complex with Im7. **(a)** Stereo diagram of the Cα trace for the two proteins: Im7 is shown in red and the DNase domain of ColE7 is in blue and green. The sequence numbers are indicated for each protein. The Zn^{2+} ion is represented by a yellow sphere. **(b)** Ribbon diagram of the overall fold of the complex drawn with the programs MOLSCRIPT [48] and Raster3D [49]. The inhibitor Im7 has a V-shaped arm that clamps the DNase domain of ColE7, thereby inhibiting the colicin's DNase activity. Coloring is as in (a).

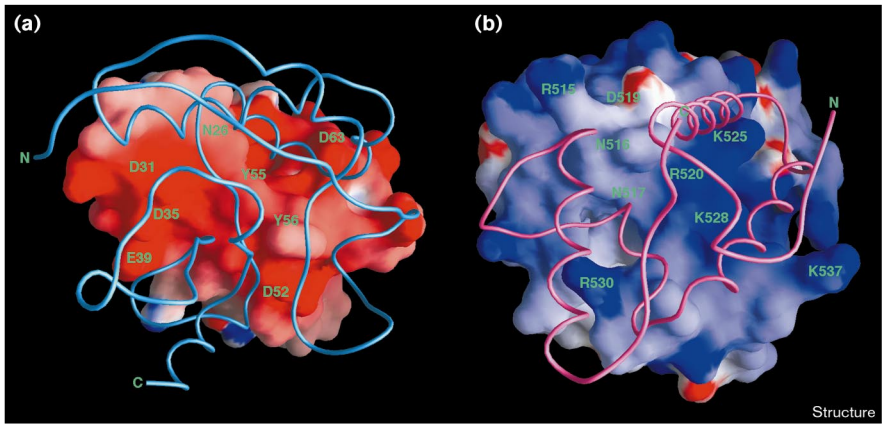
basic areas in the DNase domain of ColE7, located around residue Arg520. The electrostatic interactions must play one of the most important roles in the protein–protein interactions. The surface complementarity is not as obvious, but it can be clearly seen that Tyr56/Im7 protruding out of the interface is well accepted by a nonpolar surface pocket in the corresponding interface area in the DNase domain. The sidechains of Tyr55/Im7 and Asn516/ColE7 bulge out of the surfaces and the corresponding surface depression is observed in the counter molecules.

Residues involved in binding

The residues involved in the interactions can be identified by calculation of the solvent-accessible surface areas before and after complex formation. Figure 5 shows the difference in solvent-accessible surface areas for each of the residues. In Im7, residues Glu23 (40 Å^2), Asn26 (67 Å^2), Val27 (59 Å^2), Asp31 (50 Å^2), Asp35 (38 Å^2), Asp52 (50 Å^2), Tyr55 (90 Å^2), Tyr56 (132 Å^2) and Asp63 (66 Å^2) are most deeply buried (> 30 Å^2) at the interface, within which all the five charged residues are acidic. The two aromatic residues, Tyr55 and Tyr56, contribute the most to the buried surface areas. In the DNase domain of ColE7, the most deeply buried residues at the interface are: Ser514 (55 Å^2), Arg515 (40 Å^2), Asn516 (112 Å^2), Asp519 (43 Å^2), Arg520 (49 Å^2), Val523 (35 Å^2), Lys525 (89 Å^2), Lys528 (89 Å^2), Arg530 (67 Å^2), and Thr531 (62 Å^2). Five out of six charged residues are positively charged and Val523 is the only nonpolar residue.

In total there are sixteen direct hydrogen bonds at the Im7–ColE7 interface (see Table 1). Considering the average size of the interface areas, the number of hydrogen bonds between the two protein molecules is substantial in comparison to the number observed in other enzyme–inhibitor complexes. On average 1.37 hydrogen bonds are formed per 100 Å^2 of interface in ten enzyme–inhibitor complexes [23]. However, in the ColE7–Im7 complex, 2.17 hydrogen bonds are identified per 100 Å^2 of the interface area. Moreover, of the sixteen hydrogen

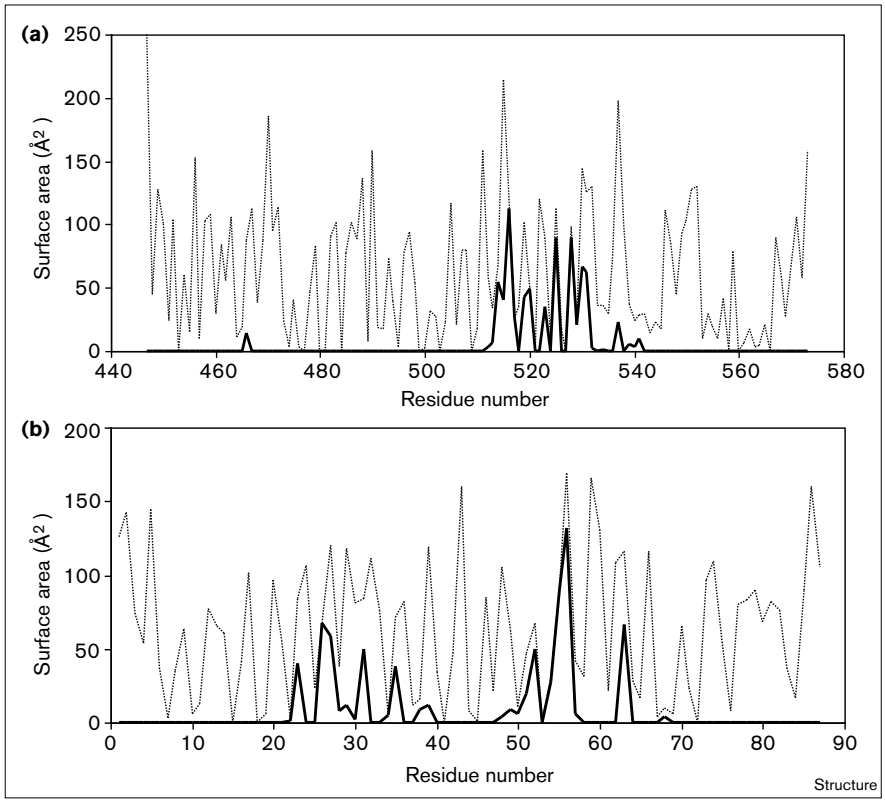
Figure 4



The molecular surfaces of Im7 and the DNase domain of ColE7. **(a)** The molecular surface of Im7 color coded according to electrostatic potential displayed with GRASP [50]. **(b)** The surface of the DNase domain of ColE7. Positive charge is in blue and negative charge in red; (a) is viewed by rotating the molecule in Figure 2 by 90°, with a horizontal axis in the plane of the paper, and (b) is obtained by rotating panel (a) by 180°, with a vertical axis in the plane of paper. An associated complex can be formed by rotating each molecule by 90° inwards around the vertical axis. The two surfaces are highly complementary in charge, in that the most acidic surface around Asp31 in Im7 interacts with the most basic surface around Arg520 in ColE7. Tyr56 in Im7, protruding out of the surface, is accommodated by a pocket in ColE7 between Asn517 and Arg530.

bonds, six involve charged donor and acceptor groups and six others involve one charged group, which should make significant contributions to the free energy of dissociation [24]. The large numbers of salt bridges and hydrogen bonds in the protein interface account for the tight interactions between colicins and immunity proteins. Only the complex of barnase and barstar shows a comparable number of hydrogen bonds, with six salt bridges out of a total of fourteen hydrogen bonds at the protein–protein interface [25,26].

Figure 5



The decrease in solvent-accessible surface area of **(a)** the DNase domain of ColE7 and **(b)** Im7, upon complex formation. The bold lines show the decreased surface and the thin dotted lines represent the solvent-accessible surface for each residue. The exposed and buried surfaces were calculated with Areaimol in CCP4 using a solvent probe radius of 1.4 Å.

Table 1

Hydrogen-bonded interactions and water-mediated hydrogen bonding between the DNase domain of ColE7 and Im7.

Im7 residues	ColE7 residues	Comments
Glu23 OE1	Asn516 OD1	Mediated by Wat850
Glu23 O	Asn516 ND2	Mediated by Wat843
Asn26 ND2	Asp519 OD1	Mediated by Wat843
Asn26 ND2	Asn516 O	d = 2.75 Å
Asn26 OD1	Arg520 NH1	d = 3.60 Å
Asn26 O	Lys525 NZ	d = 3.17 Å
Ala28 O	Lys525 NZ	d = 2.62 Å
Asp31 OD2	Lys525 NZ	d = 2.65 Å; salt bridge
Asp31 OD2	Arg520 NH1	d = 3.03 Å; salt bridge
Asp31 OD1	Arg520 NH2	d = 2.84 Å; salt bridge
Asp35 OD2	Lys528 NZ	d = 2.91 Å; salt bridge
Asp35 OD1	Thr539 OG1	Mediated by Wat835
Glu39 OE2	Lys537 NZ	d = 3.25 Å; salt bridge
Asp49 O	Thr531 OG1	d = 3.19 Å
Thr51 OG1	Thr529 O	Mediated by Wat837
Asp52 OD2	Thr529 O	Mediated by Wat837
Asp52 OD2	Thr531 N	d = 2.91 Å
Asp52 OD2	Thr531 OG1	d = 2.99 Å
Asp52 OD1	Arg530 NH1	d = 2.79 Å; salt bridge
Asp52 OD1	Gln532 N	Mediated by Wat851
Ile54 O	Asn516 OD1	Mediated by Wat839
Tyr55 O	Ser514 OG	d = 2.86 Å
Tyr55 O	Ser514 N	Mediated by Wat841
Tyr55 O	Asn517 ND2	Mediated by Wat841
Tyr56 OH	Lys528 O	d = 2.72 Å
Asp63 OD1	Arg515 N	d = 2.86 Å
Asp63 OD2	Asn516 OD1	Mediated by Wat839

Figure 6 shows the detailed interactions between ColE7 and Im7 with the two arms of Im7 displayed in (a) and (b). One of the arms of Im7 includes the region of α1*–loop12–α2* from residue 23 to 39. This region interacts with the α4 and the C-terminal loop of α4 from residue 516 to 539. The sidechains of Glu23 and Asn26 in α1* make direct or water-bridged hydrogen bonds with Asn516, Asp519, and Arg520 in the helix α4 of the DNase domain of ColE7. The sidechain and mainchain of Asn26 contribute significantly to the interactions that include three direct and one water-bridged hydrogen bond. The negative charge of the three acidic residues, Asp31, Asp35 and Glu39, in α2* of Im7 is compensated by the basic residues of Arg520, Lys525, Lys528, and Lys537 in ColE7. The sidechain of Asp31 contributes significantly to the interaction because it makes two ion pairs with Arg520 and Lys525, within which three hydrogen bonds are formed. The importance of residue Asp31 in the interactions is addressed in the Discussion section. Val27 in Im7 makes van der Waals contacts with Val523 of ColE7, and the shortest distance between the two residues is 3.9 Å.

On the other side of the V-shaped arms, as shown in Figure 6b, the helix α3* and the loops connecting to helix α3* of Im7 from residue 49 to 63 interact with the N terminus of α4 and the following loop. In this region, Asp52 of Im7 makes one salt bridge with Arg530, two hydrogen bonds with Thr531, and two water-mediated hydrogen bonds with the mainchain of Thr529 and Gln532 of ColE7. The two tyrosine residues, Tyr55 and Tyr56, not only make several hydrogen bonds, but also van der Waals contacts, with the sidechains of Lys528 and Arg530 in the DNase domain of ColE7. As mentioned earlier, these two tyrosine residues in Im7 are the most deeply buried residues at the interface; thus they must contribute substantially to the interaction. Pro57 of Im7 makes van der Waals contacts with the sidechain of Ser514. It should be noted that the important residues involved in interactions in this region, including Asp52, Tyr55, Tyr56 and Pro57, are all conserved in the DNase-type immunity-protein subfamily. Moreover, there are seven direct or water-mediated hydrogen bonds in this region involving mainchain atoms in ColE7. The choice of using more of the mainchain atoms of ColE7 in the interactions is probably because the counter residues in Im7 are conserved. The top half of Table 1 lists all the hydrogen bonds formed by the first arm of Im7 and the bottom half lists those formed by the second arm of Im7.

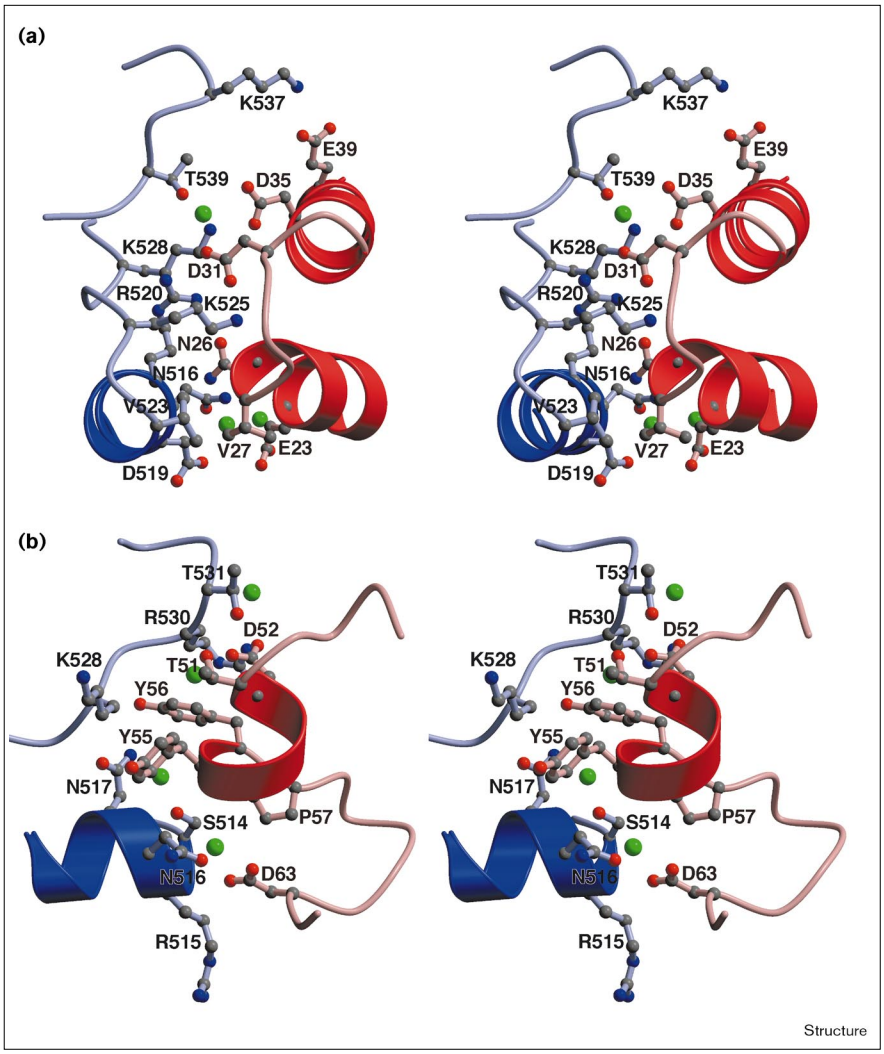
Water molecules at the interface

There are at least 28 water molecules in the vicinity of the protein-interaction interface. The average temperature factor for these water molecules is 36.4 $Å^2$, a value lower than that of the DNase domain, indicating that these waters are quite ordered. Among these water molecules, 24 are within a distance of 5 Å from both protein molecules, and seven water molecules are directly involved in the water-bridged hydrogen bonds between Im7 and the DNase domain of ColE7 that form eleven hydrogen bonds, as shown in Table 1 and Figure 6. Five water molecules are involved in a 'double-water-bridged' hydrogen bond in the style of Im7–water–water–ColE7. The remaining 12 water molecules are located at the rim of the interface and they presumably also participate in stabilizing the interface. The analyses of the surface areas in the context of the entire model with program Areaimol in the CCP4 suite shows that 8 out of the 28 interface water molecules are completely inaccessible to the bulk solvent (accessible area = 0 $Å^2$) and 5 others are mostly buried (area < 5 $Å^2$). Thus, these water molecules may also contribute to the shape complementarity between the two proteins in the complex.

Metal-binding site

The highest peak in the electron-density map (> 7σ) was assigned as a Zn^{2+} ion based on the measurement of atomic-emission spectroscopy and by virtue of its associated ligands. However, the relevance of the *in vivo* biological significance of the Zn^{2+}-bound enzyme is not certain as a recent report shows that the DNase domain of ColE9 has optimal activity in the presence of divalent metal ions of Ni^{2+}, Co^{2+}, or Mg^{2+} [27]. The Zn^{2+} ion is bound by three histidine residues:

Figure 6



Stereoview of the interactions between the DNase domain of ColE7 and Im7. **(a)** The sidechains of Im7 in the region of α1*–loop12–α2* interact with the sidechains of the residues in the DNase domain. This region of Im7 dominates its specificity. Coloring is as in Figure 3. The green spheres represent water molecules. The detailed hydrogen-bond distances are listed in the top half of Table 1. **(b)** The sidechains in the region of loop23–α3*–loop34 in Im7 interact not only with the sidechain, but also with the mainchain atoms of the DNase domain of ColE7. The detailed hydrogen-bond distances are listed in the bottom half of Table 1.

His544 (ND1), His569 (NE2), and His573 (NE2), and a water molecule (see Figure 7). The geometry of the Zn^{2+} ligand's structure is a distorted tetrahedral configuration, with the ligand–metal–ligand angles ranging from 79° to 134°. The bond distance is 2.06 Å for Zn^{2+}–ND1 of His544, 1.92 Å for Zn^{2+}–NE2 (His569), 2.57 Å for Zn^{2+}–NE2 (His573) and 2.16 Å for Zn^{2+}–OH_2. A second water is hydrogen bonded to the Zn^{2+}-bound water and ND1 of His545. The structure of a Zn^{2+} ion bound to three histidines and one water molecule in a tetrahedral geometry has been observed before in the structures of carbonic anhydrase [28], β-lactamase, and DD-Carboxypeptidase [29].

Discussion

Specific interactions between colicins and immunity proteins

The protein-recognition sites in Im7 can be described as being composed of two separate arms in a V-shaped structure. One arm is located in the α1*–loop 12–α2* region, the sequence of which is highly variable among members of the DNase-type immunity-protein subfamily. The other arm is located in the region of loop 23–α3*–loop34, and the sequence of this region is more conserved. These two arms interact with a continuous region in ColE7, the α4 helix and the following long loop. The protein-recognition site in ColE7 is the most variable region in the subfamily of DNase-type colicins (see Figure 1). Thus, a variable region in colicin interacts not only with a variable arm, but also with a more conserved arm in Im7. Almost all of the hydrogen bonds in the region of the variable arm of Im7 involve sidechains of ColE7 and there is only one hydrogen bond to a mainchain atom in ColE7. By contrast, in the region of the more conserved arm, there are three direct hydrogen bonds and four water-mediated hydrogen bonds involving mainchain atoms in ColE7. Therefore, ColE7 chiefly uses sidechains to interact with the variable residues and mainchains to interact with the conserved residues of Im7.

The specificity of each of the immunity proteins for their cognate colicin is likely to be mediated predominantly by the variable arm. In fact, the variable arm of Im7 participates in four out of five ion pairs, whereas the conserved arm participates in only one ion pair. Therefore the variable arm in Im7 is not only important for specificity, but it also contributes significantly to the tight binding between the two proteins. The hydrophobic interactions contributed by the two conserved tyrosine residues (Tyr55 and Tyr56) are also important for the high affinity. Tyr55 inserts into a nonpolar pocket between Asn516 and Lys528. This pocket may also be present in ColE2 and ColE9, in which the lysine residue is replaced by phenylalanine, and may have even higher affinity for the tyrosine sidechain. A dual-recognition model has been proposed for colicin–immunity-protein interactions [30]. It suggests that the conserved region in Im7 dominates the strength of the binding and that the neighboring variable residues control specificity. In the crystal structure of the complex of the DNase domain of ColE7 with Im7, indeed, the variable and conserved residues in Im7 both contribute to the interactions and the variable residues are probably responsible for determining the specificity of immunity.

Structural interpretation of biochemical properties of mutant immunity proteins

It has been demonstrated before that single-site mutations usually cannot change the phenotype of the DNase-type immunity proteins. For this reason, before the complex structure could be resolved, it was difficult to determine the specific residues essential for the interactions. We have changed five different negatively charged residues in Im7: Glu14, Glu23, Glu25, Asp31, and Asp35 to alanine; the mutant proteins did not exhibit any change in the *in vivo* specific immunity to ColE7 (data not shown). In a previous study, ten residues in this region of Im9 were changed, in turn, to the corresponding residues of Im8 and, except for Val34–Asp and Val37–Glu, which showed minor increases in Im8 activity, none of the other mutant proteins changed their phenotype [31]. We believe that, because there are so many residues involved in the interactions, a single-site mutation is not enough to abolish the entire interaction network; other residues are able to maintain the network.

NMR perturbation experiments showed that most of the strongly shifted amides in Im9, while Im9 was interacting with colicin E9, were located on the surfaces of α2* and α3* [32]. In another case, several chimera of Im9 were constructed that involved the substitution of different fragments of Im9 into the corresponding regions of Im2. Binding assays suggested that the specificity of immunity proteins was determined predominantly by residues on α2*, with the neighboring residues located in α1* and loop34 playing a minor role [33]. Very recently, alanine scanning mutagenesis data further revealed ten residues in Im9 important for binding the DNase domain of colicin E9 [34]. All these results are in agreement with our complex structure — that α2* and α3* indeed contribute most significantly to the interactions. However, the structure further demonstrates that the C terminus of α1* and surface loops (loop12, loop23 and loop34) are also involved.

Recently, we constructed a triple mutant of Im7 (Asp31–Asn/Asp35–Asn/Glu39–Gln) that completely lacked *in vivo* inhibitory activity against ColE7 (F-M Lu, HSY, S-J Chang and K-FC, unpublished results). Moreover, the inhibitory effect of the residues of Im7 against ColE7 can be ordered such that Asp31 > Asp35 > Glu39 (F-M Lu *et al.*, unpublished results). This result is quite consistent with the crystal structure of the complex of the DNase domain of ColE7 with Im7 in that the three acidic residues are on the same side of α2* and they all participate in the interactions in which Asp31 contributes the most (three hydrogen bonds), Asp35 the next (one direct and one water-mediated hydrogen bond) and Glu39 the least (one hydrogen bond). Mutation of the three critical residues in Im8 at the equivalent positions (Glu29, Asp33 and Glu37) eliminated specific immunity of Im8 to its cognate colicin E8 (F-M Lu, *et al.*, unpublished results). This result indicates that the residues in the region of α1*–loop12–α2* contribute significantly to the specificity of the protein–protein interactions. This conclusion is consistent with our earlier prediction based on the crystal structure of Im7 [11]: that α1*, α2* and the surface loops are involved in interactions, and, moreover, that the most acidic and variable region in Im7 located in α1*–loop12–α2* plays a major role in specific binding to colicins.

Putative DNA-binding region and DNase active site for ColE7

The DNase-type colicins have *in vivo* cytotoxic activity and a high *in vitro* activity in cleaving various DNA molecules, as previously demonstrated for ColE2 [35] and ColE9 [36]. It has been shown that the replacement of either Arg544, Glu548 and His575 with alanine in ColE9 (corresponding to Arg538, Glu542, and His569 in ColE7, respectively) completely inactivated the DNase activity. Replacement of another conserved and metal-binding residue, His579, with alanine in ColE9 (corresponding to His573 in ColE7), however, did not affect DNase activity. This result implies that His579 in ColE9 (or His573 in ColE7) is not involved in catalysis, and as the metal–ligand bond distance for the equivalent His573 of ColE7 is the longest one among the others, its binding to metal ion is probably not critical for maintaining the structure of the enzyme. On the other hand, Arg538, Glu542 and His569 in ColE7 are conserved in all of the DNase-type colicins, and disruption of the DNase activity by mutation suggests that they may be involved in the DNA cleavage reaction. These three catalytically important residues are located near the Zn^{2+} ion, where His569 is directly bonded to the metal. It has been shown previously that a Zn^{2+}-bound water molecule has a

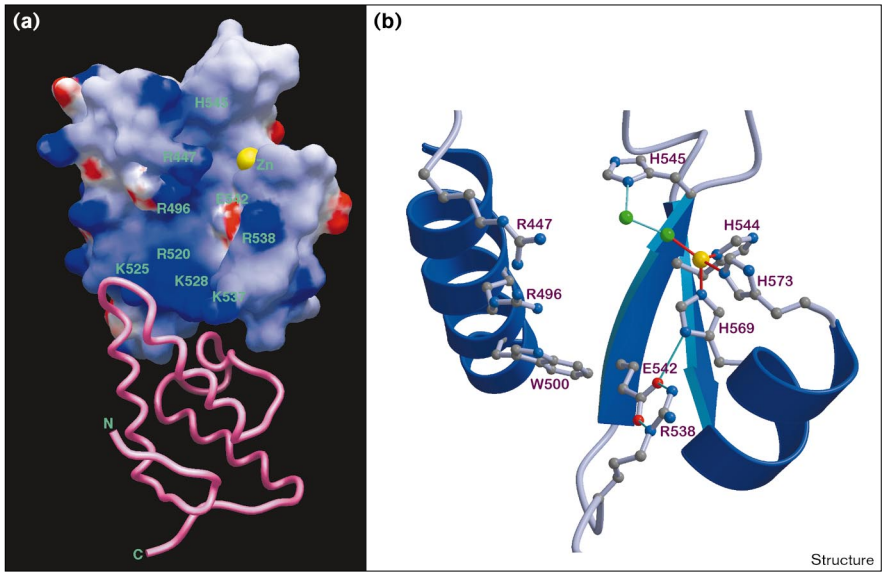
reduced pKa and is usually a critical component of the catalytically active Zn^{2+} sites [29]. We speculate that in the DNase domain of colicins, the metal-bound water molecule may serve as a nucleophilic hydroxide ion that attacks the phosphate atom. Similar mechanisms utilizing a Zn^{2+}-bound hydroxide ion to attack the carbonyl group of substrates have been observed before in carboxypeptidase A structure of *Penicillium citrinum* P1 nuclease also suggested a cleavage mechanism for the single-stranded DNA phosphodiester bonds involving the nucleophilic attack by a Zn^{2+}-activated water molecule [39].

In addition to the metal-binding site, there is also a cleft near the three important residues for activity. This cleft is located between the central β sheet and the α helices layered next to the sheet, extending from the metal-binding site to the densely positively charged surface where Im7 binds, as shown in Figure 7. Several polar or charged residues, Arg447, Arg496, Arg538, Ser540, and Glu542, flank the walls of the cleft, and the sidechains of Phe492, Trp500, Ala526, and Leu543 form the floor. Again, we speculate that the cleft is a binding site for the DNA substrate. One particularly interesting residue is Glu542, which is salt-bridged to Arg538 and also hydrogen-bonded to the NE2 atom of His569. Mutation of the three equivalent residues in ColE9 abolished the DNase activity, suggesting that they play a vital role in the hydrolysis reaction. It seems likely that these three residues are involved in carrying a proton to the leaving group. However, we must emphasize here that the mechanism for the cleavage catalyzed by colicins is completely unknown. The active site proposed here for the DNase domain of ColE7 is based purely on the structure and very limited information regarding mutagenesis and can therefore only be used as a starting point for further analysis.

The Zn^{2+}-binding site is quite distant from the protein–protein interaction interface. How a Im7 protein inhibits the DNase activity of ColE7 is unclear. The three mutants of ColE9 with no DNase activity are still able to bind DNA, implying that the major DNA-binding site for the enzyme is probably extending beyond the active site for cleavage. Assuming that the metal-binding site is indeed the active site, the binding of the immunity protein to colicin either induces conformational changes around the active site so that it loses DNase activity, or the immunity protein blocks the major DNA-binding site such that the DNA cannot bind to the DNase domain for further cleavage. A very important residue, Asp39, in barstar contributes three hydrogen bonds in the barstar–barnase interaction and was found to occupy the phosphate-binding site of barnase [26]. Asp31 in Im7, as well as other acidic residues, may play a similar role in mimicking DNA. The counter surface of Asp31 in the DNase domain of ColE7 is the most basic region, it is therefore possible that this region is primarily responsible for binding of acidic DNA molecules. Nevertheless, we cannot rule out the possibility that the metal-binding site may not be the active site as we suggest here, and that perhaps the real active site is close to or within the protein interface. The DNase domain of colicin E7 exhibits a fold not yet seen in other nucleases. As little experimental data are available, there is not enough information to elucidate the detailed mechanism for DNA-binding and cleavage. In the future, further structural or biochemical studies must be carried out in order to understand how DNase-type colicins bind and cleave DNA.

Figure 7

The metal-binding site in the DNase domain of ColE7. **(a)** The metal atom is located in a cleft, the wall of which is flanked by several charged residues. The metal-binding site is far from the protein–protein interface. **(b)** Three histidine residues, His544, His569, and His573, and a water molecule are bound to the Zn^{2+} ion in a distorted tetrahedral structure. A second water is hydrogen-bonded to the metal-bound water and His545. The water molecules are represented by green spheres and the metal atom by a yellow sphere.



Biological implications

Colicins have been used for decades in a wide variety of studies of different aspects of biology. Thirty to fifty percent of enterobacteria have been shown to contain colicin plasmids. The cytotoxicity of each of the colicins is only completely inhibited by its cognate immunity protein. We still do not fully understand the mechanism by which an immunity protein can accurately select and bind to its cognate colicin, however. The detailed structural analysis of a colicin bound by its cognate immunity protein has the potential to advance our understanding of protein recognition in general. Here, we present the structure of the DNase domain of colicin E7 in complex with its inhibitor Im7. The interaction between ColE7 and Im7 is characterized by the highly charged complementary interfaces and numerous intermolecular hydrogen bonds, 75% of which involve charged groups. The hydrophobic interactions are also present at the interface, but the protein-interface areas are less hydrophobic than other protein–protein interactions.

In the past, many structures of proteinase–inhibitor complexes and one structure of a ribonuclease–inhibitor complex have been solved, but complexes of a DNase and its inhibitor are rare. Moreover, the complex structure presented here reveals a novel fold for the DNase domain of ColE7 in which a metal ion is present. The fold of the DNase domain of ColE7 is quite dissimilar to those of other endonucleases, exonucleases, recombinases, or topoisomerases. On the basis of the coordination of the metal ion, it seems likely that the metal-bound water and the residues near the metal are involved in the phophodiester-hydrolysis reaction. The DNase domain of ColE7 may form a new class of nucleases that cleave DNA in a manner not previously seen. The structure of the DNase domain of colicin E7 provides a good starting point for the full characterization of the mechanism of DNA binding and cleavage by colicins.

Materials and methods

Construction of the vector for overexpression of the DNase domain of ColE7 and Im7

The expression vector pQE-30 (Qiagen, Germany) containing a six-histidine affinity tag at the N terminus of the cloning site was used to clone the DNase domain – *ceiE7* of the operon. *E. coli* SG13009 (*nal*^{s}, *str*^{s}, *rif*^{s}, *lac*$^{-}$, *ara*$^{-}$, *gal*$^{-}$, *mtl*$^{-}$, *recA*$^{+}$, *uvr*$^{+}$) was used as the host for the expression vector and the recombinant plasmid. All cultures were routinely grown in Luria–Bertaini (LB) broth or on plates of LB agar and were supplemented, where required, with 50 μg/ml of the appropriate antibiotic.

Plasmid pHK001 [15] containing a completed ColE7 operon was used as a template for polymerase chain reaction (PCR) amplification. A 753 base pair (bp) DNA fragment containing the DNase domain [36] and the entire *ceiE7* gene of the ColE7 operon [15] was amplified by PCR using a pair of oligonucleotide primers: PT2A-F (5′CAAATTAGGATCCGAGAGTAAACGGAATAAG3′) and PT1-R (5′TATTTTCTGCAGAACGACTCCTTGTTGTG3′). A *Bam*HI and a *Pst*I site were generated at either end of the PCR-amplified fragment (represented by underlining in primer sequence). PCR reactions were carried out in 100 μl volumes using one unit of *Taq* polymerase with 30 cycles of 94°C for one minute, 55°C for 30 seconds and 72°C for one minute. The PCR-amplified fragment was then cleaved with *Bam*H1 and *Pst*I and ligated into the expression vector pQE-30 incubated with the same restriction enzymes. The resulting recombinant plasmid, pHBH, was transformed into *E. coli* SG13009.

Purification of overexpressed ColE7 DNase domain and Im7

Overnight *E. coli* SG13009 (pHBH) cultures were diluted 100-fold in 200 ml LB. Cells were incubated at 37°C, with shaking at 200 rpm. Isopropyl-β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 1 mM to induce the expression of the DNase domain and Im7. The induced cultures were continuously incubated at 37°C for another five hours before harvesting. Cells were harvested and resuspended in a 150 ml sonication buffer (50 mM Na phosphate, 300 mM NaCl, pH 7.8) containing 150 mg lysozyme. After incubation at 37°C for one hour, cells were extensively sonicated until most of the cells were broken according to observation under a phase-contrast microscope. Cell debris of sonicated cultures was removed by centrifugation (10000× g for 30 min). The resulting crude cell extracts were then filter sterilized.

Crude cell extracts (150 ml) were loaded into a Ni-NTA resin affinity column (Qiagen, Germany). After loading, the column was washed once with 100 ml washing buffer (50 mM Na phosphate, 300 mM NaCl, pH 6.0) to remove unbound protein. The bound protein was then eluted by an imidazole gradient (0–0.5 M). The eluent was concentrated and then diluted tenfold with 10 mM Na phosphate, pH 7.0. The resulting protein solution was loaded into a CM column (CM Sepharose Fast Flow, Pharmacia, Sweden). The bound protein was eluted by a NaCl gradient (0–0.5 M). Elution of protein was monitored by spectrophotometry at 280 nm. The single-band region of the elution was collected and the protein purity was examined by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS–PAGE). The mass spectroscopy for the purified proteins indicated a molecular mass of 16,702 Da for the DNase domain of ColE7 (16,641 for the calculated value, excluding the Zn^{2+} ion) and 9894 Da for Im7 (9895 for the calculated value).

Atomic-emission measurement

The contents of metal elements in the protein complex were measured on a Jarrell–Ash model ICP 9000 inductively coupled plasma atomic emission spectrophotometer. A Zn content of 1.07 μM, a Ni content of 0.15 μM, and a Mg content of 0.56 μM were detected from 2.6 μM of protein, corresponding to the molar ratio of protein: Zn: Ni: Mg = 1: 0.4: 0.06: 0.2.

Crystallization and data collection

The crystals of the complex were obtained by the hanging-drop vapor-diffusion method. Drops of a solution containing 15 mg/ml of protein complex, 5 mM sodium citrate (pH 5.6), 0.25 M ammonium acetate, and 10% PEG4000 were set up against a reservoir of 22.5% PEG4000. Plates of crystal appeared within about a week at room temperature. These complex crystals diffracted X rays to nearly 2.0 Å resolution under cryogenic conditions and belonged to orthorhombic unit cells with a body-centered spacegroup of either I222 or $I2_12_12_1$. X-ray diffraction data were recorded with synchrotron X radiation (λ = 1.00 Å) at room temperature using the Sakabe Weissenberg camera on beamline 6B at the KEK Photon Factory (Tsukuba, Japan). The unit-cell parameters are: a = 65.5 Å, b = 76.1 Å and c = 120.1 Å. An asymmetric unit consists of one DNase domain of ColE7 complexed with Im7, and the Matthews' volume (V_m) is 2.8 Å^3/dalton, corresponding to 54% solvent content [40]. The data were processed using the HKL package [41]; Table 2 lists the statistics.

Structure determination and refinement

Several heavy-atom derivatives were tested, but none was effective for solving the crystal structure by the multiple isomorphous replacement method. However, the crystal structure was successfully determined by the molecular-replacement method using the known tertiary structure of Im7 as a search model in the program X-PLOR [42]. Several data sets and all of the orthorhombic (one) and monoclinic (four) models of Im7 were tested [11,12], but only the combination of the orthorhombic Im7

Table 2

X-ray diffraction and refinement statistics for the DNase ColE7–Im7 complex crystal.

Data collection	
Unit cell (a, b, c; Å)	65.4, 76.0, 120.0
Resolution (Å)	40–2.3
Number of observations	38,306
Unique reflections	12,090
Completeness (%)	88.5
R_{merge}*(%)	4.3
Structure refinement	
Number of reflections ($F > 2\sigma_F$)	10,498
R value based on 92% data	20.3
R_{free} for 8% test data set	27.0
Number of nonhydrogen atoms	1888
Rmsd from ideal	
Bond length (Å)	0.005
Bond angle (°)	1.012
Most favored (φ, ψ) angles excluding Gly, Pro (%)	87.6
Average B values Im7	
All atoms (Å^2)	29.7 (698)†
Backbone atoms (N, CA, C, O)	29.6 (348)†
Sidechain atoms	29.8 (350)†
Average B values DNase ColE7	
All atoms (Å^2)	41.2 (1027)†
Backbone atoms (N, CA, C, O)	40.8 (519)†
Sidechain atoms	41.6 (508)†
Average B value solvent atoms (Å^2)	52.0 (162)†

*$R_{merge} = \sum_h \sum_i | Ih,i - \langle Ih \rangle | / \sum_h \sum_i | Ih,i$, where $\langle Ih \rangle$ is the mean intensity of the i observations for a given reflection h. †Number of atoms.

model and the synchrotron data set yielded a unique solution. The orientation of the Im7 molecule in a I222 unit cell was specified by $(\theta_1, \theta_2, \theta_3) = (266°, 86°, 262°)$ and its center position by $(t_x, t_y, t_z) = (0.355, 0.389, 0.107)$. The rotation solution is 6.7σ above mean value and 2.6σ higher than the second solution. The translation solution is 6.54σ above mean value and 2.4σ higher than the second solution. The space group of $I2_12_12_1$ did not yield any solutions. This model gave an initial R value of 51.6% for data between 8 Å and 3.2 Å resolution. Subsequent rigid-body refinement using all data between 40 Å and 2.3 Å gave an R value of 46.2%. In the unit cell, Im7 molecules were packed back to back, forming four-membered clusters while the surface predicted for interactions with colicin was exposed. No bad intermolecular contacts were observed. All 87 residues in the Im7 molecule only accounted for about 40% of the asymmetric unit. However, in the electron-density map phased by this model, several fragments of densities for the DNase moiety of the complex were observed. Instead of building a protein model into the map at this stage, 109 water molecules were placed in the regions with strongest density.

Throughout the refinement, 8% of randomly selected data were set aside for cross validation with free R values, and bulk-solvent setup was also employed after the complete model was built to include the low-resolution data. After 20 separate routines of simulated annealing, the models gave R values of ~32% and an R_{free} of ~45%. All 20 trial models were combined and used as the starting models for density modification (DM) with solvent flattening and histogram matching [43]. The ensuing Fourier map clearly showed both the backbone and the sidechains of two α helices adjacent to the Im7 molecule. These were assigned as residues 490–530 in the colicin E7 sequence, as strongly implicated by the shape of the sidechain densities. In addition to this 40-residue fragment, there were still many other densities for the DNase domain. Again, 178 water molecules were added and the model was further refined by simulated annealing.

This second model yielded an R value of 29 or 30% and an R_{free} of 43–45%. Likewise, a second electron-density map was produced using DM and more amino acid residues were identified in the DNase domain. The second region built into the model comprised residues 555–570 near the C-terminal end of ColE7. The dipeptide of Trp464–Leu465 could also be seen in the map at this stage. Another 186 water molecules were added to the model, and after refinement, the R value and R_{free} were reduced to 27% and 42%, respectively. Such routines were repeated until the entire polypeptide fold of the DNase domain was elucidated. Apart from the N-terminal and C-terminal segments of the DNase domain, all of the amino acid residues were built into the map, rendering a final model comprising residues 1–87 of Im7, residues 447–573 of ColE7, 162 water molecules, and a bound Zn^{2+} ion. Nevertheless, regions around residues 470 and 550 of the ColE7 still had weak densities. The final model yielded an R value of 20.3% and an R_{free} of 27.0% for 10,498 reflections in the resolution range of 40.0–2.3 Å with $F > 2\sigma_F$; or an R value of 21.3% and an R_{free} of 0.277 for all of the 12,083 reflections with $F > 0$ in the same resolution range. All atoms had temperature factors of less than 70 Å^2, apart from residues 550 and 551 of the ColE7 and 17 water molecules, while the Zn^{2+} ion had a B factor of 51.0 Å^2. Statistics of the crystallographic refinement are also listed in Table 2, and the electron-density map is shown in Figure 2. The Ramachandran plot showed 87.6% of the amino acid residues located in the most favored region, 12.4% in the additional allowed region, and none of the residues fell in the disallowed region.

Accession numbers

The atomic coordinates of the DNase domain of ColE7 in complex with Im7 have been deposited in the Brookhaven Protein Data Bank with accession code 7CEI.

Acknowledgements

We thank Ming F Tam for the measurement of molecular mass. This work was supported by research grants and post doctoral fellowships from Academia Sinica and the National Science Council of the Republic of China to T-P Ko (Academia Sinica), K-F Chak (NSC87-2314-B010-009 and NSC88-2316-B010-016), and HS Yuan (NSC87-2311-B001-125 and NSC88-2311-B001-017).

References

1. Pugsley, A.P. & Oudega, B. (1987). Methods for studying colicins and their plasmids. In *Plasmids – A Practical Approach.* (Hardy, K.G., ed.) pp. 105-161, IRL Press, Oxford & Washington.
2. Riley, M.A. & Gordon, D.M. (1992). A survey of Col plasmids in natural isolates of *Escherichia coli* and an investigation into the stability of Col-plasmid lineages. *J. Gen. Microbiol.* **138**, 1345-1532.
3. Witkin, E.M. (1976). Ultraviolet mutagenesis and inducible DNA repair in *Escherichia coli*. *Bacteriol. Rev.* **40**, 869-907.
4. Akutsu, A., Masaki, H. & Ohta, T. (1989). Molecular structure and immunity specificity of colicin E6, an evolutionary intermediate between E-group colicins and cloacin DF13. *J. Bacteriol.* **171**, 6430-6436.
5. Bowman, C.M., Sidikaro, J. & Nomura, M. (1971). Specific inactivation of ribosomes by colicin E3 *in vitro* and mechanism of immunity in colicinogenic cells. *Nature New Biol.* **234**, 133-137.
6. James, R., Kleanthous, C. & Moore, G.R. (1996). The biology of E colicins: paradigms and paradoxes. *Microbiology* **142**, 1569-1580.
7. Lakey, J.H., van der Goot, F.G. & Pattus, F. (1994). All in the family: the toxic activity of pore-forming colicins. *Toxicology* **87**, 85-108.
8. Wiener, M., Freymann, D., Ghosh, P. & Stroud, R.M. (1997). Crystal structure of colicin Ia. *Nature* **385**, 461-465.
9. Vetter, I.R., Parker, M.W., Tucker, A.D., Lakey, J.H., Pattus, F. & Tsernoglou, D. (1998). Crystal structure of a colicin N fragment suggests a model for toxicity. *Structure* **6**, 863-874.
10. Yajima, S., *et al.*, & Uosumi, T. (1993). The three-dimensional structure of the colicin E3 immunity protein by distance geometry calculation. *FEBS Lett.* **333**, 257-260.
11. Chak, K.-F., Safo, M.K., Ku, W.-Y., Hsieh, S.-Y. & Yuan, H.S. (1996). The crystal structure of the ImmE7 protein suggests a possible colicin-interacting surface. *Proc. Natl Acad. Sci. USA* **93**, 6437-6442.
12. Hsieh, S.-Y., Ko, T.-P., Tseng, M.-Y., Ku, W.-Y., Chak, K.-F. & Yuan, H.S. (1997). A novel role of ImmE7 in the autoregulatory expression of ColE7 operon and indentification of possible ribonuclease active sites in the crystal structure of dimeric ImmE7. *EMBO J.* **16**, 1444-1454.

13. Dennis, C.A., *et al*., & Kleanthous, C. (1998). A structural comparison of the colicin immunity proteins Im7 and Im9 gives new insights into the molecular determinants of immunity-protein specificity. *Biochem. J.* **333**, 183-191.
14. Osborne, M.J., *et al*., & Moore, G.R. (1996). Three-dimensional solution structure and ^{13}C nuclear magnetic resonance assignments of the colicin E9 immunity protein Im9. *Biochemistry* **35**, 9505-9512.
15. Chak, K.-F., Kuo, W.-S., Lu, F.-M. & James, R. (1991). Cloning and characterization of the ColE7 plasmid. *J. Gen. Microbiol.* **137**, 91-100.
16. Cole, S.T., Saint-Joanis, B. & Pugsley, A.P. (1985). Molecular characterisation of the colicin E2 operon and identification of its products. *Mol. Gen. Genet.* **198**, 465-472.
17. Toba, M., Masaki, H. & Ohta, T. (1988). Colicin E8, a DNase which indicates an evolutionary relationship between colicins E2 and E3. *J. Bacteriol.* **170**, 3237-3242.
18. Eaton, T. & James, R. (1989). Complete nucleotide sequence of the colicin E9 (*cei*) gene. *Nucleic Acids Res.* **17**, 1761-1761.
19. Wallis, R., Moore, G.R., James, R. & Kleanthous, C. (1995). Protein–protein interactions in colicin E9 DNase-immunity protein complexes. 1. Diffusion-controlled association and femtomolar binding for the cognate complex. *Biochemistry* **34**, 13743-13750.
20. Wallis, R., *et al*., & Kleanthous, C. (1995). Protein–protein interactions in colicin E9 DNase-Immunity protein complexes. 2. Cognate and noncognate interactions that span the millimolar to femtomolar affinity range. *Biochemistry* **34**, 13751-13759.
21. Holm, L. & Sander, C. (1993). Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* **233**, 123-138.
22. Pavletich, N.P. & Pabo, C.O. (1991). Zinc-finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 Å resolution. *Science* **252**, 809-817.
23. Jones, S. & Thornton, J.M. (1996). Principles of protein-protein interactions. *Proc. Natl Acad. Sci. USA* **93**, 13-20.
24. Janin, J., Miller, S. & Chothia, C. (1988). Surface, subunit interfaces and interior of oligomeric proteins. *J. Mol. Biol.* **204**, 155-164.
25. Buckle, A.M., Schreiber, G. & Fersht, A.R. (1994). Protein–protein recognition: crystal structure analysis of a barnase-barstar complex at 2.0 Å resolution. *Biochemistry* **33**, 8878-8889.
26. Guillet, V., Lapthorn, A., Hartley, R.W. & Mauguen, Y. (1993). Recognition between a bacterial ribonuclease, barnase, and its natural inhibitor, barstar. *Structure* **1**, 165-176.
27. Pommer, A.J., Wallis, R., Moore, G.R., James, R. & Kleanthous, C. (1998). Enzymological characterization of the nuclease domain from the bacterial toxin colicin E9 from *Escherichia coli*. *Biochem. J.* **334**, 387-392.
28. Lindskog, S. (1997). Structure and mechansim of carbonic anhydrase. *Pharmacol. Ther.* **74**, 1-20.
29. Vallee, B.L. & Auld, D.S. (1990). Active-site zinc ligands and activated H_2O of zinc enzymes. *Proc. Natl Acad. Sci. USA* **87**, 220-224.
30. Kleanthous, C., Hemmings, A.M., Moore, G.R. & James, R. (1998). Immunity proteins and their specificity for endonuclease colicins: telling right from wrong in protein–protein recognition. *Mol. Microbiol.* **28**, 227-233.
31. Wallis, R., Moore, G.R., Kleanthous, C. & James, R. (1992). Molecular analysis of the protein–protein interaction between the E9 immunity protein and colicin E9. *Eur. J. Biochem.* **210**, 923-930.
32. Osborne, M.J., *et al*., & Moore, G.R. (1997). Identification of critical residues in the colicin E9 DNase binding region of the Im9 protein. *Biochem. J.* **323**, 823-831.
33. Li, W., Dennis, C.A., Moore, G.R., James, R. & Kleanthous, C. (1997). Protein–protein interaction specificity of Im9 for the endonuclease toxin colicin E9 defined by homologue-scanning mutagenesis. *J. Biol. Chem.* **272**, 22253-22258.
34. Wallis, R., Leung, K.-Y., Osborne, M.J., James, R., Moore, G.R. & Kleanthous, C. (1998). Specificity in protein–protein interaction: conserved Im9 residues are the major determinants of stability in the colicin E9 DNase-Im9 complex. *Biochemistry* **37**, 476-485.
35. Schaller, K. & Nomura, M. (1976). Colicin E2 is a DNA endonuclease. *Proc. Natl Acad. Sci. USA* **68**, 3989-3993.
36. Garinot-Schneider, C., Pommer, A.J., Moore, G.R., Kleanthous, C. & James, R. (1996). Identification of putative active-site residues in the DNase domain of colicin E9 by random mutagenesis. *J. Mol. Biol.* **260**, 731-742.
37. Christianson, D.W. & Lipscomb, W.N. (1989). Carboxypeptidase A. *Acc. Chem. Res.* **22**, 62-69.
38. Matthews, B.W., Jansonius, J.N., Colman, P.M., Schoenborn, B.P. & Dupourque, D. (1972). Three-dimensional structure of thermolysin. *Nature New Biol.* **238**, 37-41.
39. Volbeda, A., Lahm, A., Sakiyama, F. & Suck, D. (1991). Crystal structure of *Penicillium citrinum* P1 nuclease at 2.8 Å resolution. *EMBO J.* **10**, 1607-1618.
40. Matthews, B.W. (1968). Solvent content of protein crystal. *J. Mol. Biol.* **33**, 491-497.
41. Otwinowski, Z. (1993). Oscillation data reduction program. In *Proceedings of the CCP4 study weekend: data collection and processing.* (Sawyer, L., Isaacs, N. & Bailey, S., eds.), pp. 56-62, SERC Daresbury Laboratory, Warrington, UK.
42. Brünger, A.T. (1992). *X-PLOR, Version 3.1: A System for X-ray Crystallography and NMR.* Yale University Press, New Haven, CT.
43. Cowtan, K. (1994). DM: An automated procedure for phase improvement by density modification. *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography* **31**, 34-38.
44. Sano, Y., Matsui, H., Kobayashi, M. & Kageyama, M. (1993). Molecular structures and functions of Pyocins S1 and S2 in Pseudomonas aeruginosa. *J. Bacteriol.* **175**, 2907-2916.
45. Sano, Y. & Kageyama, M. (1993). A novel transposon-like structure carries the genes for pyocin AP41, a *Pseudomonas aeruginosa* bacteriocin with a DNase domain homology to E2 group colicins. *Mol. Gen. Genet.* **237**, 161-170.
46. Laskowski, R.A., MacArthur, M.W., Moss, D.S. & Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283-291.
47. Barton, G.L. (1993). ALSCRIPT: a tool to format multiple sequence alignments. *Protein Eng.* **6**, 37-40.
48. Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Cryst.* **24**, 946-950.
49. Merrit, E.A. & Murphy, M.E.P. (1994). Raster3D Version 2.0. A program for photorealistic molecular graphics. *Acta Cryst. D* **50**, 869-873.
50. Nicholls, A., Sharp, K.A. & Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* **11**, 281-296.