# Cloning and Sequencing of Helper Component Proteinase Gene of a New Isolate of Potato Virus Y in *E. coli.*

**WANG Yuanying**
**Tobacco Research Institute, Chinese Academy of Agricultural Sciences**

**Tennessee,USA**
**Sept. 2008**

# contents

- **INTRODUCTION**
- **MATERIALS AND METHODS**
- **RESULTS**
- **DISCUSSION**

# INTRODUCTION

- The potato virus group is the largest of the 34 known plant virus groups and families.
- The representative of this group is potato virus Y (PVY). PVY can be transmitted non-persistently by aphids and is mainly hazardous to potato and tobacco plants.
- The PVY transmission by aphids depends on two proteins encoded by PVY, including HC-Pro and CP.
- HC-Pro is one of the most multifunctional and complex virus proteins and has played an important role in all aspects of virus life history, such as protease activity, participating in polyproteasome processed, and transmission of the virus by aphids mainly through the N-terminal KITC sequence, C-terminal PTK sequence

# INTRODUCTION

- In addition, HC-Pro also has other functions, such as participating in transmission of the virus in some seeds, systemically infecting the host plant and suppressing gene silencing so as to achieve associative development.
- The viral RNA corresponding to open reading frames, regulatory elements, non-coding sequences or silent mutations encodes a single and large polypeptide.
- In order to understand the virus life cycle and design anti-viral strategy, a novel Potato virus Y strain was isolated from Qingzhou in Shandong Province. The encoding DNA sequence of the HC-Pro region was obtained by RT-PCR and DNA splicing methods.
- HC-Pro was expressed in E. coli , induced by IPTG and purified by electroelution.

aphid(*Myzus persicae*)

Symtom of potato virus Y

# MATERIALS AND METHODS

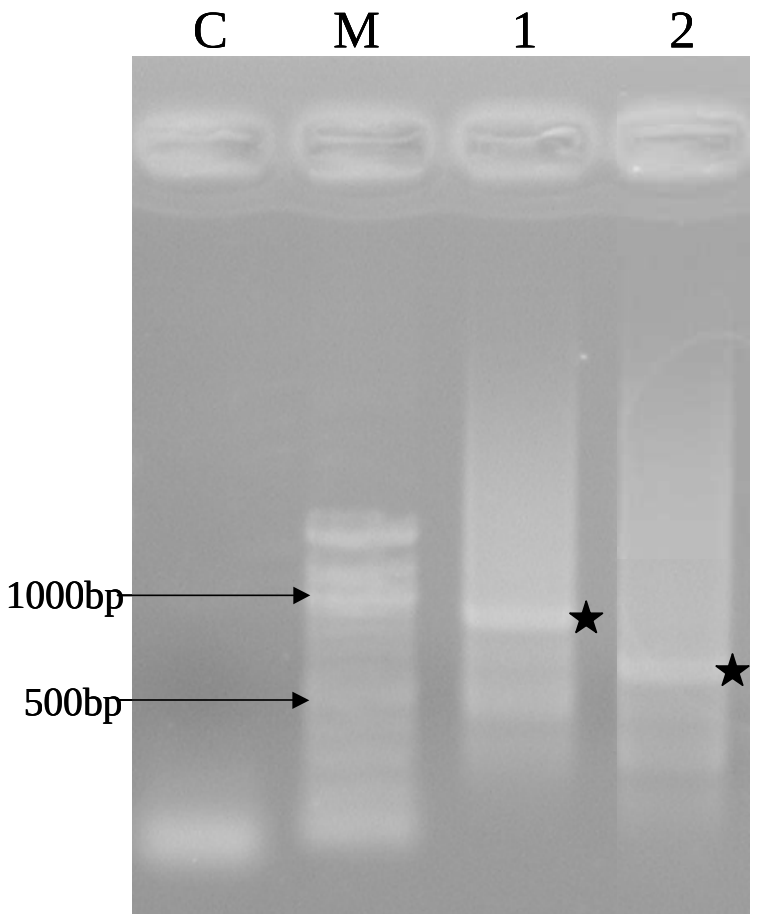
- **Materials and Reagents**
- **Materials:**
- **PVY was isolated by Tobacco Research Institute of CAAS in Qingzhou, Shandong Province.**
- **The host for propagating and preserving PVY was tobacco NC89. E. coli BL21 (D3)**
- **The expression vector pET28a+ were supplied by Plant Protection Institute of CAAS.**
- **Reagents:**
- **dNTPs, T4-DNA ligase, DNA restriction enzyme (NcoI, BsmI and XhoI), 100bp DNA ladder marker were from NEB Company.**
- **The virus RNA extraction kit and PCR purification kit were from Qiagen Company.**
- **RT-PCR reagent kits were AccessQuickTM RT-PCR System of Promega.**

# MATERIALS AND METHODS

- **Methods**
- RNA extraction from the virus
- RT-PCR Amplification of two specific fragments
- Construction of prokaryotic expression vector pET-28a+-HC-Pro
- Prokaryotic expression of HC-Pro in *E.coli* BL21
- Determination of HC-Pro
- Purification of HC-Pro

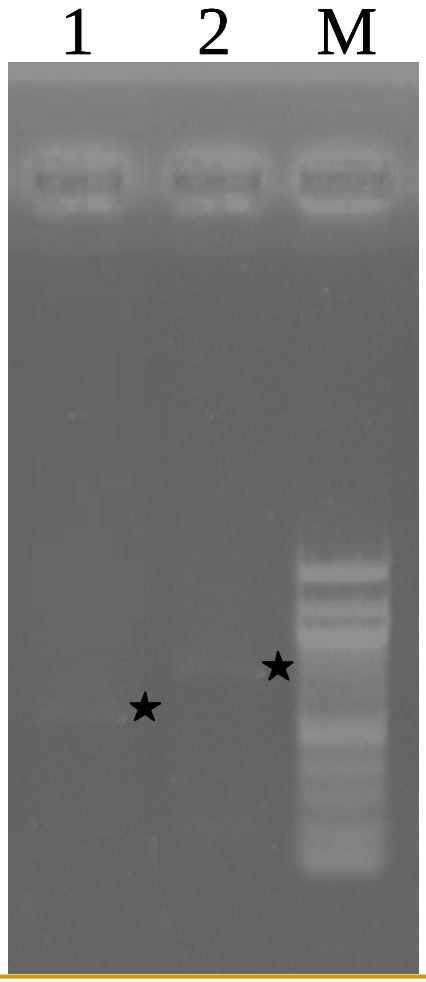
# RESULTS-1

- **RT-PCR amplification of HC-Pro**
- Two fragments of HC-Pro were amplified by RT-PCR.




- **Fig.1 RT-PCR amplification for two fragements.**
- **Lane1: 920 bp fragment amplification (downstream) by RT-PCR;**
- **lane 2: 630 bp fragment amplification (upstream) by RT-PCR. Both of the positive fragments were marked by star. Lane M: 100 bp DNA ladder marker. Lane C: negative control of RT-PCR reaction**

# RESULTS-2



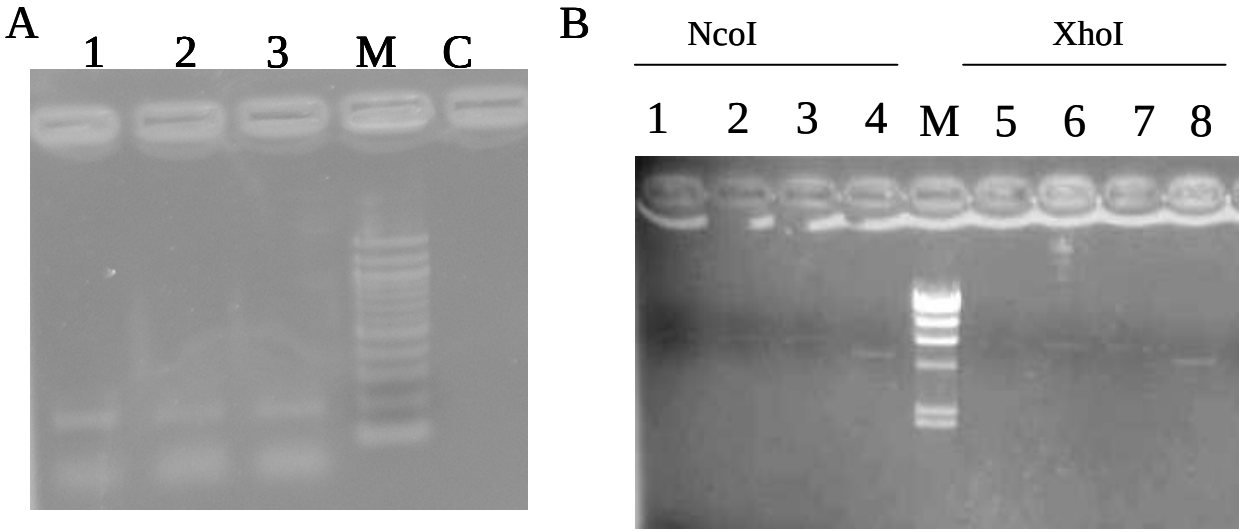

- **Construction of pET-28a+-HC prokaryotic expression vector**
- **The 630 bp fragments were cut with NcoI and BsmI and 920 bp fragments were cut with BsmI and XhoI.**
- **The double-digestion products were separated by 1% agarose gel electrophoresis.**

- **Fig.2 Double digestion of the two fragments.**
- **Lane 1: double digestion of 630 bp fragments.**
- **Lane 2: double digestion of 920 bp fragments. fivestars indicate fragments amplified**
- **Lane M: 100 bp DNA ladder marker**

# RESULTS-2

**The products of double-digestion in Lane 1 and 2 were purified and ligated into the linear vector pET-28a+ and then transformed into *E. coli* BL21 (DE3) competent cells. The clones were detected by PCR with p2 and p3 primers. The results were shown in Figure 3A.**

**The positive recombinant plasmids were extracted by the alkaline lysis method and detected by digestion with NcoI and XhoI. The results were shown in Figure 3B.**




- **Fig.3 The recombinant plasmids were detected by PCR and digestion with NcoI or XhoI .**
- A: **Lane1 to lane 3 are the PCR products amplified with p3 and p2 primers for the 180 bp fragment. Lane M is 100 bp DNA ladder marker. Lane C is negative control of PCR amplification. Fivestar indicates fragments amplified with p3 and p2 primers.**
- **B: Lane1 to lane 3 are the products with NcoI digestion. Lane 4 is the liner pET-28a+ produced with NcoI digestion. Lane 5 to lane 7 are the products with XhoI digestion. Lane 8 is the liner pET-28a+ produced with XhoI digestion. Lane M is λ DNA with Hind III digestion marker. Fivestars indicate fragments digested.**

# RESULTS-3

## The Full-length Sequence Analysis of HC-Pro

```
1    ATG GAT TCA ATG GTT CAG TTC TCA AGC GCT GAA AGC TTT TGG AAG GGA TTG GAC GGC AAT TGG GCA CAA ATG AGA TAT CCT ACA GAT CAT   90
1     M  D  S  M  V  Q  F  S  S  A  E  S  F  W  K  G  L  D  G  N  W  A  Q  M  R  Y  P  T  D  H   30

91   ACA TGT GTG GCA GGC TTA CCA GTT GAA GAC TGT GGC AGA GTT GCA GCG ATA ATG ACA CAC AGT ATT TTA CCG TGC TAC AAG ATA ACC TGC  180
31    T  C  V  A  G  L  P  V  E  D  C  G  R  V  A  A  I  M  T  H  S  I  L  P  C  Y  K  I  T  C   60

181  CCT ACC TGT GCC CAA CAA TAC GCC AAC TTG CCA GCC AGT GAC TTG CTT AAG ATA TTA CAC AAG CAC GCA AGT GAT GGT TTA AAT CGA TTG  270
61    P  T  C  A  Q  Q  Y  A  N  L  P  A  S  D  L  L  K  I  L  H  K  H  A  S  D  G  L  N  R  L   90

271  GGG GCG GAC AAA GAT CGC TTT GTG CAT GTC AAA AAG TTC TTG ACA ATC TTA GAG CAC CTA ACT GAA CCG GTT GAT CTG AGT CTA GAA ATT  360
91    G  A  D  K  D  R  F  V  H  V  K  K  F  L  T  I  L  E  H  L  T  E  P  V  D  L  S  L  E  I  120

361  TTC AAT GAA GTA TTC AAG TCT ATA GGG GAG AAG CAA CAA TCA CCT TTC AAA AAC CTG AAT ATT CTG AAT AAT TTC TTT TTG AAA GGA AAG  450
121   F  N  E  V  F  K  S  I  G  E  K  Q  Q  S  P  F  K  N  L  N  I  L  N  N  F  F  L  K  G  K  150

451  GAA AAT ACA GCT CGT GAA TGG CAG GTG GCT CAA TTA AGC TTA CTT GAA TTG GCG AGA TTC CAA AAG AAC AGA ACG GAT AAC ATC AAG AAA  540
151   E  N  T  A  R  E  W  Q  V  A  Q  L  S  L  L  E  L  A  R  F  Q  K  N  R  T  D  N  I  K  K  180

541  GGC GAC ATC TCG TTC TTT AGG AAT AAA CTA TCT GCC AAA GCA AAT TGG AAC TTG TAT CTG TCA TGT GAT AAC CAG CTG GAT AAG AAT GCA  630
181   G  D  I  S  F  F  R  N  K  L  S  A  K  A  N  W  N  L  Y  L  S  C  D  N  Q  L  D  K  N  A  210

631  AAC TTC CTG TGG GGA CAG AGA GAA TAT CAT GCT AAG CGA TTT TTC TCG AAC TAT TTC GAG GAA ATT GAT CCA GCG AAG GGC TAT TCA GCA  720
211   N  F  L  W  G  Q  R  E  Y  H  A  K  R  F  F  S  N  Y  F  E  E  I  D  P  A  K  G  Y  S  A  240

721  TAT GAA AAT CGT TTG CAT CCG AAT GGG ACA AGA AAG CTT GCA ATT GGA AAC TTA ATT GTT CCA CTT GAT CTG GCT GAG TTT AGG CGG AAG  810
241   Y  E  N  R  L  H  P  N  G  T  R  K  L  A  I  G  N  L  I  V  P  L  D  L  A  E  F  R  R  K  270

811  ATG AAA GGT GAT TAT AAA AGA CAG CCA GGG GTG AGT AAG AAG TGC ACG AGC TCG AAG GAT GGA AAC TAC GTG TAT CCC TGT TGT TGC ACT  900
271   M  K  G  D  Y  K  R  Q  P  G  V  S  K  K  C  T  S  S  K  D  G  N  Y  V  Y  P  C  C  C  T  300

901  ACA CTT GAT GAT GGC TCA GCT GTT GAA TCA ACA TTT TAC CCG CCA ACT AAG AAG CAC CTC GTA ATA GGT AAT AGT GGC GAC CAA AAG TAT  990
301   T  L  D  D  G  S  A  V  E  S  T  F  Y  P  P  T  K  K  H  L  V  I  G  N  S  G  D  Q  K  Y  330

991  GTT GAT CTT CCA AAA GGG AAT TCT GAA ATG TTA TAC ATT GCC AGG CAA GGA TTT TGT TAT ATT AAT ATA TTC CTT GCG ATG TTG ATA AAC 1080
331   V  D  L  P  K  G  N  S  E  M  L  Y  I  A  R  Q  G  F  C  Y  I  N  I  F  L  A  M  L  I  N  360

1081 ATT AGT GAA GAA GAT GCA AAG GAT TTC ACG AAG AAG GTT CGT GAC ATG TGT GTG CCA AAG CTT GGA ACC TGG CCA ACC ATG ATG GAT CTG 1170
361   I  S  E  E  D  A  K  D  F  T  K  K  V  R  D  M  C  V  P  K  L  G  T  W  P  T  M  M  D  L  390

1171 GCT ACT ACT TGT GCT CAA ATG AAA ATA TTT TAT CCT GAT GTT CAT GAT GCA GAG CTG CCC AGA ATA CTA GTC GAC CAT GAA ACT CGG ACA 1260
391   A  T  T  C  A  Q  M  K  I  F  Y  P  D  V  H  D  A  E  L  P  R  I  L  V  D  H  E  T  R  T  420

1261 TGC CAT GTG GTC GAC TCG TTT GGT TCG CAA ACA ACT GGG TAT CAT ATT TTG AAA GCA TCT AGC GTA TCC CAA CTT ATT TTG TTT GCT AAT 1350
421   C  H  V  V  D  S  F  G  S  Q  T  T  G  Y  H  I  L  K  A  S  S  V  S  Q  L  I  L  F  A  N  450

1351 GAT GAG TTG GAG TCT GAC ATT AAG CAC TAT AGA GTT GGT                                              1389
451   D  E  L  E  S  D  I  K  H  Y  R  V  G                                                        463
```

**HC-Pro full-length DNA sequence is 1 389 bp with an ORF encoding a protein of 463 amino acid residues. The molecular mass and pI predicted with BioEdit are 52 kD and 8.10, respectively.**

**Fig.4 The nucleotide and amino acid sequences of HC-Pro.**

# RESULTS-3

- **Tabe1. The alignment of the nucleotide and amino acid sequence of HC-Pro in the research with those originated from some other strains**

| PVY★ | nucleotides(%) | Amino acids(%) | NCBI accession number |
|---|---|---|---|
| PVY NE11 | 98 | 99.8 | DQ157180.1 |
| PVY NTN | 97 | 99.8 | EF026075.1 |
| PVY N | 96.6 | 99.3 | AJ890342.1 |
| PVY O | 96.2 | 99.1 | AJ890349.1 |
| PVY C | 84 | 92.3 | AJ890348.1 |
| TVMV | 59 | 50 | NC_001768 |
| TEV | 43 | 45.5 | NC_001555 |

- **The gene had more than 96% homology with other PVY strain in nucleotide sequences and more than 92% homology in putative amino acid sequences.**

# RESULTS-3

**The conserved domains of HC-Pro were predicted by using NCBI software. HC-Pro posses a C6 peptidase, a zinc finger domain (Figure 5I) which was combined with N-terminal and metal ions, KITC and PTK domains (Figure 5 II and V) which control the transmission by aphids, FRNK domain(Figure 5 III) which is involved in proliferation and infection of the virus. The VRD RNA-binding domain (Figure 5 VI) and CCC domain (Figure 5 IV) which is involved in participation of virus necessary for long-distance transport .**

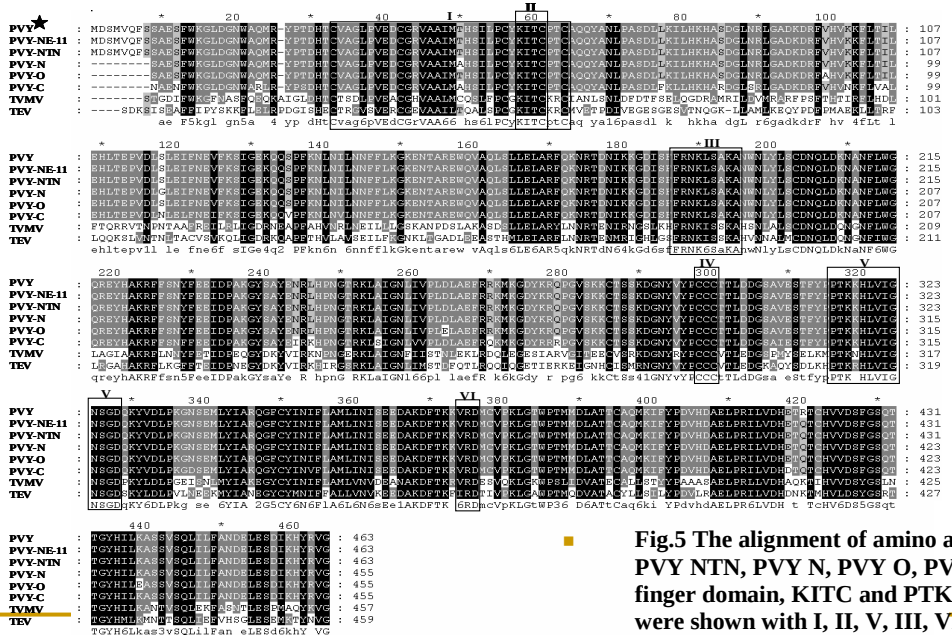



**Fig.5 The alignment of amino acid sequences encoded by HC-Pro genes from PVY NE-11, PVY NTN, PVY N, PVY O, PVY C, TVMV, TEV and PVY★ (this study). The zinc finger domain, KITC and PTK domain, FRNK domain, VRD domain and CCC domain were shown with I, II, V, III, VI and IV, respectively.**

# RESULTS-4

- **Prokaryotic expression of HC-Pro in *E.coli* BL21(DE3)**
- **A 52 KD protein band was induced in accord with the prediction.**

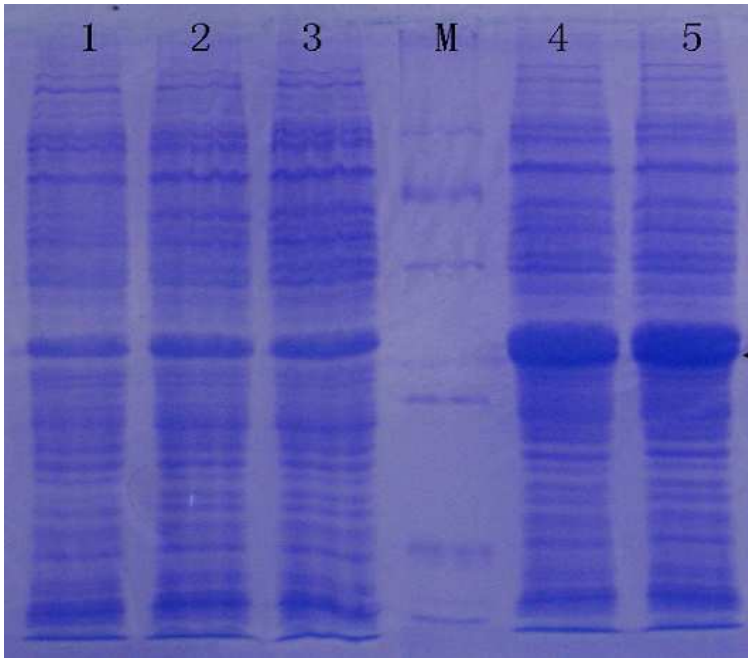



**Fig.6 Prokaryotic expression of HC-Pro in *E.coli* BL21(DE3).**

**Lane 1: total proteins of BL21(DE3).**

**Lane 2 and 3: total proteins of BL21 (DE3) with pET28a+-HC-Pro before the IPTG induction.**

**Lane4 and 5: total proteins of BL21 (DE3) with pET28a+-HC-Pro after the IPTG induction.**

**Lane M: protein molecular weight standards. Fivestar indicates HC-Pro protein expressed in *E.coli* BL21(DE3).**

# RESULTS-5

- **The purification of HC-Pro protein with electroelution:**
- Purification of HC-Pro protein was performed by the electroelution method. A single band was visualized after electroelution (Lane 1) and dialysis (Lane 2).

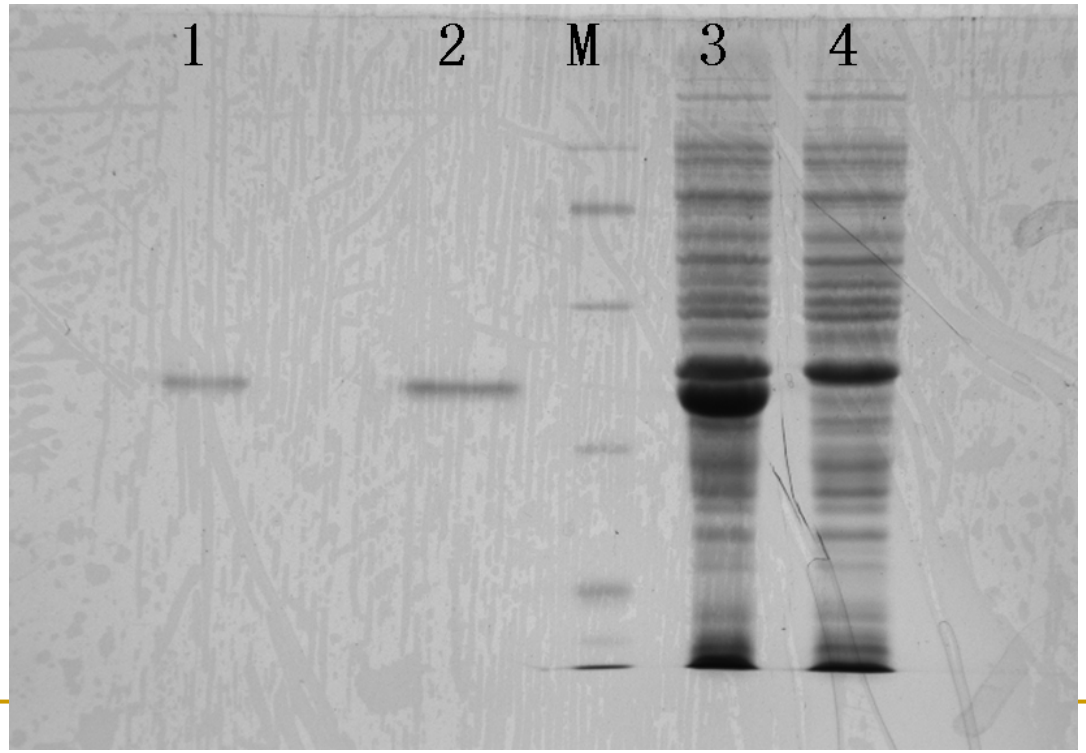



**Fig.7 Purification of HC-Pro protein by electronelution.**

**Lane 1: the HC-Pro protein after electroelution.**

**Lane 2, the HC-Pro protein after dialysis,**

**Lane 3: the total proteins of BL21 (DE3) + pET28a+-HC-Pro after the IPTG induction,**

**Lane 4: the total protein of BL21 (DE3) + pET28a+-HC-Pro before the IPTG induction. Fivestar indicates HC-Pro protein.**

# DISCUSSION

- In this study, we cloned a cDNA of HC-Pro from a novel isolate of PVY. Because the full length of HC-Pro is 1.4 kb approximately, and the DNA length amplified by Taq polymerase was limited to 1 kb, two sets of primers were designed to amplify upstream fragment (p1, p2) and downstream fragment (p3, p4). The full length of HC-Pro cDNA was then cloned by RT-PCR and DNA splicing.
- Positive clones were detected by PCR amplification with primers of p3 and p2 for amplifying the 180 bp overlap fragment between upstream and downstream so as to determine the two fragments that had been recombined correctly with the vector pET28a+.

# DISCUSSION

- According to the alignment of amino acid sequences for HC-Pro protein, the HC-Pro protein is a functional protein.
- To study the biological functions, we prepared the HC-Pro protein by prokaryotic expression.
- The protease activity and other functions, such as participating in polyproteasome processed and transmission of the virus by aphids for the HC-Pro protein are now studied in my lab.

# Thanks for your attention!