1.6-Å Crystal Structure of EntA-im

A BACTERIAL IMMUNITY PROTEIN CONFERRING IMMUNITY TO THE ANTIMICROBIAL ACTIVITY OF THE PEDIOCIN-LIKE BACTERIOCIN ENTEROCIN A*

Received for publication, February 7, 2005, and in revised form, March 4, 2005
Published, JBC Papers in Press, March 7, 2005, DOI 10.1074/jbc.M501386200

Line Johnsen‡§, Bjørn Dalhus¶, Ingar Leiros**, and Jon Nissen-Meyer‡

From the ‡Program for Biochemistry and Molecular Biology, Department of Molecular Biosciences, University of Oslo, 0316 Oslo, Norway, §Institute of Medical Microbiology, Section for Molecular Microbiology, National University Hospital, N-0027 Oslo, Norway, and **Macromolecular Crystallography Group, European Synchrotron Radiation Facility, BP220, F-38043 Grenoble Cedex 9, France

Many Gram-positive bacteria produce ribosomally synthesized antimicrobial peptides, often termed bacteriocins. Genes encoding pediocin-like bacteriocins are generally cotranscribed with or in close vicinity to a gene encoding a cognate immunity protein that protects the bacteriocin-producer from their own bacteriocin. We present the first crystal structure of a pediocin-like immunity protein, EntA-im, conferring immunity to the bacteriocin enterocin A. Determination of the structure of this 103-amino acid protein revealed that it folds into an antiparallel four-helix bundle with a flexible C-terminal part. The fact that the immunity protein conferring immunity to carnobacteriocin B2 also consists of a four-helix bundle (Sprules, T., Kawulka, K. E., and Vedera, J. C. (2004) Biochemistry 43, 11740–11749) strongly indicates that this is a conserved structural motif in all pediocin-like immunity proteins. The C-terminal half of the immunity protein contains a region that recognizes the C-terminal half of the cognate bacteriocin, and the flexibility in the C-terminal end of the immunity protein might thus be an important characteristic that enables the immunity protein to interact with its cognate bacteriocin. By homology modeling of three other pediocin-like immunity proteins and calculation of the surface charge distribution for EntA-im and the three structure models, different charge distributions were observed. The differences in the latter part of helix 3, the beginning of helix 4, and the loop connecting these helices might also be of importance in determining the specificity.

The pediocin-like bacteriocins constitute an important and well-studied class of ribosomally synthesized antimicrobial peptides produced by lactic acid bacteria. At least 20 different pediocin-like bacteriocins have been characterized (1–3). They contain between 37 and 48 residues (1–3), and they are all cationic, display antilisteria activity, and kill target cells by permeabilizing the cell membrane (4, 5). Moreover, they have very similar primary structures, especially in the N-terminal domain (6). NMR studies indicate that pediocin-like bacteriocins contain two structural units, a cationic N-terminal β-sheet domain and a hydrophobic C-terminal domain (7–9), which in most of these peptides appears to form a hairpin-like structure (7). These two domains are separated by a flexible hinge that enables the two domains to move relative to each other. The N-terminal domain binds to the target cell surface (10, 11), whereas the more hydrophobic C-terminal domain enters the hydrophobic part of the cell membrane (3, 12, 13). Despite their extensive sequence similarity, the pediocin-like bacteriocins differ markedly with respect to the bacteria they kill (6, 14). Recent studies show that the hydrophobic C-terminal domain of pediocin-like bacteriocins is important in determining this target cell specificity (15).

Immunity proteins protect bacteriocin-producing organisms against the antimicrobial activity of their own bacteriocin. The gene encoding an immunity protein is generally cotranscribed with or in close vicinity to the gene encoding the cognate pediocin-like bacteriocin (16–21). The primary structures of most immunity proteins of pediocin-like bacteriocins have been deduced from DNA sequences (22). The proteins display 5–85% sequence similarities, contain between 88 and 115 amino acids, and have been classified into three subgroups (A, B, and C) based on sequence homology (14, 22). The immunity proteins show a high degree of specificity with respect to the bacteriocin they recognize, although some immunity proteins render cells immune to a few pediocin-like bacteriocins in addition to their cognate bacteriocin (22). Using hybrid immunity proteins and hybrid bacteriocins, it was recently shown that the C-terminal half of the immunity proteins contains a region that is involved in specific recognition of the C-terminal hairpin-like domain of the bacteriocin they confer immunity to (15, 23). The mode of action of the immunity proteins is not known, but their effectiveness is strain-dependent, and the functionality of the immunity proteins thus appears to depend partly on strain-dependent factors (22, 23). It has been suggested that they may act by disturbing the interaction between the bacteriocin and a (putative) membrane-located bacteriocin receptor (19, 24, 25).

The immunity proteins for pediocin-like bacteriocins have been shown to be located intracellularly, with only a small proportion (about 1%) possibly being associated with the cell membrane (17, 19). Circular dichroism studies indicate that the immunity proteins do not interact extensively with mem-
The addition of Benzonase Nuclease (20 mg/ml) and Protein Extraction Reagent (20 ml/liter cell culture; Novagen) and 200 ml of minimal medium, and the cells were starved for 4–8 h at 37 °C before induction, the cells were spun down, the pellet was resuspended in 200 µl of minimal medium, and the cells were harvested by centrifugation at 37 °C for 4 h. The cells were harvested by centrifugation, and the supernatant containing the EntA-im gene (22), was used as template in the PCR. EntA-im contains only 1 methionine residue, which is located in the N terminus (start codon). To ensure an anomalous signal in the x-ray diffraction experiment, 2 leucine residues (Leu31 and Leu80), were mutated to methionine. The resulting mutated version of EntA-im is termed EntA-im-(L31M/L80M). The Leu31 and Leu80 residues were selected on the basis of a secondary structure prediction (Predator (27)), which indicated that these 2 residues are located in α-helix regions. The mutations were made by site-directed mutagenesis, and the DNA sequence of the mutated plasmid, pET3a/entA-im-(L31M/L80M), was verified by automated DNA sequencing using a MegaBase DNA analysis system and the DYEnamic reverse transcription dye terminator cycle sequencing kit (GE Healthcare).

**Experimental Procedures**

**Construction of Enterocin A Immunity Protein Expression Plasmid**—PCR fragments containing the EntA-im gene with flanking Ndell and BamHI sites were cloned into the vector pET3a (Novagen), resulting in the plasmid pET3a/entA-im. Specific primers for PCR amplification were synthesized (Eurogentec), and the plasmid pEAM1, which contains the EntA-im gene (22), was used as template in the PCR. EntA-im contains only 1 methionine residue, which is located in the N terminus (start codon). To ensure an anomalous signal in the x-ray diffraction experiment, 2 leucine residues (Leu31 and Leu80), were mutated to methionine. The resulting mutated version of EntA-im is termed EntA-im-(L31M/L80M). The Leu31 and Leu80 residues were selected on the basis of a secondary structure prediction (Predator (27)), which indicated that these 2 residues are located in α-helix regions. The mutations were made by site-directed mutagenesis, and the DNA sequence of the mutated plasmid, pET3a/entA-im-(L31M/L80M), was verified by automated DNA sequencing using a MegaBase DNA analysis system and the DYEnamic reverse transcription dye terminator cycle sequencing kit (GE Healthcare).

**Protein Expression and Purification**—The selenomethionine (Se-Met)-labeled mutant version of EntA-im (Se-Met-EntA-im-(L31M/L80M)) was overexpressed in B834(DE3) pLysS cells transformed with pET3a/entA-im-(L31M/L80M) and subsequently purified to homogeneity. Initially, B834(DE3) pLysS cells transformed with pET3a/entA-im-(L31M/L80M) were grown at 37 °C to an A600 of 1.0 in 200 ml of minimal medium containing methionine (50 mg/liter). Before induction, the cells were spun down, the pellet was resuspended in 200 ml of minimal medium, and the cells were starved for 4–8 h at 37 °C before the addition of seleno-L-methionine (50 mg/liter). The cells were then incubated for 30 min before isoelectric focusing in a pH 4.0–5.6 ampholine range was added to a final concentration of 0.4 M, and the induction was conducted at 37 °C for 4 h. The cells were harvested by centrifugation (5,000 × g for 10 min) and frozen before resuspension in Bugbuster Protein Extraction Reagent (20 ml/liter cell culture; Novagen) and addition of Benzonase Nuclease (20 µl/liter cell culture; Novagen). Clarified cell extract was obtained after incubation for 20 min at room temperature and centrifugation (13,000 rpm for 20 min). The cell extract was dialyzed in 50 ml of phosphate buffer (pH 6) before it was applied on a SP Sepharose Fast Flow cation-exchange column (GE Healthcare). Excess protein was removed by washing with phosphate buffer before Se-Met-EntA-im-(L31M/L80M) was eluted with 0.4 M NaCl. The 0.4 M NaCl fraction from the cation exchanger was diluted in phosphate buffer and then applied at a flow rate of 1 ml/min on a low-pressure Resource Mono S column (GE Healthcare). Se-Met-EntA-im-(L31M/L80M) binds to the column and was detected as the major peak on the optical density profile upon elution with a 0–0.25 M linear NaCl gradient. The protein solution was dialyzed against water overnight before being concentrated by ultrafiltration using Centricon concentrators (molecular weight cutoff, 3,000; Millipore). The primary structure was confirmed by mass spectrometry on a Voyager-DE RP matrix-assisted laser desorption time-of-flight mass spectrometer (PerSeptive Biosystems), with 3.5-dimethoxy-4-hydroxy-cinnamic acid used as the matrix. The experimental procedure for purification of native EntA-im has been described previously (28).

**Crystallization and Data Collection**—Crystals of Se-Met-EntA-im-(L31M/L80M) were obtained by vapor diffusion using the hanging drop method in combination with microseeding. The reservoir solution contained 15–30% (w/v) polyethylene glycol 3350 and 200 mM sodium tartrate. Two µl of protein solution (typically 3–5 mg/ml protein) were mixed with the same volume of precipitant solution. The drops were allowed to equilibrate for about 24 h before crystal nuclei were transferred from previously obtained thin needle-shaped crystals by use of a cut whisker. Crystals grew within 7–14 days at room temperature. The new crystals were generally thicker than the crystals obtained without seeding. The experimental procedures for crystallization of native EntA-im and subsequent data collection have been described previously (28).

Data for Se-Met-EntA-im-(L31M/L80M) were collected at beamline ID29 at the European Synchrotron Radiation Facility (Grenoble, France) using a ADSC Q210 charge-coupled device detector. Statistics for the data collection are summarized in Table I. All data were indexed and integrated with MOSFLM (29) and thereafter scaled and merged using SCALA (30). The intensities were converted to structure factors using TRUNCATE in CCP4 (30).

**Structure Solution and Refinement of Se-Met-EntA-im-(L31M/L80M)**—Selenium was introduced by the SeMet-EntA-im-(L31M/L80M) model was solved using SOLVE. Data collected at the peak of the selenium K-absorption edge (λ = 0.979 Å) extending to 2.2 Å resolution were scaled but not merged in SCALA (30) and used as input to SOLVE (31) in “SAD mode.” Two selenium positions were located automatically in the asymmetric unit, and the resulting overall figure of merit after SOLVE was 0.36 with an overall Z-score of 9.3. These initial phases were further improved with statistical solvent flattening using RESOLVE (31), which resulted in a much improved electron density map and gave an overall figure of merit of 0.65. Subsequently, ARP/wARP (32) was used to auto-trace the structure, and 80 of the 103 residues in a single protein chain, constituting the asymmetric unit, were correctly placed. Refinement of the Se-Met-EntA-im-(L31M/L80M) model was performed with CNS1.1 (33), and several cycles of refinement interspersed with manual rebuilding gave final Rwork and Rfree values of 21.0% and 26.2%, respectively.

About one-third of the 47 water molecules were located automatically with the water-picking procedure within CNS1.1, whereas the rest of the water molecules were built into the present model by manual inspection of the mFo − DF, and 2mFo − DF, difference Fourier maps in O (34). The final model consists of 79 residues. The statistics from phasing and refinement are shown in Table I.

**Structure Solution and Refinement of Native EntA-im—**Diffraction data for a crystal with native EntA-im were collected at the Swiss-Norwegian Beamline at the European Synchrotron Radiation Facility. Protein crystallization, data collection procedures, and processing and analysis of the data have been presented (28). The refined Se-Met-EntA-im-(L31M/L80M) model was used as a search model in molecular replacement against the data of the native protein, using the program EMIR (35). Two unique solutions were located, giving a model R-factor of ~40. A third molecule was identified and auto-built using ARP/wARP (~250 residues in total were built). Subsequent refinement and water picking in CNS (36) gave an Rwork of 0.224, an Rfree of 0.27, mFo − DF, and 2mFo − DF, difference Fourier maps in O (34), has resulted in the present model containing 264 residues and 361 water molecules as well as two copies of a citrate anion from the crystallization solution. Final Rwork and Rfree values are 21.9% and 24.3%, respectively. The addition of citrate or tartrate proved to be crucial for crystal formation, and in the case of native EntA-im with two citrate anions bonded to the protein, these negatively charged polyanions seem to stabilize the molecule. Two selenium positions were located automatically in the asymmetric unit, and many positively charged residues in these two molecules.

Atomic coordinates and structure factor data for both native EntA-im and Se-Met-EntA-im-(L31M/L80M) have been deposited in the Protein Data Bank with accession numbers 2BL8/2RBLSF and 2BL7/2RBL7SF, respectively.

**Model Quality**—More than 98% of the located residues in both the native EntA-im and Se-Met-EntA-im-(L31M/L80M) have backbone conformation within the most favored regions of the Ramachandran plot (36) and gives a representative view of the quality of the mFo − DF, and 2mFo − DF, map after final refinement for native EntA-im.

**Sequence Alignment and Homology Modeling**—The sequence alignments were performed using ClustalW (37). The sequences were imported to Swiss-Pdb Viewer (38), and after adjustment of the alignment, they were fitted to the EntA-im structure. Side chain conformations were selected on the basis of the known crystal structure, except in cases with obvious steric conflicts. Such problems were resolved by selecting the energetically most favored rotamer using the built-in tools of Swiss-Pdb Viewer.
RESULTS AND DISCUSSION

Crystallization—The native immunity protein, EntA-im, and the Se-Met-labeled mutant version, EntA-im-(L31M/H110/L80M), were crystallized under slightly different conditions, resulting in two different forms. Native EntA-im crystallized in the monoclinic space group C2 in the presence of citrate, whereas Se-Met-EntA-im-(L31M/H110/L80M) crystallized in the triagonal space group P3121 in the presence of tartrate (Table I). In the crystal containing native EntA-im, there are three molecules in the asymmetric unit (molecules A, B, and C; see Fig. 2). The crystallographically independent native EntA-im molecules have an r.m.s.d. of <0.9 Å for C-α superimpositions of corresponding residues between the three monomers, A–C. The crystal with Se-Met-EntA-im-(L31M+L80M), on the other hand, has only a single protein chain in the asymmetric unit. There are no significant differences between the native and mutant EntA-im as far as main chain conformation is concerned.

Overall Structure—The crystal structure of EntA-im reveals that it folds into an antiparallel four-helix bundle with a flexible C-terminal part; the final model of Se-Met-EntA-im-(31M+L80M) consists of residues 3–81 with no electron density for the first 2 residues and the last 22 residues. There is a well-defined electron density for the backbone part of the four helices of the α-bundle and in all three loops connecting them. Conversely, the three crystallographically independent mole-

---

**Table I**

Data collection and refinement statistics for Se-Met-EntA-im-(L31M+L80M) and native EntA-im

|                      | Se-Met-EntA-im-(L31M+L80M) | Native EntA-im |
|----------------------|-----------------------------|----------------|
| **Data collection**  |                             |                |
| Beamline             | ID29 (ESRF)                 | SNBL (ESRF)    |
| Wavelength (Å)       | 0.9792                      | 0.8727         |
| Temperature (K)      | 110                         | 110            |
| Space group          | P3,21                       | C2             |
| Unit cell (Å, °)     | 75.80 75.80 36.59           | 116.3 42.4 66.2 |
| Resolution (Å)       | 90 90 120                   | 90 111.3 90    |
| R/merge              | 50.0 to 2.2                 | 30.0 to 1.60   |
| R (3% test set)      | 0.11 (0.40)                 | 0.07 (0.29)    |
| r.m.s.d. bond distance (Å) | 16.5 (2.2) | 10.5 (4.4)    |
| r.m.s.d. bond angle (°) | 99.8 (98.8) | 95.6 (92.7)   |
| r.m.s.d. dihedrals (°) | 8.6 (4.2)   | 2.9 (2.1)     |
| Completeness (%)     | 95.9                        | NA             |
| Anomalous completeness (%) | 4.2              | NA             |
| Anomalous multiplicity | 3.6                       | NA             |
| **Refinement statistics** |                      |                |
| Resolution range (Å) | 50 to 2.20                  | 30 to 1.60     |
| No. of unique reflections | 6364                    | 38,124         |
| No. of test reflections | 193 (3%)                | 1165 (2.9%)    |
| R (working set)      | 0.210                       | 0.219          |
| R (3% test set)      | 0.262                       | 0.242          |
| r.m.s.d. bond distance (Å) | 0.005                     | 0.005          |
| r.m.s.d. bond angle (°) | 1.00                      | 1.18           |
| r.m.s.d. dihedrals (°) | 16.0                      | 17.5           |
| Total no. of non-H atoms | 723                        | 2618           |
| No. of amino acid residues | 3–81 A: 4–39, 47–103 | B: 4–84 C: 4–81, 93–103 |
| No. of solvent molecules | 47                         | 361            |
| Average overall B-factor (Å²) | 34.8                    | 23.8           |
| CV Luzzati coordinate error | 0.35                    | 0.24           |

**FIG. 1.** A stereoview of the electron density (2mF₀ – F₀ map), with a contour level of 1, after final refinement. Residues Tyr93, Arg94, Ala95, Asp96, and Tyr97 from molecule A and Tyr93 from molecule B are shown. Water is shown as red spheres.

**FIG. 2.** Ribbon diagram of the three independent molecules in the asymmetric unit in native EntA-im (molecule A is red, molecule B is yellow, and molecule C is green).
molecules in the structure of native EntA-im display individual differences with respect to the number of well-defined loop regions (Fig. 2; Table I). The structure of molecule A has a flexible loop between helix 2 and 3, whereas the entire C-terminal part adopts a well-defined random coil conformation (Figs. 2 and 3). The structure of molecule C has a flexible loop between helix 4 and the last 10 residues, which form a terminal coil with the same conformation and relative position as in molecule A (Fig. 2). This last observation suggests that the C-terminal might actually have a preferred conformation. However, the absence of a single main chain conformation of the C-terminal part in the structure of Se-Met-EntA-im-(L31M/L80M) and that of molecule B points to a flexible C terminus. In molecule A, the distinct folding of the C terminus, being located between molecules A and C, could be due to molecular packing in the crystal, and not an intrinsic structural element. The same is true for the well-ordered short C terminus of molecule C, which has 111 residues. The EntA-im and CarB2-im proteins display 13% sequence identity (Fig. 4A). Despite their low sequence similarity, these two proteins have some common three-dimensional structural features. They are both left-turning, four-helix bundles that fold around a hydrophobic core (Fig. 3). CarB2-im has a short α-helical turn between helices 1 and 2, whereas EntA-im has a one-turn α-helix between helices 2 and 3. The most obvious structural difference is that CarB2-im has an additional fifth helix in the C-terminal end, which is either flexible or random coil in EntA-im. Their α-helices also pack together with different interhelical angles (Fig. 3). The C-α superposition of CarB2-im and EntA-im (molecule B with complete α-helix bundle core) gives an r.m.s.d. of 2.1 Å. The fact that two immunity proteins that belong to different subclasses consist of a four-helix bundle strongly indicates that this is a conserved structural motif in many, if not all, pediocin-like immunity proteins.

Comparison of Molecular Surfaces of Subgroup A Immunity Proteins—Calculation of the surface charge distribution in EntA-im (using Swiss-Pdb Viewer (38)) reveals a relatively large positively charged region formed by residues in helices 2 and 3 (Lys33, Lys37, and Arg53) and the loop between the two helices (Lys40 and Lys42) (Fig. 5, a and b). The latter 2 residues are flexible because no single dominant conformation is apparent.
ent in the electron density map. However, with an all-trans extended side chain conformation, the residues will be part of the positively charged region. On the opposite side of the molecule (Fig. 5, c and d), a negatively charged surface region is formed by residues in helix 1 (Glu$^{11}$) and helix 4 (Glu$^{71}$, Glu$^{74}$, and Glu$^{75}$).

Based on the high sequence identity with EntA-im (>40%; Fig. 4B) and the use of homology modeling, structure models were predicted for three other subgroup A immunity proteins: LeuA-im (39), MesY-im (40), and DivI-im (41). Structural models were also predicted for two hybrid immunity proteins (EntA/LeuA-im and LeuA/EntA-im) that have been constructed by combining the N- and C-terminal halves of two subgroup A immunity proteins (EntA-im and LeuA-im) (23). Both of these hybrid immunity proteins were active (23), indicating that the three-dimensional structure of LeuA-im is similar to that of EntA-im. Because the C-terminal random coil could be flexible in solution, only the four-helix bundle core was used as a

**FIG. 5.** Surface charge distribution of EntA-im and models of LeuA-im, DivI-im, MesY-im, EntA/LeuA-im, and LeuA/EntA-im. The C-terminal is omitted from residue 82 for EntA-im and for all the models. a, view of helices 1 and 2. b, view of helices 2 and 3. c, view of helices 3 and 4. d, view of helices 1 and 4. Blue indicates positive charge, and red indicates negative charge.
template. The molecular surface charge distribution was calculated for all predicted structure models (Fig. 5, a–d).

The positive region observed in EntA-im is less prominent in LeuA-im and DivI-im (Fig. 5, a and b), although Lys\textsuperscript{37} and Arg\textsuperscript{53} are conserved in all four proteins. The negative region observed in EntA-im (Fig. 5d) is more pronounced in LeuA-im and DivI-im, whereas it is positive in MesY-im. This is in good agreement with the calculated pI values of 8.7 for EntA-im, 7.8 for LeuA-im, 5.7 for DivI-im, and 8.6 for MesY-im.

Recent studies have shown that the C-terminal part of the immunity proteins for the pediococin-like bacteriocins contains a domain that is involved in the specific recognition of the bacteriocins they confer immunity to (23). Whereas EntA-im and LeuA-im have different specificity (i.e. they recognize and confer immunity to different bacteriocins), EntA-im and the hybrid LeuA/EntA-im have similar specificity. Likewise, the LeuA-im and hybrid EntA/LeuA-im proteins also have similar specificity. Most of the charge differences discussed above are not located in the C-terminal part of these proteins and are consequently not likely to be involved in the specific recognition of the cognate bacteriocin. Certain differences were, however, noticed in the C-terminal part, and these might be of importance for specificity. A positive patch due to Lys\textsuperscript{73} is present in LeuA-im and EntA-im, whereas a negative patch due to Glu\textsuperscript{71} is present in EntA-im and LeuA/EntA-im (Fig. 5c). Another difference is the positive patch due to Arg\textsuperscript{29} in LeuA-im and EntA-im, which is absent in EntA-im and LeuA/EntA-im (Fig. 5c). It is also worth noting that both EntA-im and LeuA/EntA-im have a protruding aromatic residue (Tyr) in position 63 that is not found in LeuA-im or EntA/LeuA-im (Fig. 5, b and d).

**CONCLUSIONS**

The 1.6-Å x-ray structure of EntA-im reveals that it is a globular protein consisting of an antiparallel four-helix bundle with a flexible C-terminal part. The four-helix bundle appears to be a conserved structural motif in the pediococin-like immunity proteins (at least in subgroups A and C) (26). It is clear that the C-terminal part of these immunity proteins (directly or indirectly) specifically interacts with their cognate bacteriocins and thereby confers immunity. In this study we determined the structure of EntA-im and thereby gained insight into structural features of the C-terminal half that might be important for specificity. Interestingly, our results indicate that approximately the last 20 residues of the C-terminal part seem to form a flexible structural element that might be involved in specific interactions with bacteriocins or other cellular components that may interact with bacteriocins. Another possibility is that the specificity is determined by the different charged patches found in the latter part of helix 3, helix 4, and the loop combining these two helices. These results on the structure of the pediococin-like immunity proteins will make it easier to design new hybrid immunity proteins as well as immunity protein mutants that will clarify which structured features and/or specific residues are determinants for recognition of cognate bacteriocins.

**Acknowledgments**—We thank the scientific staff at beamline ID29 (European Synchrotron Radiation Facility) and the Swiss-Norwegian Beamline.