Structural Basis for Action of the External Chaperone for a Propeptide-deficient Serine Protease from Aeromonas sobria*

Received for publication, October 31, 2014, and in revised form, March 16, 2015. Published, JBC Papers in Press, March 16, 2015, DOI 10.1074/jbc.M114.622852

Hidetomo Kobayashi†1, Toru Yoshida†1, Takuya Miyakawa†1, Mitsuru Tashiro‡, Keinosuke Okamoto**, Hiroyasu Yamanaka‡, Masaru Tanokura‡, and Hideaki Tsuge§2

From the 1Laboratory of Molecular Microbiological Science, Faculty of Pharmaceutical Sciences, Hiroshima International University, Kure, Hiroshima 737-0112, Japan, the 2Department of Bioresource and Environmental Sciences, Faculty of Life Sciences, Kyoto Sangyo University, Kamigamo-Motoyama, Kyoto 603-8555, Japan, the 3Department of Applied Biological Chemistry, Graduate School of Agricultural and Life Sciences, University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan, the 4Department of Chemistry, College of Science and Technology, Meisei University, Hodokubo, Hino, Tokyo 191-0042, Japan, and the 5Department of Pharmacogenetics, Graduate School of Medicine, Dentistry and Pharmaceutical Sciences, Okayama University, Okayama 700-8530, Japan

Background: ASP is propeptide-deficient subtilase; thus the maturation mechanism is unclear.

Results: We found ORF2 as the external chaperone and revealed ORF2 and ORF2-ASP complex structures.

Conclusion: The N- and C-terminal regions of ORF2 are requisite for its chaperone activity.

Significance: ASP and its homolog form a novel family of subtilases having an external chaperone.

Subtilisin-like proteases are broadly expressed in organisms ranging from bacteria to mammals. During maturation of these enzymes, N-terminal propeptides function as intramolecular chaperones, assisting the folding of their catalytic domains. However, we have identified an exceptional case, the serine protease from Aeromonas sobria (ASP), that lacks a propeptide. Instead, ORF2, a protein encoded just downstream of asp, appears essential for proper ASP folding. The mechanism by which ORF2 functions remains an open question, because it shares no sequence homology with any known intramolecular propeptide or other protein. Here we report the crystal structure of the ORF2-ASP complex and the solution structure of free ORF2. ORF2 consists of three regions: an N-terminal extension, a central body, and a C-terminal tail. Together, the structure of the central body and the C-terminal tail is similar to that of the intramolecular propeptide. The N-terminal extension, which is not seen in other subtilisin-like enzymes, is intrinsically disordered but forms some degree of secondary structure upon binding ASP. We also show that C-terminal (ΔC1 and ΔC5) or N-terminal (ΔN43 and ΔN64) deletion eliminates the ability of ORF2 to function as a chaperone. Characterization of the maturation of ASP with ORF2 showed that folding occurs in the periplasmic space and is followed by translocation into extracellular space and dissociation from ORF2, generating active ASP. Finally, a PSI-BLAST search revealed that operons encoding subtilases and their external chaperones are widely distributed among Gram-negative bacteria, suggesting that ASP and its homologs form a novel family of subtilases having an external chaperone.

Molecular chaperones such as GroEL/GroES and DnaK/DnaJ/GrpE assist with the folding and unfolding of proteins, as well as with their assembly and disassembly (1), in part by preventing newly synthesized polypeptide chains and assembled subunits from collecting into nonfunctional aggregates. In addition, some proteins contain a chaperone domain that lower the energy barrier between the native and partially folded states of the main protein and is thus essential for the proper folding of that protein, which would not fold spontaneously (2). The best described examples of steric chaperones are the propeptide regions of the bacterial proteases subtilisin and α-lytic protease (2–6). These enzymes are expressed with an N-terminal propeptide often referred to as an “intramolecular chaperone” because it is essential for the proper folding of the tertiary structure of the catalytic domain. Once the catalytic domain is properly folded, the propeptide is cleaved, generating the active protease, after which the free propeptide is degraded. Furthermore, with the exception of protease, lipase-specific foldase has been known to play as a separate protein chaperone (7).

Aeromonas species are Gram-negative facultative anaerobic bacteria found ubiquitously in aquatic environments (8). The main syndrome caused by infection with Aeromonas is gastroenteritis, although in severe cases sepsis may occur (9–11). Aeromonas species secrete a number of putative virulence factors, including hemolysins, enterotoxins, and proteases (12–15). We purified a 65-kDa serine protease from the culture supernatant of A. sobria (ASP)3 and found that the enzyme induced vascular leakage, reduced blood pressure through activation of the kallikrein/kinin system (16), promoted human
plasma coagulation through activation of prothrombin (17), and caused the formation of pus and edema through the action of anaphylatoxin C5a (18). This suggests ASP is a potent virulence factor in diseases caused by Aeromonas infection.

ASP is a member of subtilisin-like serine proteases (subtilase) superfamily and belongs to the Kexin subfamily, which includes mammalian proprotein convertases such as furin (19). The structure of Kexin family proteins includes a signal peptide, a partially conserved propeptide pivotal for maturation of the enzyme, a highly conserved serine protease domain, a partially conserved region called the P-domain, a transmembrane domain, and a cytoplasmic domain. Like mammalian Kexin proteins, ASP has a P-domain, which is in contrast to other bacterial subtilisins, and at present ASP is thought to be the only bacterial member of the Kexin family (20). Moreover, ASP lacks a propeptide. As far as we know, ASP is one of only two propeptide-deficient subtilases, the other being the intracellular serine protease from Bacillus subtilis (21). We recently determined the crystal structure of active ASP (19) and suggested that a protein encoded just downstream of ORF2 (ORF2, ~15 kDa) functions as a chaperone for ASP. This was based on the finding that ASP was not detected in either the periplasmic or extracellular space unless it was co-expressed with intact ORF2 (14). However, ORF2 shows no sequence homology with the propeptide from any known subtilase or with any other protein, so its precise function could not be predicted.

Here we report the crystal structure of the ORF2-ASP complex at 1.4 Å resolution. Surprisingly, the structure was similar to that of the complex formed between subtilase with its cognate propeptide, despite the lack sequence homology between ORF2 and the propeptide. We further report the solution structure of free ORF2 obtained using NMR, which, unlike the subtilase propeptide, has tertiary structure. ORF2 contains an intrinsically disordered N-terminal extension that develops some secondary structure upon ASP binding and that is not seen in the subtilisin propeptide. Biochemical analysis of ORF2 revealed that the C-terminal tail, which inhibits the active site, and the N-terminal extension are both critical for the folding of ASP. We also show the pH-dependent dissociation of the ORF2-ASP complex after translocation of the protein from the periplasmic to the extracellular space. Finally, we show that operons encoding a subtilase (ASP homolog) and its chaperone (ORF2 homolog) are broadly distributed among Gram-negative bacterial species, prompting us to conclude that ASP is a member of a novel family within subtilase.

**EXPERIMENTAL PROCEDURES**

**Protein Expression and Purification**—ASP was overexpressed and purified as described previously (19). ORF2 (WT, ΔC1, ΔC5, ΔN11, ΔN43, or ΔN63) was overexpressed using pET11 vector and Escherichia coli BL21(DE3) cells and purified using nickel affinity and gel filtration chromatography. For co-expression of ORF2 and ASP, genes encoding ASP (S336A) and His6-tagged ORF2 were cloned into pSA19 plasmid (14). The proteins were then co-expressed in E. coli HB101 cells grown in LB broth and purified using nickel-Sepharose Fast Flow and Sephacryl S-200 HR gel filtration columns (GE Healthcare) with 0.5 M NaCl. We initially found that the ORF2-ASP complex was stable only under high NaCl conditions (see Fig. 1A). However, to keep the NaCl concentration as low as possible for crystallization, we determined that the complex was also stable in low NaCl if PIPES buffer was used. We therefore dialyzed the purified sample containing the complex against 10 mM PIPES (pH 7.5) containing 50 mM NaCl before crystallization.

**Measurement of ASP Activity**—To measure the enzymatic activity of ASP, we used a fluorogenic synthetic peptide substrate, Boc-Glu-Lys-Lys-MCA (Peptide Institute), as previously described (22). One unit was defined as the production of 1 pmol of 7-amino-4-methyl-coumarin (AMC)/min/μl.

**Analysis of the Inhibitory Action of ORF2 against ASP**—To determine whether ORF2 inhibits ASP activity according to the formation of the ORF2-ASP complex, we carried out the following enzymological analysis (23). To determine the IC50 for ASP inhibition, the activity of ASP (20 nM) in the presence of various amounts of purified ORF2 (0–200 nM) was measured using a synthetic substrate, Boc-Glu-Lys-Lys-MCA (66 μM), as described above. The IC50 was defined by calculating the ORF2 concentration at which 50% of ASP activity was blocked. To then investigate the inhibition mechanism, ASP activity was measured using 2 nM active ASP and Boc-Glu-Lys-Lys-MCA over a concentration range spanning 8.3–266.7 μM. In addition, the ORF2 concentration was varied from 1 to 6 nM. Least squares nonlinear regression was used to determine the mode of inhibition and calculate Ki. The data were fitted to the mixed inhibitor equation \( v = \frac{V_{max}([1 + [I]/(αK_{I})]S)}{K_m([1 + [I]/K_{I}]/(1 + [I]/(α'K_{I})) + [S])} \) using GraphPad Prism 6 software. The α value reports on the mechanism of inhibition; large α values are indicative of competitive inhibition, whereas a value of 1 suggests purely noncompetitive inhibition. The value for ORF2 was found to be 3.1. All experiments were performed in triplicate.

**Immunological Detection of the ORF2-ASP Complex**—Purified ORF2 and ASP were mixed together and incubated in 20 mM phosphate buffer (pH 7.5), with 0, 0.1, and 0.5 M NaCl, after which the mixture was diluted 10-fold with 20 mM phosphate buffer (pH 7.5) containing 0.5 M NaCl, 5 mM imidazole, and 0.5% Tween 20. Samples, including a negative control, were then added to the wells of a Ni-NTA HisSorb microtiter plate (Qiagen) and allowed to adsorb onto the surface, after which the wells were washed twice with 20 mM PBS containing 0.05% Tween 20. ASP-bound ORF2 was then monitored using an enzyme-linked immunosorbert assay with anti-ASP IgG. To know whether the effect of NaCl in the formation of the ORF2-ASP complex is specific, we further performed similar experiments using KCl (0.1 and 0.5 M) or NH4Cl (0.1 and 0.5 M) instead of NaCl.

**Crystallization and Determination of the Structure of the ORF2-ASP (S336A) Complex**—Crystals of the purified ORF2-ASP (S336A) complex were grown at 4 °C using the hanging drop vapor diffusion method. The protein complex (12 mg/ml) was then mixed with an equal volume of reservoir solution containing 0.1 M HEPES (pH 7.5), 0.1 M ammonium sulfate, and 15% (w/v) PEG 3000. The diffraction data set was collected at 100 K using a Rigaku R-AXIS VII with a Rigaku rotating anode x-ray generator on Beamline PF-AR NW12A at the Photon Factory. The structure of the ORF2-ASP (S336A) complex was determined using a new survey method.
solved through molecular replacement using the program PHASER (24) with the ASP structure (Protein Data Bank code 2OXA). Iterative model building and refinement were subsequently accomplished using COOT (25) and Refmac5 in the CCP4 software suite (26). The atomic coordinates and structure factor of the ORF2-ASP complex have been deposited in the Protein Data Bank under accession code 3WQB.

Determination of the ORF2 Solution Structure Using NMR—All NMR spectra for structure determination were collected using 0.9 mM 13C/15N-labeled ORF2, 18 mM MOPS (pH 7.0), and 10% (v/v) D2O at 25 °C on an Agilent (Varian) Unity INOVA 600 NMR spectrometer equipped with a cryogenic probe (Agilent). Signal assignments were mainly accomplished using triple resonance and HCCH total correlation spectroscopy experiments, and distance restraints were obtained from three-dimensional 15N- and 13C-edited NOESY spectra. Structure calculations were done using the CYANA software package version 2.1 (27). The NMR structure has been deposited in the Protein Data Bank under accession code 2MK4.

In Vitro Translation—In vitro translation was carried out according as described previously (28) using an in vitro translation system kit (5 PRIME). Briefly, the pIVEX derivative plasmid including the asp gene (62.5 ng) was added into the in vitro translation reaction mixture. ASP molecules were expressed from the plasmid encoding the asp gene by incubation of the reaction mixture at 30 °C for 9 h in the presence of 0.1 nmol of purified ORF2 protein (or its mutant protein). After the incubation, a portion of the reaction mixture was further incubated with trypsin (Sigma-Aldrich, 2.5 mg/ml) to examine the stability of the ASP molecules produced in vitro against the trypsin digestion. Using the rest of the reaction mixture, we measured the ASP activity of the reaction mixture as described above.

ITC Measurements—ITC experiments were performed at 25 °C on an iTC200 microcalorimeter (GE Healthcare). ASP (S336A) and ORF2 (WT, ΔC1, ΔC5, ΔN11, ΔN43, or ΔN63) were dialyzed against 20 mM sodium phosphate buffer (pH 7.4) containing 0.5 M NaCl. The sample cell was filled with an ASP (S336A) solution (204 μl) at the concentration of 12 μM (ΔN63), 19 μM (ΔN11), 20 μM (WT or ΔN43), or 50 μM (ΔC1 or ΔC5). For each titration, 37 consecutive 1.0-μl aliquots of a WT (0.16 mM), ΔC1 (0.40 mM), ΔC5 (0.80 mM), ΔN11 (0.20 mM), ΔN43 (0.18 mM), or ΔN63 (0.20 mM) solution were injected at 120-s intervals. The first injection volume was 0.2 μl, and the observed thermal peak was excluded in data analyses. Individual ITC data were subtracted with the reference value obtained by the titration of the buffer solution with the ORF2 solutions. Data analyses were performed using the Origin-ITC analysis package (GE Healthcare) in the “one set of sites” mode.

RESULTS

ORF2 Binds to ASP and Inhibits Its Activity under High Salt Conditions at pH 7.5—We previously showed that the interaction between ORF2 and ASP is requisite for formation of the active structure of ASP (29). We first examined the effect of NaCl concentration on the formation of the ORF2-ASP complex. To detect the interaction of ORF2 with ASP, independently purified ASP and Histagged ORF2 were mixed in buffer containing 0, 0.1 or 0.5 M NaCl, after which the mixture was added to a Ni-NTA HisSorb plate, where ASP complexed with ORF2 would be immobilized and could be immunologically detected using an anti-ASP antibody. As shown in Fig. 1A, at pH 7.5 the ORF2-ASP complex was detected in samples containing 0.1 or 0.5 M NaCl (high salt conditions), but little or no complex was detected in the absence of NaCl. Thus ORF2 forms a stable complex with ASP under high NaCl conditions at pH 7.5 in vitro.

When we then measured the enzymatic activity of ASP in the samples at pH 7.5, we found little activity in samples containing 0.1 or 0.5 M NaCl but strong activity in samples lacking NaCl (Fig. 1B). Moreover, in the absence of ORF2, the addition of 0.1 or 0.5 M NaCl had little effect on ASP activity (see Fig. 1C, lower panel). This suggests the reduced activity seen in samples containing 0.1 or 0.5 M NaCl was related to the formation of the ORF2-ASP complex. To verify that possibility, we performed an enzymological analysis in the presence of 0.5 M NaCl using Boc-Glu-Lys-Lys-MCA as the substrate (23). Under those conditions, the IC50 of ORF2 was 46.09 (± 1.37) nM (Fig. 1C). A more detailed kinetic analysis revealed that the ORF2 inhibition data were poorly fit by a simple competitive inhibition model but were much better fit by a mixed model with a significant non-competitive contribution (Fig. 1D). The calculated Kᵢ for ORF2 was 2.01 (± 0.89) μM. Taken together, these findings indicate that in the presence of at least 0.1 M NaCl at pH 7.5, ORF2 binds.
to ASP, forming a stable complex and inhibiting ASP enzymatic activity.

**Crystal Structure of the ORF2-ASP Complex**—To gain further insight into the function of ORF2, we carried out the purification and crystallization of the ORF2-ASP (S336A) complex. To successfully purify the complex, we reasoned that although the complex was maintained under high NaCl conditions, the presence of an ORF2-free ASP fraction would reduce recovery of the purified ORF2-ASP complex because of ASP proteolytic activity. For these studies, therefore, we used the ASP S336A mutant, which does not have proteolytic activity. In addition, although the interaction of ORF2 with ASP is thought to occur in the periplasmic space after translocation across the inner membrane in *A. sobria* (29), the prompt secretion of the two proteins into the extracellular space made it difficult to collect the ORF2-ASP complex from the periplasmic space (19). We therefore co-expressed ORF2 and ASP in *E. coli* and purified the complex from the periplasmic preparation using nickel-Sepharose column chromatography and Sephacryl S-200 HR gel filtration chromatography under 0.5M NaCl conditions (Fig. 2, A and B). Finally, we used high resolution analytical gel filtration chromatography to determine whether the purified proteins formed a complex (Fig. 2C). Because ORF2 has only a small number of amino acids that absorb ultraviolet light (one phenylalanine and two tyrosines of 130 residues), the absorbance at 280 nm was very weak when ORF2 was eluted alone. In addition, the retention time of the purified sample was shorter than that for independently purified ASP, which is consistent formation of an ORF2-ASP complex within the sample.

After crystallization of the purified sample, we were able to reveal the structure of the ORF2-ASP complex (Fig. 3A and Table 1). We found that ASP consists of a catalytic domain, a P-domain, and six “extra regions” that are not seen in subtilisin or Kexin (19). ORF2 consists of an N-terminal extension, a central body, and a C-terminal tail. The N-terminal extension is not seen in the propeptides of subtilisin. The structure of ASP within the complex differed little from that of active free ASP (Protein Data Bank code 2OXA), as indicated by the root mean square deviation of 0.21 Å. However, the surface of the ASP catalytic triad (Asp78, His115, and Ser336) was blocked by the C-terminal tail of ORF2, which bound within a deep cleft in ASP. The C-terminal tail bound to ASP was clearly observable in the F o - F c omit map (Fig. 3B) and formed numerous hydrogen bonds with the ASP catalytic domain. This would account for the ORF2-mediated inhibition of ASP activity (Fig. 1, C and D). On the other hand, there was no detectable electron density for two sections of the ORF2 N-terminal extension, extending from Gln1 to Tyr11 and from Pro24 to Gln43, indicating that those regions were disordered. The N-terminal extension formed only two hydrogen bonds with ASP (Fig. 3C), and the
loop region (Gln53–Glu59) bound to the hydrophobic region (Ala225, Gly226, Tyr228, Leu515, and Gln519) in the additional domain of ASP formed from three of its extra regions. ORF2 did not interact with the P-domain. A portion of the N-terminal extension had secondary structure, four anti-parallel β-strands, which together with a region of the C-terminal tail extending from Glu121 to Glu126 formed a β-sheet (further details about interaction of ORF2 and ASP are shown in Fig. 4). Numerous water molecules were present in the space between ORF2 and ASP, and ASP formed 24 hydrogen bonds with ORF2, including 14 with the C-terminal tail, 8 with the central body, and 2 with the N-terminal extension. In addition, the ORF2 central body, in particular the side chains of Leu97, Ile93, and Leu95 on its surface, form hydrophobic interactions with the surface of ASP, mediating a tight interaction between the two proteins.

The structure of the ORF2-ASP complex exhibits several similarities to that of subtilisin E in complex with its cognate propeptide (Fig. 5A): 1) the C terminus of the propeptide binds to the cleft of subtilisin E, blocking the surface of its catalytic triad (Asp32, His64, and Ser221); 2) despite an absence of sequence homology, the root mean square deviation between the structure of the ORF2 central body and C-terminal tail and

FIGURE 3. Crystal structure of the ORF2-ASP complex. A, top panel, overall structure. ORF2 is depicted as a cartoon in which the domains are color-coded, and ASP is depicted as a surface on which the domains are color-coded. The ASP catalytic triad is shown in cyan. Bottom panel, ORF2 and ASP domain structures. Dashed lines indicate the disordered regions of ORF2. Also shown are the domain structures of Kexin and subtilisin for comparison to ASP. Among the six extra regions of ASP, which are not seen in Kexin, three regions make up an additional domain shown in deep blue; the other three regions are shown in light blue. B, interaction of the ORF2 C-terminal tail with ASP. The Fo−Fc omit electron density map for ORF2 is contoured at 1.5 σ (magenta). Dashed lines indicate hydrogen bonds. C, interaction of the ORF2 N-terminal extension with ASP. Black dashed lines show the hydrogen bonds within ORF2, and red dashed lines show the hydrogen bonds between ORF2 and ASP. Only the main chain of ORF2 is shown as a stick model, with the exception of the side chains of Gln53 and Glu126. All residues and labels are color-coded according to their domain color.
the structure of the subtilisin E propeptide is small (2.5 Å) (Fig. 5B); and 3) ORF2 and propeptide bind in the same direction to ASP and subtilisin E, respectively. These structural similarities suggest that the function of ORF2 is similar to that of the subtilisin E propeptide. On the other hand, the subtilisin E propeptide contains no region analogous to the N-terminal extension of ORF2.

In sum, the structure of ASP within the ORF2-ASP complex purified from the periplasmic space of *E. coli* differed little from that of active free ASP. Within the complex, however, ASP activity was inhibited because of blockade of its catalytic triad by the ORF2 C-terminal tail. These results indicate that the crystal structure of the ORF2-ASP complex represents the state of the molecules after the folding of ASP is complete (postfolding state) and that the folding of ASP is completed within the periplasmic space. In addition, it appears the N-terminal extension peculiar to ORF2 binds to the additional domain peculiar to ASP.

Solution Structure of ORF2 in an ASP-free State—The structure of ASP-free ORF2 was solved using NMR based on the assigned signals of the ORF2 (data not shown) and NOE distance constraints (Table 2). The structure showed that ORF2 contains some of its tertiary structure when free in solution (Fig. 6). This is noteworthy because when the propeptide region

### TABLE 1

| X-ray data collection and refinement statistics |
|-----------------------------------------------|
| The statistics in the highest resolution shell are given in parentheses. |

| Data collection       | P63             |
|-----------------------|-----------------|
| Space group           | P63             |
| Cell dimensions       | a, b, c (Å)     |
| α, β, γ (°)           | 90.0, 90.0, 120.0 |
| Wavelength (Å)        | 1.0000          |
| Rmean (%)             | 9.7 (46.6)      |
| Rfree (%)             | 3.4 (25.8)      |
| CC1⁄2                 | 0.996 (0.826)   |
| I/m (%)               | 32.8 (4.5)      |
| Completeness (%)      | 98.9 (96.3)     |
| Redundancy            | 7.0 (3.0)       |

| Refinement            |                |
|-----------------------|-----------------|
| Resolution (Å)        | 50.0-1.41 (1.46-1.41) |
| No. unique reflections| 130,244 (12,739) |
| Rwork (%)             | 18.7 (19.3)     |
| Rfree (%)             | 20.7 (23.7)     |
| No. waters/calciums   | 392/3           |
| B factors (Å²)        | ASP 14.6        |
|                       | ORF2 27.3       |
|                       | Water 21.5      |
|                       | Calcium 13.7    |
| Root mean square deviations |
| Bond lengths (Å)      | 0.007           |
| Bond angles (°)       | 1.15            |
| Ramachandran plot     |
| Most favored regions (%) | 97.5            |
| Additional allowed regions (%) | 2.5            |
| Disallowed regions (%) | 0.0             |

* Correlation coefficient between intensities from random half data sets.

---

**TABLE 2**

| X-ray data collection and refinement statistics |
|-----------------------------------------------|
| The statistics in the highest resolution shell are given in parentheses. |

| Data collection       | P63             |
|-----------------------|-----------------|
| Space group           | P63             |
| Cell dimensions       | a, b, c (Å)     |
| α, β, γ (°)           | 90.0, 90.0, 120.0 |
| Wavelength (Å)        | 1.0000          |
| Rmean (%)             | 9.7 (46.6)      |
| Rfree (%)             | 3.4 (25.8)      |
| CC1⁄2                 | 0.996 (0.826)   |
| I/m (%)               | 32.8 (4.5)      |
| Completeness (%)      | 98.9 (96.3)     |
| Redundancy            | 7.0 (3.0)       |

| Refinement            |                |
|-----------------------|-----------------|
| Resolution (Å)        | 50.0-1.41 (1.46-1.41) |
| No. unique reflections| 130,244 (12,739) |
| Rwork (%)             | 18.7 (19.3)     |
| Rfree (%)             | 20.7 (23.7)     |
| No. waters/calciums   | 392/3           |
| B factors (Å²)        | ASP 14.6        |
|                       | ORF2 27.3       |
|                       | Water 21.5      |
|                       | Calcium 13.7    |
| Root mean square deviations |
| Bond lengths (Å)      | 0.007           |
| Bond angles (°)       | 1.15            |
| Ramachandran plot     |
| Most favored regions (%) | 97.5            |
| Additional allowed regions (%) | 2.5            |
| Disallowed regions (%) | 0.0             |

---

**FIGURE 4. Interaction of ORF2 and ASP within the crystal structure of the ORF2-ASP complex.** A, ORF2 is depicted as a ribbon model colored in a rainbow spectrum from blue (N-terminal) to red (C-terminal). ASP and the waters located at the interface between ORF2 and ASP are depicted as white and blue surfaces, respectively. B, close-up views of the black squares shown in A. ORF2 is depicted as a stick model colored as in A. ORF2 amino acid residues not located at the interface between ORF2 and ASP are not shown to facilitate visualization. Hydrogen bonds are shown as black dashed lines and labeled in red numbers. Three hydrophobic amino acid residues of ORF2 located at the interface are labeled in black characters. C, table of the pairs of amino acid residues from ASP and ORF2 that form the hydrogen bonds shown in B.
of subtilisin was expressed independently its catalytic domain, it formed no tertiary structure (30). When we then superimposed the 20 lowest energy structures of the central body region of free ORF2, they showed the proper folding and were identical to the structure within the ORF2-ASP complex. In addition, three hydrophobic residues, Leu<sup>67</sup>, Ile<sup>93</sup>, and Leu<sup>95</sup>, were arranged on the antiparallel β-sheet, suggesting that the central body region is ready for the hydrophobic interaction with ASP, even in the ASP-free state. On the other hand, the N-terminal extension and C-terminal tail were largely disordered with a flexible orientation. In particular, the calculated NMR structures of ORF2 indicated the N-terminal extension to be a long disordered chain with no secondary structural elements. Nonetheless, SSP results showed two peaks with negative scores (−0.92 for Thr<sup>13</sup> and −0.53 for Tyr<sup>21</sup>) between Lys<sup>9</sup> and Leu<sup>22</sup> (data not shown). This implies the region has the potential to form an extended structure, e.g. the β-strand. In fact, this region contained an anti-parallel β-sheet within the structure of the complex (Fig. 3A).

**Both the ORF2 C-terminal Tail and N-terminal Extension Are Important for Binding ASP**—Within the crystal structure of the ORF2-ASP complex, the ORF2 C-terminal tail appears to form numerous hydrogen bonds with ASP. To confirm the interaction, ITC thermograms for ASP (S336A) upon ORF2 binding were examined using ORF2 WT and two deletion mutants lacking either one or five residues from the C-terminal tail (ΔC1 or ΔC5, respectively) (Fig. 7A and Table 3). ORF2 WT strongly bound to ASP (S336A) with a dissociation constant (K<sub>D</sub>) of 1.8 nM (Fig. 7B). On the other hand, the binding affinity of ORF2 ΔC1 was substantially reduced (K<sub>D</sub> = 0.31 μM), and that of ORF2 ΔC5 was reduced even further (K<sub>D</sub> = 8.0 μM). These results indicate that the C-terminal tail tightly binds to the active site of ASP, contributing to the formation of the ORF2-ASP complex and are consistent with the structural observation that the flexible ORF2 C-terminal tail adopts a fixed conformation upon binding to ASP.

We were also interested in characterizing the intrinsically disordered N-terminal extension of ORF2, which is missing...
We therefore assessed the binding to ASP (S336A) of three deletion mutants lacking 11, 43, or 63 residues from the N-terminal extension (ΔN11, ΔN43, and ΔN63, respectively). The binding affinity of ORF2 ΔN11 was nearly the same as the WT protein ($K_D = 2.5$ nM), whereas the affinities of ΔN43 and ΔN63 were somewhat reduced ($K_D = 15$ and 26 nM, respectively) (Fig. 7B). These results show that the contribution of the N-terminal extension to the binding of ORF2 and ASP is smaller than that of the C-terminal tail and are consistent with the finding that the number of hydrogen bonds between ASP and the N-terminal extension is much lower than between ASP and the C-terminal tail.

Both the ORF2 C-terminal Tail and N-terminal Extension Are Essential for ASP Folding—We next used an in vitro protein translation system to express ASP and ORF2 (WT, ΔC1, ΔC5, ΔN11, ΔN43, or ΔN63) and verified the importance of the from the bacterial subtilisin. We therefore assessed the binding to ASP (S336A) of three deletion mutants lacking 11, 43, or 63 residues from the N-terminal extension (ΔN11, ΔN43, and ΔN63, respectively). The binding affinity of ORF2 ΔN11 was nearly the same as the WT protein ($K_D = 2.5$ nM), whereas the affinities of ΔN43 and ΔN63 were somewhat reduced ($K_D = 15$ and 26 nM, respectively) (Fig. 7B). These results show that the contribution of the N-terminal extension to the binding of ORF2 and ASP is smaller than that of the C-terminal tail and are consistent with the finding that the number of hydrogen bonds between ASP and the N-terminal extension is much lower than between ASP and the C-terminal tail.

Both the ORF2 C-terminal Tail and N-terminal Extension Are Essential for ASP Folding—We next used an in vitro protein translation system to express ASP and ORF2 (WT, ΔC1, ΔC5, ΔN11, ΔN43, or ΔN63) and verified the importance of the
ORF2 C-terminal and N-terminal amino acids for the proper folding of ASP. We initially confirmed that the expressed ASP and ORF2 (WT, ΔC1, ΔC5, ΔN11, ΔN43, or ΔN63) migrated precisely as expected in Western analyses (Fig. 7C, upper panel, Non-treated). We then assessed ASP folding based on the protein tolerance to trypsin digestion, because the completely folded structure of ASP is resistant to trypsin digestion. As shown in Fig. 7C (lower panel, + Trypsin), ASP in a reaction mixture containing ORF2 WT was not digested by trypsin, whereas substantial digestion of ASP was seen in reaction mixtures containing ORF2 ΔC1 or ΔC5. These results suggest that ORF2 WT enabled ASP to fold properly, but the C-terminal-deficient ORF2 mutants (ΔC1 and ΔC5) did not.

To verify that idea, we measured the enzymatic activity of ASP in the reaction mixtures. As expected, the reaction mixture expressing ORF2 WT and ASP showed strong ASP activity, but those expressing C-terminally deficient ORF2 (ΔC1 or ΔC5) and ASP did not (Fig. 7D). The C-terminal tail of ORF2 thus appears essential for the proper folding of ASP.

We also evaluated trypsin tolerance to assess ASP folding in the presence of ORF2 ΔN11, ΔN43, or ΔN63. As shown in Fig. 7C, ASP in a reaction mixture containing ORF2 ΔN11 was resistant to trypsin digestion, whereas substantial digestion of ASP was seen in reaction mixtures containing ORF2 ΔN43 or ΔN63. In addition, reaction mixtures containing N-terminally deficient ORF2 ΔN43 or ΔN63 and ASP showed no ASP activity. Thus the intrinsically disordered N-terminal extension also appears important for the proper folding of ASP.

ORF2 and ASP Are Secreted as a Complex from the Periplasmic to the Extracellular Space—The function of ORF2 during secretion of ASP from the periplasmic to the extracellular space is unknown. When ORF2 and ASP (S336A) were co-expressed in A. sobria, both proteins were detected in the periplasmic and extracellular spaces (Fig. 8A). This indicates that both proteins are secreted, but it does not suggest whether they are secreted

### Table 3
Parameters calculated from the ITC thermogram of ORF2 and its deletion mutants (ΔC1, ΔC5, ΔN11, ΔN43, and ΔN63) in Fig. 7B

| Mutation | K_M (M) | ΔG (kcal mol⁻¹) | ΔH (kcal mol⁻¹) | TΔS (kcal mol⁻¹) |
|----------|---------|----------------|----------------|-----------------|
| WT       | 5.24 ± 0.64 x 10⁶ | -11.9 ± 0.3 | -18.6 ± 0.3 | 6.7 |
| ΔC1      | 3.20 ± 0.16 x 10⁶ | -8.9 ± 0.5 | -22.9 ± 0.5 | 14.0 |
| ΔC5      | 1.05 ± 0.03 x 10⁶ | -6.9 ± 0.6 | -5.9 ± 0.6 | -1.0 |
| ΔN11     | 3.66 ± 0.25 x 10⁵ | -11.7 ± 0.0 | -18.6 ± 0.3 | 6.9 |
| ΔN43     | 6.45 ± 0.36 x 10⁻² | -10.6 ± 0.4 | -18.5 ± 0.4 | 7.9 |
| ΔN63     | 4.59 ± 0.34 x 10⁻³ | -10.5 ± 0.6 | -22.8 ± 0.6 | 12.3 |

* Molar ratio of ORF2-ASP binding is calculated as the quotient of ORF2 concentration over ASP concentration.

**Figure 8. Characterization of ORF2 and ASP in vivo and in vitro.** A: immunological detection of ORF2 and ASP in vivo. ORF2 WT was expressed alone in A. sobria T94 cells or was co-expressed with ASP WT or ASP S336A (inactive mutant). ORF2 and ASP in the periplasmic space (Peri) and culture supernatant (CS) were detected using Western analysis. B, effect of pH, salt concentration, and cation type in the interaction of ORF2 with ASP. The relative amount of ORF2-ASP complex in 20 mM Tris-HCl (pH 7.5 or 9.0) buffer were immunologically determined by measuring absorbance of the reaction mixture at 450 nm. C, inhibition of the ASP activity by the interaction with ORF2. Under various conditions examined in Fig. 8B, the proteolytic activities were measured (upper panel). As a control, we performed similar experiments in the absence of ORF2 (lower panel). Purified ASP (10 nM) were mixed with or without ORF2 (100 nM) in 20 mM Tris-HCl (pH 7.5 or 9.0) containing any salts. One unit was defined as the production of 1 pmol AMC/min/μL. D, optimal pH for ASP activity. The activities were measured using a fluorogenic synthetic peptide substrate in 50 mM citrate buffer (circles), 50 mM phosphate buffer (squares), or 50 mM carbonate buffer (triangles). One unit was defined as the production of 1 pmol AMC/min/μL. E, degradation of ORF2 by ASP (65 kDa). 56 μM purified ORF2 treated with the indicated concentrations of purified ASP was separated by SDS-PAGE and visualized with Coomassie staining. The location of intact ORF2 is indicated by the arrow.
as a complex or independently. On the other hand, when ORF2 was expressed alone, it was not detected in the extracellular space, which means ORF2 secretion does not occur unless ASP is also present in the periplasmic space. Given that the structure of the ORF2-ASP complex represents the postfolding state, that the folding of ASP is completed within the periplasmic space, and that the osmotic strength of the periplasmic space is estimated to be 300 mOsm (31), sufficient to mediate ORF2-ASP binding in vitro (Fig. 1A), it seems most likely that ORF2 is secreted in complex with ASP from the periplasmic to the extracellular space. The importance of the osmotic strength in the formation of the ORF2-ASP complex was further supported by our additional experiments. As shown in Fig. 8B, the ORF2-ASP complex was markedly formed in the buffer (pH 7.5) containing any salts (not only NaCl but also KCl or NH₄Cl) but not in the same buffer without salts. This suggests that the salt effect in the formation of the ORF2-ASP complex was not specific because the effect was seen using any salts. That is, ionic strength largely affected the interaction between ORF2 and ASP.

ASP Dissociates from ORF2 and Then Degrades ORF2 in the Extracellular Space in a pH-dependent Manner—Interestingly, when ORF2 and ASP WT were co-expressed in A. sobria, both proteins were detected in the periplasmic space, but only ASP was detected in the extracellular space (Fig. 8A). On the other hand, if ORF2 was co-expressed with the inactive mutant ASP (S336A), ORF2 was detected in extracellular space. This suggests that ORF2 is degraded by ASP after their secretion. Moreover, because it appears ASP is secreted in complex with ORF2, we suggest ASP dissociates from ORF2 prior to degrading it. To test that idea, we endeavored to determine the driving force for the dissociation of the ASP-ORF2 complex. Under the culture conditions (LB broth, pH 7.5) for A. sobria, the NaCl concentration was ~0.17 M, which was sufficient to mediate ORF2-ASP binding in vitro (Fig. 1A), and that ionic strength did not vary with time. By contrast, the pH of the culture medium increased from 7.5 to 9.0 over a period of about 16 h. This suggests that the driving force for dissociation could be the extracellular pH. Consistent with that idea, when independently purified ASP and His₆-tagged ORF2 were mixed in pH 7.5 or 9.0 buffer, immunological assays using Ni-NTA HisSorb showed that much less complex remained at pH 9.0 than at pH 7.5, irrespective of the NaCl, KCl, or NH₄Cl concentration (Fig. 8B, right columns). Furthermore, we measured the proteolytic activities of the samples containing ASP at pH 7.5 and 9.0 with or without addition of ORF2. Comparing with controls (Fig. 8C, lower panel), the proteolytic activities were remarkably inhibited by external addition of ORF2 into the sample buffer (pH 7.5) containing any salts, but not in the cases using the sample buffer (pH 9.0) containing any salts (Fig. 8C, upper panel). Because the pH level in the periplasmic space is thought to be neutral, we suggest the change in pH from the periplasmic to the extracellular space leads to the dissociation of the ORF2-ASP complex and the generation of active ASP. Furthermore, pH 9.0 is the optimal pH for ASP activity in vitro (Fig. 8D), and ASP dose-dependently degrades ORF2 at that pH (Fig. 8E).

**DISCUSSION**

In this study, we first revealed that formation of a stable ORF2-ASP complex for crystallographic analysis requires the presence of high ionic strength. In addition to Na⁺, other relevant cations present in the periplasmic space, including K⁺ and NH₄⁺, are also able to mediate ORF2-ASP interaction (Fig. 8, B and C). Based on the fact that we did not observe any ions at the ASP-ORF2 interface, we think that the salt effect is nonspecific and that ionic strength is one of important factor to form the ORF2-ASP complex. Interestingly, this interaction was uncoupled at pH 9.0. At that pH, both proteins carry a negative
charge, because the pI values for ORF2 and ASP are 6.8 and 5.6, respectively.

The crystal structure of the ORF2-ASP complex showed that the structure of ORF2 within the complex is similar to that of the propeptide in complex with its cognate subtilase. Furthermore, ORF2 is essential for the folding of ASP, just as the propeptide is essential for the folding of its cognate subtilase. Although this suggests that ORF2 functions as a chaperone in a manner similar to propeptide, there are several substantive differences: 1) ORF2 is not an intramolecular chaperone; it is an external chaperone encoded separately from ASP; 2) ORF2 contains an extra 63-amino acid N-terminal extension (Fig. 5B) that is intrinsically disordered but develops a degree of secondary structure upon ASP binding; and 3) the central body region of ORF2 has a novel structure that is not present in the propeptide.

![FIGURE 10. Sequence alignment of ASP and ORF2 homologs.](image)

Conserved and similar residues are shown in green and orange, respectively. A, sequence alignment of ASP homologs. Inverted red triangles show residues of the ASP catalytic triad. Double-headed arrows indicate extra regions that Kexin family proteins do not have (19). B, sequence alignment of ORF2 homologs. Double-headed arrows show the domain structure of ORF2 as in Fig. 3.
of ORF2 has tertiary structure in solution, whereas both NMR and CD studies showed that the propeptide expressed independently of subtilase has no tertiary structure under physiological conditions (30, 32, 33).

The C-terminal tail of ORF2 forms numerous hydrogen bonds with ASP within the structure of the ORF2-ASP complex, and ITC measurements showed the importance of the C-terminal amino acids for the interaction between ORF2 and ASP. In addition, the huge $\Delta H$ decrease of ORF2 $\Delta C5$ caused by ASP binding (Table 3) supports the formation of numerous hydrogen bonds between the C-terminal tail and the active site of ASP. The importance of the ORF2 C-terminal amino acids for ASP folding was further indicated by our observation that the ORF2 $\Delta C1$ and $\Delta C5$ deletion mutants had markedly less ability to mediate folding than ORF2 WT. This means that the ability of ORF2 to bind ASP correlates positively with its ability to mediate folding.

We observed that the ORF2 N-terminal extension containing the two antiparallel $\beta$-sheets (Ser$^{12}$–Ile$^{23}$ and Gly$^{48}$–Val$^{62}$) interacts with the ASP additional domain. One $\beta$-sheet (Ser$^{12}$–Ile$^{23}$) is the region in which the potential formation of a $\beta$-strand was implied by SSP analysis (data not shown). Interestingly, the ORF2 $\Delta N43$ and $\Delta N63$ deletion mutants obviously lost the ability to mediate folding, although their binding constants were reduced by an order of magnitude or less. In contrast to the C-terminal tail, the function of the N-terminal extension is apparently not to inhibit ASP. Given that the contribution of the N-terminal extension of ORF2 to the binding to the folded ASP is small, it seems likely that interactions not observed in the crystal structure are important for ORF2 to function as a chaperone during folding. We conclude that all three regions, including the central body, contribute in assisting the folding of ASP. The mechanism by which protein-specific chaperones act remains undetermined, although there are a few reports in which interactions between propeptide and subtilase have been identified (34, 35). Further studies will be necessary to determine how the intrinsically disordered C-terminal tail and N-terminal extension of ORF2 interacts with ASP during folding.

The maturation of subtilase involves at least three distinct steps: 1) folding from a metastable intermediate, which is mediated by the propeptide domain; 2) autoproteolytic cleavage between the propeptide and catalytic domains; and 3) degradation of the cleaved propeptide (36). Maturation mediated by ORF2 appears to be different and divided into two major steps, without a cleavage step: 1) ASP folding mediated by ORF2 and 2) dissociation of ORF2 from the mature ASP, leading to ORF2 degradation (Fig. 9). It is known that ORF2 and ASP interact within the periplasmic space (29). Our finding that the crystal structure of the ORF2-ASP complex is in a postfolding state indicates that ASP folding is completed within the periplasmic space. In addition, Western analysis showed that ORF2 is secreted from the $A$. sobria periplasmic space into the extracellular space in complex with ASP. This suggests that ORF2 forms a stable complex with ASP within the periplasmic space, where it mediates folding. Interaction with ORF2 also likely protects the unfolded ASP from intrinsic protease, because ASP is not observed in the periplasmic space of $E$. coli unless it is co-expressed with ORF2 (14). Once the ORF2-ASP complex is translocated to the extracellular space, ASP dissociates from
ORF2, enabling it to act as a virulence factor. We showed that the pH change from the periplasmic to the extracellular space is sufficient to trigger that dissociation, which ultimately leads to the degradation of ORF2 by ASP.

ASP is thought to belong to the Kexin family because it contains a P-domain, which other subtilase families lack (Fig. 3A) (20). However, ASP differs from other Kexin proteins in that it has no N-terminal propeptide; instead, asp and orf2 form an operon, and ORF2 mediates ASP folding after expression. In addition, ASP contains six extra regions not seen in other Kexin proteins (19). These differences permitted us to reassess the phylogenetics of ASP. When we performed a PSI-BLAST search with the ORF2 sequence, we found that operons consisting of genes homologous to asp and orf2 are present in five genera: Aeromonas, Vibrio, Shewanella, Chromobacterium, and Pseudoalteromonas. Although the sequence similarity among the ORF2 homologs is very low, the sequence identity among the ASP homologs is 23% (Fig. 10). This tendency is similar to that seen with other subtilases, where the amino acid sequence of the catalytic domain is well conserved, but there is little conservation of the sequence of the N-terminal propeptide, which functions as the intramolecular chaperone (6). In addition, no ASP homologs had an N-terminal propeptide, and all had extra regions, like the ASP examined in the present study (Fig. 10). Taken together, these findings indicate that ASP homologs differ from Kexin family proteins and form a new family relying on external chaperones for folding (Fig. 11).

Acknowledgments—We thank Y. Tsumori for crystallization. We are grateful to the beamline staff at KEK Photon Factory for help with data collection.

REFERENCES

1. Hendrick, J. P., and Hartl, F. U. (1993) Molecular chaperone functions of heat-shock proteins. Annu. Rev. Biochem. 62, 349–384
2. Pauwels, K., Van Molle, I., Tommassen, J., and Van Gelder, P. (2007) Chaperoning Anfinsen: the sterile foldases. Mol. Microbiol. 64, 917–922
3. Zhu, X. L., Ohta, Y., Jordan, F., and Inouye, M. (1989) Pro-sequence of subtilisin can guide the refolding of denatured subtilisin in an intermolecular process. Nature 339, 483–484
4. Eder, J., Rheinheimer, M., and Fersht, A. R. (1993) Folding of subtilisin BPN’; characterization of a folding intermediate. Biochemistry 32, 18–26
5. Sohl, J. L., Jaswal, S. S., and Agard, D. A. (1998) Unfolded conformations of α-lytic protease are more stable than its native state. Nature 395, 817–819
6. Shinde, U., and Thomas, G. (2011) Insights from bacterial subtilases into the choice between kinetic and thermodynamic protein folding and stability in subtilases. Biochemistry 43, 14348–14360
7. Kobayashi, H., Takaeshi, E., Oguma, K., Fujii, Y., Yamanaka, H., Negishi, T., Arimoto-Kobayashi, S., Tsuji, T., and Okamoto, K. (2006) Cleavage specificity of the serine protease of Aeromonas sobria, a member of the kexin family of subtilases. FEMS Microbiol. Lett. 256, 165–170
8. Gamble, M., Künze, G., Dodson, E. J., Wilson, K. S., and Jones, D. D. (2011) Regulation of an intracellular subtilisin protease activity by a short propeptide sequence through an original combined dual mechanism. Proc. Natl. Acad. Sci. U.S.A. 108, 3536–3541
9. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. J. Appl. Crystallogr. 40, 658–674
10. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and current developments of the CCP4 suite and current development of Coot. Acta Crystallogr. D Biol. Crystallogr. 66, 486–501
11. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A. C., Nicholos, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011) Overview of the CCP4 suite and current developments. Acta Crystallogr. D Biol. Crystallogr. 67, 235–242
12. Lee, J., and Güntert, P. (2003) Influence of the completeness of chemical shift assignments on NMR structures obtained with automated NOE assignment. J. Struct. Funct. Genomics 4, 179–189
13. Kobayashi, H., Tateshii, A., Tsuge, H., Takahashi, E., Okamoto, K., and Yamanaka, H. (2009) The carboxy-terminal tail of Aeromonas sobria serine protease is associated with the chaperone. Microbiol. Immunol. 53, 647–657
14. Nomura, T., Fujii, Y., Yamanaka, H., Kobayashi, H., and Okamoto, K. (2002) The protein encoded at the 3′ end of the serine protease gene of Aeromonas sobria functions as a chaperone in the production of the protease. J. Bacteriol. 184, 7058–7061
15. Buevich, A. V., Shinde, U. P., Inouye, M., and Baum, J. (2001) Backbone dynamics of the natively unfolded pro-peptide of subtilisin by heteronuclear NMR relaxation studies. J. Biol. Chem. 276, 233–249
16. Stock, J. B., Rauch, B., and Roseman, S. (1977) Periplasmic space in Salmonella typhimurium and Escherichia coli. J. Biol. Chem. 252, 7850–7861
17. Eder, J., Rheinheimer, M., and Fersht, A. R. (1993) Folding of subtilisin BPN’; role of the pro-sequence. J. Mol. Biol. 233, 293–304
33. Shinde, U., Li, Y., Chatterjee, S., and Inouye, M. (1993) Folding pathway mediated by an intramolecular chaperone. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 6924–6928
34. Inouye, M., Fu, X., and Shinde, U. (2001) Substrate-induced activation of a trapped IMC-mediated protein folding intermediate. *Nat. Struct. Biol.* **8**, 321–325
35. Tanaka, S., Matsumura, H., Koga, Y., Takano, K., and Kanaya, S. (2009) Identification of the interactions critical for propeptide-catalyzed folding of Tk-subtilisin. *J. Mol. Biol.* **394**, 306–319
36. Jain, S. C., Shinde, U., Li, Y., Inouye, M., and Berman, H. M. (1998) The crystal structure of an autoprocessed Ser221Cys-subtilisin E-propeptide complex at 2.0 Å resolution. *J. Mol. Biol.* **284**, 137–144
37. Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948
38. Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **28**, 2731–2739