Direct Formation of Human Interleukin-11 by Cis-Acting System of Plant Virus Protease in Escherichia coli

Tohru TAKAHASHI, Michiko NAKANISHI, Yoshio YAO, Ichiro UYEDA, and Nobufusa SERIZAWA

Biomedical Research Laboratories, Sankyo Co. Ltd., 2-58 Hiromachi 1-chome, Shinagawa-ku, Tokyo 140-8710, Japan

1Laboratory of Plant Virology and Mycology, Department of Agrobiology and Bioresources, Faculty of Agriculture Hokkaido University, Kita 9, Nishi 9, Kita-ku, Sapporo, Hokkaido 060-8589, Japan

Received October 30, 1997

To produce a large amount of recombinant proteins in Escherichia coli, we constructed a unique cis-acting expression system using a plant virus protease. This new expression system could directly produce recombinant proteins, that had a biologically active form. A gene of nuclear inclusion-a (Nia), which had a specific amino acid sequence, was fused with a foreign protein gene at the same protein reading frame. One of the Nia-specific cleavage amino acid sequences, Gln-Ala, was also contained at the protein-protein junction. In the case of human interleukin-11 (hIL-11), a 23-kDa specific signal band was obtained from recombinant bacteria. N-terminal sequencing of the 23-kDa protein showed that Nia specifically cleaved the fusion protein at Gln-Ala, producing Ala-hIL-11. Furthermore, we could produce the mature rHL-11 by extending the culture time. This 23-kDa protein had the same biological activity as hIL-11 in a mouse plasmacytoma, T1165. Combined with fermentation control, we produced mature rHL-11 in E. coli.

Key words: gene expression; clover yellow vein virus; potyvirus; nuclear inclusion-a; interleukin-11

Among the many approaches taken to improve the yield and purity of recombinant proteins, one particularly useful procedure is to express the desired polypeptide as part of a larger fusion protein with, for example, the glutathione S-transferase (GST) expression system or protein-A fusion system. Such fusion proteins can be highly purified in good yield in a single step by passing them through columns of an appropriate affinity matrix. However, in some cases, fusion proteins form inclusion bodies in E. coli. These insoluble fusion proteins cannot be applied to an affinity column, and also the linker peptides cannot be digested by specific proteases. In fact, we unsuccessfully attempted to produce some clinically important proteins like cytokines and allergenic antigens using direct or GST-fusion systems.

Human interleukin-11 is a clinically important protein, which acts as a platelet differentiation factor or adipogenesis inhibitory factor. First, we tried to express the mature human interleukin-11 (hIL-11) gene by a simple adjustment of the ATG codon upstream from the gene, but failed. We also tried to make rHL-11 using a GST expression system. Purified GST-IL-11 fusion protein could be partially digested by enterokinase or factor Xa, but an extra amino acid residue was retained at the N-terminus of the product. Thus we had to remove this N-terminal extra amino acid by further enzyme treatment. To solve these problems, we constructed a unique fusion protein expression system using the autocatalytic action of a plant virus protease.

Clover yellow vein virus (CIYVV) is a member of the genus Potyvirus, which belongs to the picorna-like virus superfamily. Potyviruses have an uninterrupted large open reading frame (ORF) in their genomic ssRNA. A single polypeptide precursor of about 350 kDa is cleaved by cis- and trans-acting viral-coded proteases, resulting in the maturation of a number of different low molecular weight proteins. One of these viral proteases, corresponding to the nuclear inclusion-a (Nia) protein, is responsible for several proteolytic processing events of the large viral polypeptide precursor. Specifically cleaves Gln-Ala, Gln-Gly, or Gln-Ser in a substrate peptide; and these reactions occur not only in cis (cleaving the polyprotein containing the protease) but also in trans (cleaving different polyproteins) fashion. We used this CIYVV-Nia as a driving motor of this unique expression system.

Unlike most of well-known fusion systems, like a GST, that are driven by trans-acting systems, our plant virus protease fusion system is driven by a cis-acting system. In the cis-acting system, the target protein is produced as a protease-fusion protein, and after that, it will mature in the host cells.

In the expression vector described here, the foreign gene is placed downstream of the CIYVV-Nia protease gene. With this expression system, we could produce the protein with an alanine, glycine, serine, or proline residue at its N-terminus.

* To whom correspondence should be addressed: Biomedical Research Laboratories, Sankyo Co. Ltd., 2-58 Hiromachi, 1-chome, Shinagawa-ku, Tokyo 140-8710, Japan. Tel: +81-3-3492-3131; Fax: +81-3-5436-8565; E-mail: SERIZA@shina.sankyo.co.jp.
Materials and Methods

cDNA synthesis and molecular cloning. Isolate No. 30 of CIYVV was originally obtained from *Phaseolus vulgaris* and propagated in *Vicia faba* in a greenhouse.\(^1\) A necrotic strain of PVY was also originally isolated from potato plants and propagated in *Nicotiana sylvestris*.\(^2\) The viruses were purified from the infected plants using the method previously described,\(^1\) and the viral RNA was extracted by an alkaline sucrose density gradient method.\(^3\) Complementary DNA synthesis and molecular cloning were done by the method of Gubler and Hoffman.\(^4\)

Nucleotide sequencing. The resultant clones, pNS51 from CIYVV and p\(^\text{pVTDt88}\) from PVY-N, were subcloned into M13mp18 and M13mp19, and a series of sequential deletions of the inserts were made for sequencing using Exonuclease-III. Sequences were analyzed by the dideoxy-chain termination method.

Construction of pKSun9 expression vector. An NL\(a\)/hlIL-11 fusion gene was made by the polymerase chain reaction (PCR). To add the initiation codon ATG and a recognition site for the restriction enzyme *NcoI* suitable for cloning on the 5' terminus of CIYVV-NL\(a\) (NL\(a\)'), PCR primers were prepared. The sequences prepared in these experiments are shown in Table I. One \(\mu\)g of the template DNA, pNS51, and 20 pmol of a set of primers (NSATG, NSX1) were mixed in 100 \(\mu\)l of a reaction solution containing 8 \(\mu\)l of dNTPs (5 mM each of dATP, dCTP, dGTP, and dTTP) and 5 units of *Pfu* DNA polymerase (Stratagene). The PCR program consisted of a cycle of 92°C for 1 min, 37°C for 1 min, and 72°C for 2 min, which was repeated 20 times. The resulting amplified DNA was extracted with phenol and separated by 5% polyacrylamide gel electrophoresis. The DNA (NL\(a\)') band thus obtained was cut out and recovered by electroelution. The plasmid pKK388X was made from pKK388-1 (Clontech) by replacing the *SacI* site with an *XhoI* site. The eluted DNA was then digested with the restriction enzymes *NcoI* and *XhoI*, and inserted into the plasmid pKK388X (pKN\(\text{N}15\)').

Plasmid pcD20-2 contains the cDNA coding for an hIL-11 precursor sequence. This plasmid was cleaved with the restriction enzymes *BamHI* and *Apal*, and a region containing both the pre-mature IL-11 and SV40 promoter was excised. The fragment was inserted into the *BamHI* and *Apal* sites of pBLUESCRIPT II SK+ and then cleaved again with the restriction enzymes *XhoI* and *KpnI*, resulting in a gene coding for mature IL-11 protein that was devoid of the N-terminus of the signal and pre-sequences. The mature IL-11 fragment was inserted between the *XhoI* and *KpnI* restriction sites of the pKN15', and the resulting plasmid was named pKN15IL.

To connect the specific sequence at the C-terminus of NL\(a\) with mature IL-11 via alanine while keeping the same reading frame, we synthesized four kinds of PCR primers. The first PCR reaction was done using a pair of primers, NSX2 and NSJ002N, with pNS51 DNA as a template to amplify the coding region for the C-terminus of NL\(a\) (CN13). On the other hand, another PCR reaction was also done using the pair of primers NSJ001P and ILSAC, with pcD20-2 as template to amplify the coding region for the N-terminus of the hIL-11 peptide (5'IL). After PCR, each product was separated on a 5% polyacrylamide gel electrophoresis (PAGE), and stained with ethidium bromide. Each eluted DNA fragment, CN13 and 5'IL, was combined and re-amplified using another pair of primers, NSX2 and ILSAC. After PCR, the fused DNA was isolated by 5% PAGE. The recovered fused DNA was digested with the restriction enzyme *XhoI*, and inserted into plasmid pKN15IL, which had been previously digested with *XhoI*. The resulting plasmid was named pKSun9. The recombinant plasmid was introduced into *E. coli* JM109.

Expression of NL\(a\)-hIL-11 fused gene in *E. coli*. *E. coli* JM109 carrying plasmid pKSun9 was inoculated into LB medium containing 40 \(\mu\)g/ml ampicillin, and the culture was shaken at 37°C overnight. Two ml of the overnight culture was then added to 200 ml of fresh LB medium with 40 \(\mu\)g/ml of this antibiotic, and the culture was further incubated at 37°C. When the bacterial culture reached an optical density of approximately 1.0 at 600 nm, 1 mM isopropylthio-\(\beta\)-D-galactoside (IPTG) was added induce expression of the recombinant chimeric gene. Following additional growth at 28°C for 12 or 36 h, the cells were harvested by centrifugation at 3,500 \(\times\)g for 15 min at 4°C. The cell pellets were stored at -80°C before purification of the protein.

Western blot. Purified protein (20 ng each) or crude *E. coli* cell debris (1 \(\mu\)l of culture each) was on 12.5% SDS-polyacrylamide gel electrophoresed (SDS-PAGE), and transferred to a PVDF membrane by the method of Towbin et al.\(^5\) After having been blocked with 5% skim-milk, the PVDF membrane was incubated with rabbit anti-hIL-11 polyclonal antiserum at a 1:10,000 dilution, followed by incubation with a 1:5,000 dilution of horseradish peroxidase-conjugated donkey anti-rabbit IgG(H+L) (Amersham). The PVDF membrane was then treated with ECL detection reagent (Amersham), and the bands that reacted with the anti-hIL-11 antibody were made visible by bringing the PVDF membrane into contact with an X-ray film for a period of 2 min and 30 sec.
Extraction and purification of recombinant hIL-11 from E. coli. The frozen cells were resuspended in ten volumes (w/v) of an extraction buffer containing 50 mM borate-NaOH (pH 9.0) and 1 mM DTT. The suspension was homogenized using a French-Press homogenizer (Ohtake Seiko Japan). Following disruption, it was then centrifuged at 10,000 × g for 30 min at 4°C, and the clear supernatant was collected. The supernatant was then put through weak ion-exchange column chromatography using a CM-Toyopearl Pack 650M column (Toso Japan). The concentration gradient used was a linear gradient from eluent A (10 mM borate-NaOH, 13 mM KCl, pH 9.0) to eluent B (10 mM borate-NaOH, 13 mM KCl, 400 mM NaCl, pH 9.0) over a period of 120 min. Each eluted fraction was subsequently western blotted to identify fractions containing rhIL-11.

N-terminus amino acid sequencing. The fractions containing a substance that reacted with rabbit anti-hIL-11 antibody were pooled, concentrated with Centiprep-10 (Amicon), and then run on a 12.5% SDS-PAGE gel. After electrophoresis, the gel was electro-botted to a ProBlott membrane (Perkin Elmer). The bands on the membrane were stained with Coomassie Brilliant Blue R250. The bands that reacted with anti-hIL-11 antibody were excised and used for amino acid sequencing with a 477A protein sequencer (Perkin Elmer).

Biological activity of rhIL-11 toward plasmacytoma cell line T1165. The effects of purified rhIL-11 on plasmacytoma cell survival assay were assayed with the T1165-834 cell line,\(^1\) which is dependent on IL-11 for survival. In brief, serially diluted samples were cultured with 1 × 10\(^4\) cells/100 µl on a microtiter plate for 48 h. The cell survival rate was measured by a 3-[4,5-dimethylthiazol-2-yl]-2,5-diphenyltetrazolium bromide (MTT) assay.\(^10\) The rhIL-11 activity was expressed in comparison with that of standard rhIL-11, which was expressed in CHO cells transfected with pSRα1-14 (IL-11 cDNA driven by SRα promoter).

Miscellaneous procedures. Isolation of plasmid DNA, preparation of DNA fragments, DNA synthesis reactions, DNA ligations, and screening of plasmid-containing colonies were done essentially as described by Maniatis et al.\(^17\) Transformation of E. coli JM109 was done by the manufacturer’s protocol. Oligonucleotides were synthesized with a Perkin Elmer model 394.

Results
Characterization of the cDNA encoding CIYVV-N1a region
The cDNA clone pNS51, having about 6.5 kbp, was assigned to the 3′-terminal region of the CIYVV genome by Southern blotting hybridization and by restriction enzyme mapping (Fig. 1). From this restriction map, we supposed that N1a was located between Pst I and Sal I (PS fragment) and Sal I and Sal I (SS fragment) sites. The nucleotide sequence data reported in this paper will appear in the DDBJ, EMBL, and GenBank nucleotide sequence databases with the following accession number AB002698. A comparison with TVMV cDNA found 55% similarity for the nucleotide sequence, and the predicted amino acid sequence had 49% identity. It has been shown that most of the structural proteins of potyviruses are cleaved at the Gln-Ala, Gln-Gly, or Gln-Ser sites. Based on the predicted cleavage site together with the amino acid sequence homology, the N1a coding region was assigned to the sequence from 1560 to 2861.

Isolation of the N1a structural gene
The N1a gene was isolated from pNS51 by use of the polymerase chain reaction method (PCR), and the hIL-11 gene was isolated from pcD20-2 carrying hIL-11 cDNA.\(^3\) These genes were joined with a specific linker that coded for a proteolytic processing amino acid sequence (Gln-Ala), and were inserted into the E. coli expression vector pKK388-1. The resulting vector, pKSun9, which carried the N1a-Gln-Ala-hIL-11 fused gene, was expressed by the trc promoter (Fig. 2A). For comparison with other expression systems, we also prepared a mature hIL-11 type vector with an additional initiation codon added (pMIL), and other constructs of the N1a-fusion vector (Fig. 2B). Another type of N1a was isolated from a necrotic strain of potato virus Y (PVY),\(^12\) which is a well-known type member of the genus Potyvirus. In the case of GST- or other protein fusion systems, we found that the N-terminus proline cluster of hIL-11 inhibited protease cleavage (data not shown). Consequently, we made two types of expression vectors from PVY-N1a. The pAPVIL expression vector has the PVY-N1a original cleavage site, but some extra sequences inserted upstream of the proline cluster-deficient hIL-11. On the other hand, pPVIL was constructed in the same way as pKSun9.

Expression of hIL-11 by the auto-cleavage system of N1a
The expression vector plasmid pKSun9 and others were used to transform E. coli JM109 (hereinafter, the E. coli clones containing these plasmids are abbreviated as KSUN9, MIL, APVIL, and PVIL). The methionine-
Each gene is inserted into the E. coli expression vector pKK388-1, and promoted by the tac promoter. Plasmid pMIL is simply one with ATG in front of the N-terminal proline of the mature hIL-11. Plasmid pΔPVIL contains PVV-Nla and hIL-11 genes joined in the same reading frame, but in this plasmid, the hIL-11 gene is incomplete (lack from +1 to +7: N'PGPPPGP-C'). Plasmid pΔPVIL has an original specific cleavage site and 55 amino acids of the N-terminus of Nlb. These genes are fused in the same reading frame. Plasmid pPVIL has an original specific cleavage site and joined in the same reading frame with complete hIL-11 gene.

Fig. 2A. Schematic Illustration of hIL-11 Expression System under Direct of Nla-Fusion Systems.

The C1VV-Nla and hIL-11 fused genes are controlled by the tac promoter. The remaining segment was derived from pKK388-1. All known functional genes present in the plasmid and the direction of transcription are marked by arrows. The white box on the sequence shows the specific cleavage site of the C1VV-Nla.

fused hIL-11 or Nla/hIL-11 fusion proteins were expressed by induction with 1 mM IPTG. The expressed recombinant proteins were analyzed by western blotting. Cell extracts were blotted with an anti-hIL-11 polyclonal antibody. Twenty-three-kDa and 60-kDa protein signal bands, specific to the cell extract of KSUN9, were detected (Fig. 3). The 23-kDa band migrated at the same rate as mature hIL-11. The yield, measured by western blotting, was about 45 mg from a liter of culture. In the case of ΔPVIL, we also detected two bands, one at 25 kDa and one at 60 kDa. In contrast, only the 60-kDa band was detected in the cell extract of PVIL, and no band was detected in the cell extract of MIL.

**Fig. 2B. Principal Elements of pKSun9.**

The C1VV-Nla and hIL-11 fused genes are controlled by the tac promoter. The remaining segment was derived from pKK388-1. All known functional genes present in the plasmid and the direction of transcription are marked by arrows. The white box on the sequence shows the specific cleavage site of the C1VV-Nla.

**Fig. 3. Western Analysis of the Expressed Recombinant hIL-11.**

Lane 1: whole cell extract of E. coli JM109 containing no plasmid. Lane 2: MIL. Lane 3: KSUN9. Lane 4: PVIL. Lane 5: ΔPVIL. These samples were electrophoresed on a SDS-PAGE and then immunoblotted with anti hIL-11 rabbit polyclonal antibody. The positive control (lane 6: 50 ng, and 7: 100 ng) was purified from recombinant S. picha, which expressed recombinant hIL-11 lacking the N-terminal proline.

**Purification and Identification of 23-kDa Protein in E. coli KSUN9**

The 23-kDa recombinant protein was partially purified from E. coli KSUN9 by ion-exchange column chromatography. The eluted protein was analyzed by a direct antibody sandwich enzyme linked immunosorbent assay (DAS-ELISA) using hIL-11 polyclonal and monoclonal antibodies. The fractions positive by DAS-ELISA were collected and concentrated. The concentrated fraction was run on SDS-PAGE and transferred onto a nylon membrane. The specific protein band, migrating at the same rate as rhIL-11, was cut out and used for amino acid sequence analysis (Fig. 4). The N-terminal amino acid sequence of the 23-kDa protein band, which was purified 12 hr after IPTG induction (23 kDa-12 h), was Ala-Pro-Gly-Pro-Pro-Pro-Gly. This sequence corresponded to the amino acid sequence from position -1 to position +6 of the mature hIL-11 protein. On the other hand, the amino acid sequence of the N-terminus of the 23-kDa protein band, which had been purified 36 h after IPTG induction (23 kDa-36 h) was Pro-Gly-Pro-Pro-Gly-Pro. This sequence corresponded to that from position +1 to +7 of the mature hIL-11. Together with data obtained from mass spectrum analysis, it was found that more than 95% of the 23 kDa-36 h protein had a proline residue at the N-terminus (data not shown).

**Biological Activity of rhIL-11 on T1165 Stromal Cell Line**

The biological activity of rhIL-11 (23 kDa-36 h) was
Fig. 4. SDS-PAGE Analysis of hIL-11 Protein.
Each sample was electrophoresed on a SDS-15% polyacrylamide gel under reducing conditions. Lane 1: proteins size markers. Lane 2: whole cell extract obtained from E. coli KSUN9. Lane 3: hIL-11 partially purified by CM-Toyopearl column chromatography.

Fig. 5. Biological Activity of Recombinant hIL-11.
Dose-response of survival of T1165-834 cell line treated with E. coli (●) or CHO (■) expressed hIL-11. Cell survival was measured by the MTT colorimetric assay. Several 3-fold dilutions of the hIL-11 were prepared for the assay, the initial concentration of hIL-11 was 100 ng/ml.

compared with that of rhIL-11 expressed in the CHO cell line, with the IL-11-dependent mouse T1165 stromal cell line used as the indicator. As shown in Fig. 5, rhIL-11 expressed in this E. coli expression system showed the cell survival activity in a dose-dependent manner, to the same extent as the authentic rhIL-11 derived from the CHO cell line.

Discussion
Many kinds of proteins are produced using recombinant technology. Of these techniques, fusion protein expression systems are most useful. However, in some cases, fusion proteins formed insoluble inclusion bodies and the requisite proteins cannot be isolated from them by protease digestion. Human IL-11 is one example of a protein difficult to produce by the GST-fusion system. Although the GST-IL11 fusion protein was not insoluble, the purified fusion protein could not be easily digest-
ed with trans-acting systems by using enterokinase and factor-Xa endoproteases (data not shown). This was also the case for some allergenic antigen proteins. For this reason, we tried to create a cis-acting proteolytic processing system using plant virus processing enzymes that can directly produced biological active mature proteins in E. coli.

Human IL-11, a stromal cell-derived cytokine, is known to play important roles in the hematopoietic system. It supports megakaryocyte colony formation and maturation, and acts as an autocrine growth factor in a megakaryoblastic cell line. Paul et al.1) and others2,3) have previously isolated hIL-11 cDNA from the stromal cell lines PU-34 and KM-102, and expressed it in COS-1 and CHO cells. In their in vitro studies, they used hIL-11 purified from COS-1 cells. Because of the low yield by the CHO expression system, we decided to develop other expression systems to get larger amounts of rhIL-11 for the examination of in vitro properties of hIL-11 in detail. First of all, we tried to express the mature hIL-11 gene with the simple adjustment of the ATG codon upstream from the mature gene. However, when an extract derived from E. coli 3JM109/MIL was analyzed by western blotting, we could not detect any immunoreactive pre-mature hIL-11 with a methionine residue at the N-terminus of mature hIL-11. The failure of this attempt was attributed to the mRNA secondary structure at the translation initiation site. In a previous work on the production of hIL-6,18,19) it was also difficult to express pre-mature hIL-6 from cDNA with an initiation codon followed by the coding sequence. Because of the high GC-content in hIL-6 mRNA, ribosomes could not easily scan this mRNA during the translation process. On the other hand, hIL-6 mRNA could be expressed when there was a change from a high to a low GC-content in one third of the 5' end sequence of the mRNA. We could not change the 5' end of hIL-11 mRNA by using alternative codons, because it had oligo-proline residues at its N-terminus.

In these experiments, we investigated the use of two typical potyviruses: potato virus Y12) and clover yellow vein virus.11,20) In an in vitro translation experiment, PVY-Nla and CIYVV-Nla also had protease activity (data not shown). To investigate whether these proteases could process the Nla and hIL-11 fused protein, we made PVY-Nla/hIL-11 and CIYVV-Nla/hIL-11 fusion protein expression vectors. In these experiments, we succeeded in the direct expression of hIL-11 when CIYVV-Nla (E. coli JM109/KSUN9) was used. On the other hand, when PVY-Nla (E. coli JM109/PVIL) was used, no hIL-11 was processed from the fusion protein. Although we also examined the use of the GST fusion system for rhIL-11, complete cleavage of the fusion protein was difficult using enterokinase. In comparison to PVY-Nla, CIYVV-Nla has higher specificity and broader utility than other potyvirus proteases, and therefore, it can be used in other fusion protein expression systems. In fact, we succeeded in producing other clinically important proteins like the KM-102 derived reductase like-factor,21) interferon and some allergenic antigens (data not shown). These proteins are also difficult to
produce by commonly used expression systems.

Surprisingly, we were able to produce mature hIL-11 with a proline residue at its N terminus. In the culture experiments, the induced Nla and hIL-11 fusion protein was digested by Nla's auto-catalytic action at the early stage in the host cell. During this stage, E. coli had Nla and pre-mature hIL-11, which had an alanine residue at its N terminus. Following the proteolytic digestion, the alanine residue was completely removed from the N terminus of the pre-mature hIL-11 by extending the cultivation time. It was previously reported that, if the penultimate amino acid was proline, the N terminal methionine was removed by endogenous methionine aminopeptidase.22,23 Our results suggest that the N terminal alanine residue can also be removed from the N terminal of pre-mature rhIL-11 by an endogenous enzyme such as aminopeptidase P.

We expect that our unique auto-catalytic Nla fusion expression system will prove useful for producing proteins with an alanine, serine, glycine, or proline residue at the N terminus. In addition, we recently found that CIYVV-Nla is not only a cis-active protease, but also a trans-active protease in vitro and in vivo. Especially in vivo, CIYVV-Nla, which was quite active in mammalian cell like a 293 human cell line and completely digested the G-CSF receptor fusion protein.25 Because of this protease activity, CIYVV-Nla will become widely become useful as specific endoprotease like enterokinase.

Acknowledgments

We thank Ryuta Koishi and Chigusa Yoshimura for assaying the biological activity of rhIL-11.

References