UNIVERZITA PALACKÉHO V OLOMOUCI
Přírodovědecká fakulta

DISERTAČNÍ PRÁCE

2011 Jana Rohožková

Univerzita Palackého v Olomouci

Přírodovědecká fakulta

Katedra botaniky

a

Katedra buněčné biologie a genetiky

Jana Rohožková
(roz. Petrusová)

Molekulární determinace interakce rostlina/Potyvirus

Doktorská disertační práce

Studijní obor botanika

Školitel: Prof. RNDr. Milan Navrátil, CSc.

Olomouc

2011

Poďakovanie

Rada by som tymto poďakovala mojmu školitelovi Prof. RNDr. Milanovi Navrátilovi CSc. za jeho odborné vedeni a srdečný prístup. Rovnako by som sa chcela poďakovať Mgr. Daně Šafářové, PhD. za jej pomoc pri riešení problémov, prínosné konzultácie a priateľský prístup počas celého obdobia doktorského studia. Ďakujem aj všetkým členom Katedry buněčné biologie a genetiky, ktorí mi vytvorili príjemné pristredie na prácu. V neposlednom rade ďakujem svojmu manželovi, že to so mnou vydržal.

Prohlášení

Prohlašuji, že jsem tuto disertační práci vypracovala samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů, a že jsem měla hlavní podíl na získání experimentálních dat uvedených v práci. Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona v platném znění.

V Olomouci dne..... podpis

BIBLIOGRAPHICAL IDENTIFICATION

Author's Name: Jana Rohožková (roz. Petrusová)

Title: Molecular determination of interaction plant/Potyvirus

Type of Thesis: Ph.D. Thesis

Department: Botany

Supervisor: Prof. RNDr. Milan Navrátil, CSc.

The Year of Presentation: 2011

Abstract:

Pea seed-borne mosaic virus (PSbMV) a member of the genus Potyvirus, can be transmitted by seeds and causes serious lost in the legume crop yields. Its genome is consisted of 3 peptidases – P1, HC-Pro and NIa. Interestingly only P1 peptidase localized on the N terminal part of the viral polyprotein still remains without any concrete function influencing viral infection. We have expressed truncated part of P1 (Δ33-143) for immunizing of rabbit to produce polyclonal antibody against P1. Using this antibody we have found, that P1 (42 kDa) can act in the truncated form (30 kDa) during viral infection. Its interacting partner is CI. Immunohistochemical localization of P1 in the protoplasts obtained from infected pea plant demonstrated localization of P1 in the cytoplasm as well as in the nucleus. We have described C2H2 zinc finger motif (ZnF), which is non-standard of that presented in the eukaryotic world. We demonstrate *in silico* predicted protein models of found ZnF. After identification of this ZnF we found P1 of other 6 viruses (4 Potyvirus, 2 Tritimovirus) which are containing it too. Predicted protein models are confirming canonical secondary structure of found ZnF.

PSbMV isolates are differed into 4 pathotypes (P1, P2, P3 and P4). This differentiation was done due to the path of viral infection on the pea plants carrying genes of resistance *sbm-1* and *sbm-2*. Viral protein VPg was considered as matching gene of virulence. Viral determinant of virulence corresponding to the interaction with *sbm-2* was mapped as P3-6K1 protein of PSbMV. We analyzed sequences of coding regions for VPg and P3- 6K1 Czech isolates of PSbMV to class found isolates to the pathotype group. Interestingly, we found that isolates are distributed into all three pathotypes with a slight difference in the sequence of PSB204CZ isolate. We have designed novel primers bordering coding region of interest. To confirm the results coming from the molecular experiments we made biological testing.

Keywords: Potyvirus, PSbMV, P1 peptidase, VPg, P3-6K1

Number of Pages: 82

Language: English

BIBLIOGRAFICKÁ IDENTIFIKACE

Jméno: Jana Rohožková

Téma: Molekulární determina interakce rostlina/Potyvirus

Typ práce: Disertační práce

Obor: Botanika

Školitel: Prof. RNDr. Milan Navrátil, CSc.

Rok obhajoby: 2011

Abstrakt:

Pea seed-borne mosaic virus (PSbMV; Virus mozaiky hrachu prenosný semenom) patrí do rodu Potyvirus, je prenosný semenom a spôsobuje vážne straty na úrode strukovín. Genóm tohto vírusu kóduje 3 peptidázy: P1, HC-Pro a NIa. Jedine P1 peptidáza (lokalizovaná na N terminálnom konci polyproteinu) však ostáva bez jasnej funkcie v priebehu vírusovej infekcie. Exprimovali sme skrátenú oblasť P1 (Δ33-143), ktorou sme následne imunizovali králika a vyrobili si tak vlastnú polyklonálnu protilátku proti tomuto peptidu. Imunoprecipitáciou z úplného lyzátu pripraveného z infikovanej rastliny hrachu sme zistili, že P1 (42 kDa) sa môže vyskytovať v priebehu infekcie v skrátenej forme (30 kDa). Pravdepodobný interakčný partner je vírusový CI. Ďalej sme P1 lokalizovali v protoplastoch infikovanej rastliny hrachu imunofluorescenčne a to v cytoplazme a v jadre. Popísali sme motív Cys2His2 zinkového prstu (C2H2 ZnF), ktorý je čiastočne podobný tomu vyskytujúcemu sa u eukryotov, kde bol popísaný. Následne sme navrhli model tohto ZnF v *in silico* podmienkach. Pokúsili sme sa nájsť homológiu tohto motívu i u iných Potyvírusov, čo sa nám podarilo, a to u ďalších 6 vírusov (4 Potyvírusov, 2 Tritimovírusov).

Izoláty PSbMV sa rozdeľujú do 4 patotypov (P1, P2, P3 a P4). Takéto rozdelenie bolo rozhodnuté na základe priebehu vírusovej infekcie na rastlinách hrachu nesúcich gény rezistencie *sbm-1* a *sbm-2*. Vírusová proteín VPg bol nájdený ako gén virulencie odpovedajúci rastlinnému génu *smb-1*. Vírusový determinant virulencie interagujúci s *sbm-2* bol nájdený v oblasti P3-6K1. Cieľom našej práce bolo zaradiť izoláty PSbMV nájdené v Českej Republike do jednotlivých patotypov a to pomocou sekvenovania oblastí VPg a P3- 6K1. Sekvenčná analýza poukázala na distribúciu izolátov do troch patotypov. Ďalej sme postupovali v overení tohto výsledku pomocou biologický testov.

Klíčová slova: Potyvirus, PSbMV, P1 peptidáza, VPg, P3-6K1

Počet stran: 82

Jazyk: Anglický

Abbreviations

aa	amino acid
bp	base pairs
C2H2 ZnF	Cys2His2 Zinc finger
DNA	deoxy-ribonucleic acid
HC-Pro	Potyviral Helper component protein
ICTV	International Committee on Taxonomy of Viruses
IgG	Immunoglobulin G
nt	nucleotides
P1	Potyviral peptidase
P3-6K1	Potyviral gene of virulence matching with *sbm-2* plant gene of resistance
PSbMV	Pea seed-borne mosaic virus
PVY	Potato virus Y
RNA	ribonucleic acid
VPg	Potyviral protein covalently bound to the N terminus of viral RNA
	Potyviral gene of virulence matching with *sbm-1* plant gene of resistance
ZnF	Zinc finger

Summary

1 Background

The *Potyviridae* family (named after its type member Potato virus Y) is the largest of the 65 plant virus groups and families currently recognized (ICTV). Quality and activity explanation of viral proteins in this group, with more than about 100 different species in six gnera (López-Moya and García 2008; Berger *et al.* 2000) is nowadays in the middle of the scientific interest. However was offered to this topic over 30 years there remain still a lot of non-discovered properties connected with viral infectious cycle and its system spreading as well as identification of the "key" to the viral evolution resulting to the host specificity. Within the Potyviridae family, the following genera are recognized: Brambyvirus, Potyvirus, Rymovirus, Macluravirus, Ipomovirus and Tritimovirus. A seventh genus, Bymovirus, contains viruses with bipartite genomes. Recent attempts to expand the genera include Blackyvirus (Susaimuthu *et al.* 2008), Poacevirus (Tatineni *et al.* 2009) and Susmovirus (Xu *et al.* 2010).

The viruses of this family cause significant losses in agricultural, pastoral, horticultural and ornamental crops (Ward and Shukla 1991). Virions are flexuous and rod-shaped, 80 to 900 nm long and 11 to 15 nm wide (Fig.1A). The nucleocapsid is composed of around 2000 units of a single structural protein, surrounding one molecule of nucleic acid (Carrington and Dougherty 1988). A feature shared by all of the potyviruses is the induction of characteristic pinwheel or scroll-shaped inclusion bodies in the cytoplasm of infected cells (Fig.1B) (Edwardson 1974). The predominant way of transmission of potyviruses is through aphids. Other possible means of the virus' spread are by mites and whiteflies. The genome is created by single-stranded RNA of positive sense, around 10 000 bp long. In most potyvirid species with a monopartite genome, their genomic sequence encodes nine multifunctional proteins plus the capsid protein.

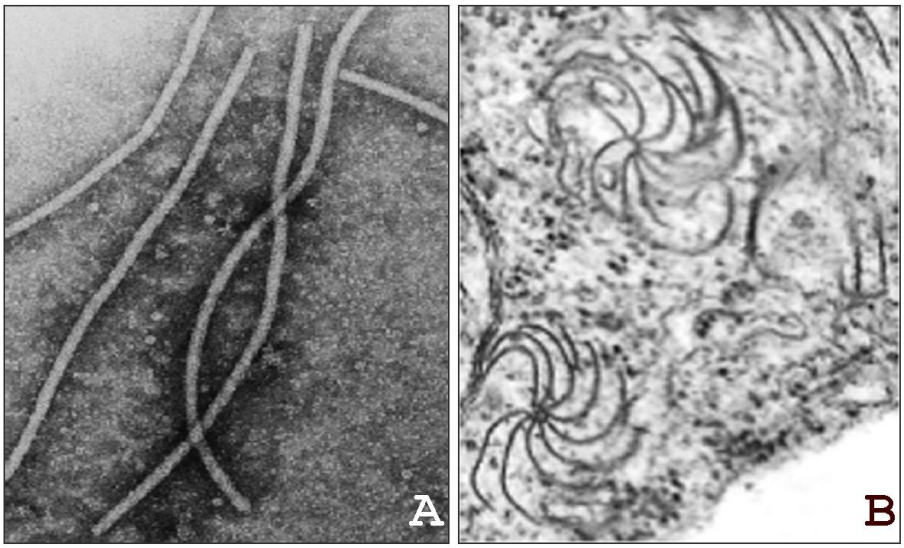


Fig.1 Electron micrograph of potyvirid rod shaped virion (A) and characteristic pinwheel inclusion bodies in the cytoplasm of infected cells (B). *Photo by JA García, Centro Nacional de Biotecnología-CSIC, Madrid,*

Pea seed-borne mosaic virus (PSbMV) is economically important seed transmitted virus infecting pea plants (most cultivars of *Pisum sativum*), lentil, chickpea and broad bean. Infection by PSbMV belongs to the group of important legume diseases and causes stunting, downward rolling of leaflets, swelling of leaf veins (Fig.2) and delayed flowering as well as serious crop lost varying between 11-82% (Torok 2006). The virus is transmitted by aphids in non-persistent manner and vertically through seeds (Hampton *et al.* 1973). The PSbMV genome, as well as most of the *Potyviridae* members, is composed from ssRNA of messenger-polarity. Size of genome varies according to isolate between 9342 and 10820 nt including a 3´-poly(A) tail and a virus-encoded protein (VPg) covalently bound on the 5´-end (Kekarainen *et al.* 1999, Oruetxebaria *et al.* 2001). Viral genome encodes one polyprotein, which is finally processed by its three proteases P1, HC-Pro and NIa (Dougherty and Semler, 1993) to 9 non-structural and 1 capside protein (Fig.3).

Fig.2 Symptoms of PSbMV virus infection on the pea plant (*cultivar Merkur*) presented by stunting, downward rolling of leaflets, swelling of leaf veins (*right*) in comparison to the healthy plant (*left*) (A) and serious crop lost varying between 11-82% compared to the yield from the healthy (*left*) and infected (*right*) pea plant (B). *Photo by Dana Šafářová, Dept. of cell biology and genetics, UP Olomouc*

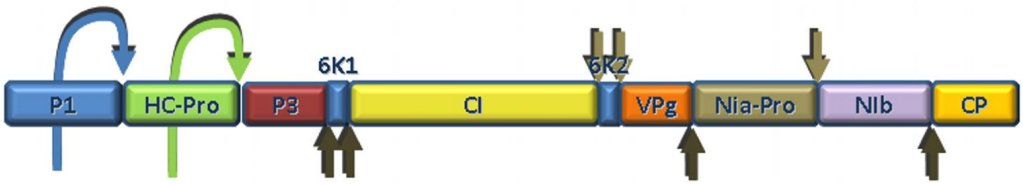


Fig.3 Scheme of Potyviral proteins with signed proteolytical Cleavage of the proteins from the viral polyprotein. P1 pepteidase and HC-Pro cleaves itself out of the viral polyprotein. NIa-Pro cleaves itself from NIb and 6K2 (*hell brown*). As the last are released 6K1 from CI, NIa-Pro from VPg and CP (*dark brown*).

PSbMV isolates are recognized as pathotypes P-1, P-2, and P-4 by their infection profiles on a panel of *Pisum sativum* lines (Alconero *et al.* 1986; Kasimor *et al.* 1997). Correspondingly, *P. sativum* lines can be divided into categories A, B, C, and D, depending on their susceptibility to the PSbMV pathotypes (Johansen *et al.* 2001). Category A lines are susceptible to all three pathotypes, category B lines are resistant to all three pathotypes, category C lines are resistant only to pathotype P-2, and category D lines are resistant to pathotypes P-1 and P-2. Thus, categories C and D are sufficient to differentiate the three known pathotypes. Resistance to PSbMV in *P. sativum* is inherited as single recessive genes, designated *sbm-2* and *sbm-1* in category C and D lines, respectively (Johansen *et al.* 2001; Provvidenti and Alconero 1988).

1.1 P1 peptidase

P1 peptidase was the last of the three peptidases to have been identified as processing the potyviral polyprotein (Verchot *et al.* 1991). P1 peptidase, the first protein of the viral polyprotein, is the most divergent protein with regard to both length and amino acid sequence (Adams *et al.* 2005). Its size between potyviruses varies in 30-83 kDa. Its role in virus infection cycle has remained unclear. There are just hints regarding P1 successful adaptation of the potyviruses into a wide range of host species (Valli *et al.* 2007). P1 peptidase belongs to family serine-like peptidases and catalyses cleaving itself from viral helper component (HCPro; Mavankal and Rhoads 1991, Verchot *et al.* 1991, 1992). Highly conservative protease domain is identified in C-terminal part of coding region 143AA-END. Serine catalytic triad is composed of His-307, Asp-331 and Ser-348 (*by P1 PSbMV isolate DPD1, GenBank NP734419*) (Verchot *et al.* 1992). Specific motif surrounding putative active-site Ser residue (Gly-Xaa-**Ser**-Gly) is recognized by all examined potyviruses (Dougherty and Semler 1993). Mutational analysis knocking out Ser (*bold in formula*) led to proteolyticaly non-functional P1 proteinase (Verchot *et al.* 1995).

In the microbial world, peptidases are not uncommon. Serine peptidases were found in various viruses (Rawlings and Barrett 1993). They have acquired their importance because of their functional involvement in the processing of proteins of viruses that cause certain fatal diseases such as AIDS and cancer. All virus-encoded proteases are endopeptidases. The mature protease is released by the autolysis of the precursor (Rao *et al.* 1998). Serine proteases are characterized by the presence of a serine group within their active site. The numerous examples of serine proteases among viruses suggest that they are vital to the organisms (Barret 1994). With the intent of preventing serious losses to several major crop plants, strong efforts continue into research on the viral

Till present there have been done a lot of studies trying to explain P1 quality as well as its activity during viral infection. Due to the expanding study and conclusive overview I would like to refer the review Rohožková and Navrátil (2011; *see Chapter 6.1: Research articles and reviews*) in which we would like to offer comprehensive view on the P1 peptidase topic through viruses in family *Potyviridae*.

In present, scientific groups offer high effort in decoding the true processes of viral proteins during infection, breaking the defense mechanism of host organism, cell as well as system spreading with focus on concrete host endogenous interacting proteins and molecular

pathways. We considered the P1 peptidase exploration as very useful not just in way for expanding known facts about plant viruses, but also with the wider application of those viruses concerning and affecting humans.

1.2 VPg and P3-6K1

PSbMV isolates are differed into 4 pathotypes (P1, P2, P3 and P4). This differentiation was done due to the path of viral infection on the pea plants carrying genes of resistance *sbm-1* and *sbm-2* (Alconero 1986). Both genes are active on the level of active barring of virus local and systemic spread. Gene of resistance sbm-1 was identified as homolog of eukaryotic tranclational initiation factor (eIF4E) carrying few mutations and signed eIF(*iso*)4E (Gao *et al.* 2004). Viral protein VPg was considered as matching gene of virulence (Hjulsager *et al.* 2006). During viral infectious cycle has VPg role in a long-distance movement, cell-to-cell movement, and replication. This protein is covalently bound to the 5´ end of viral RNA and acts as a primer for annealing of viral RNA replicase and other cell translational factors (Rajamäky and Valkonen 2009).

Viral determinant of virulence corresponding to the interaction with sbm-1 was mapped as P3-6K1 protein of PSbMV (Johansen *et al.* 2001). The P3-6K1 cistron has been shown genetically to be strictly essential for productive potyvirus infection at the cellular level (Klein *et al.* 1994). However, a corresponding molecular function of the P3-6K1 gene product(s) still remains elusive (Urcuqui-Inchima *et al.* 2001). Our experiments were dealing with molecular and biological characterization of the PSbMV isolates obtained in the Czech Republic.

2 Aims of study

We have focused our effort on the two different protein of PSbMV:

1. P1 peptidase and expanding the present knowledge of novel proteomic, biochemical and evolutional data;

2. Sequential and phylogenetic analysis of coding regions P3 and 6K1, which are interacting regions with the genes of resistance in host plants.

2.1 P1 peptidase

For discovering proteomic details about P1 peptidase of PSbMV we tried to fulfill these aims:

- *In vitro* expression of P1 peptidase of PSbMV selected isolate and preparation of polyclonal antibody against it
- Localization of P1 in infected pea plant protoplasts using immunohistochemical methods
- Identification of proteins specifically interact with P1 peptidase during viral infection
- Prediction of P1 peptidase active domains and *in silico* modeling of the secondary structure

2.2 P3+6K1 and VpG coding regions interacting with the genes of resistance of host plant

Second part of the study is dealing with phylogenetic analysis of VpG and P3+6K1 coding regions in PSbMV selected isolates and we tried to diversify the pathotype between Czech found isolates fulfilling following aims:

- Sequential analysis of VPg and P3+6K1 coding region of 11 known isolates of PSbMV identified in Czech Republic
- Diversification of isolates to their biological characteristics (pathotype) and confirmation of found data by biological testing on different pea cultivars.

2.1 P1 peptidase

2.1.1 *In vitro* expression of P1 peptidase of PSbMV selected isolate and preparation of polyclonal antibody against it.

For *in vitro* expression of P1 peptidase we tried 3 different strategies: 1. Expression of whole P1 protein, expressing of peptides; 2. Expression of peptides localized in the very variable N terminus (P1/I in position 15-98 aa) and in contrary short peptide localized in the middle of the conserved part of the P1 peptidase (P1/II in position 146-215 aa); 3. Expression of one peptide (110 aa) fulfilling criteria of surface display predicted to the secondary structure topology of P1 and immunogenicity of reached fragment. The third strategy was regarded as the most successful. Shortened peptide P1 Δ33-143 expressed in prokaryotic expression system pQE30. This peptide was further used for rabbit immunization with regard to production of polyclonal IgG.

Produced antibody was further tested on WB of total cell lyzate prepared form infected pea plant. We have also immunoprecipitated P1 peptidase via this polyclonal antibody from infected plant lyzate. In both cases we have detected protein of 30 kDa. Protein sequencing confirmed that aa of it covers P1 coding region between 80-277 aa (further signed as ΔP1). We can suggest, that there is undergoing alternative splicing of P1 after its proteolytic cleavage from the viral polyprotein.

2.1.2 Localization of P1 in infected pea plant protoplasts by immunofluorescence

We were able to detect P1 peptidase in WB from infected pea plant total cell lyzate. However we detected also non-specific interaction of this polyclonal antibody with Rubisco. This was maintained as problem for further localization of P1 in the cell, whereas non-specific interaction with Rubisco enables to identify specific localization pattern. To solve this problem we used specific antibody against Rubisco, which blocks epitopes possibly interacting with rabbit polyclonal immunoglobulin. Result was satisfying with obvious difference between non-blocked and blocked samples.

Localization of P1 in protoplasts isolated from infected pea plant using immunofluorescence after blocking Rubisco was successful. We detected this viral peptidase spread all over the cytoplasm but excluded from the cytoplasmic membrane and other membranous structures as endoplasmic reticulum and Golgi apparatus. Clearly positive localization was detected in the cell nucleus. We haven't localized P1 in the intact plant tissues so we can't say anything about localization of P1 in the interstitial spaces between the plant cells.

2.1.3 Identification of protein specifically interact with P1 peptidase during viral infection

To clarify the processes in which is 30kDa P1 peptidase involved we used method co-immunoprecipitation which enables to pull down P1 peptidase and proteins interacting with it during infectious cycle. In comparison to the negative control with ΔP1 (30 kDa) we have precipitated viral CI protein (71 kDa) and other 48 kDa un-known protein. Found interaction with CI can be explained by two suggested theories: 1) All viral proteins are synthesized near cytoplasmic inclusions; 2) Cytoplasmic inclusions are sites of potyviral RNA replication. P1-CI can be expanded by the hypothesis of Verchot and Carrington (1995) that P1 peptidase could participate in virus replication or regulatory factor stimulating genome amplification.

2.1.4 Prediction of P1 peptidase active domains and *in silico* modeling of the secondary structure

We analyzed aa sequence in region 1-277 which confirms the 30 kDa predicted size and analyzed it by SMART: Sequence analysis database (Schultz et al. 1998, Letunic *et al.* 2009). There was predicted Zinc finger (ZnF) with the specific motif Cys2His2 (C2H2) localized in the N-terminal part of the ΔP1. This motif was in the P1 coding region organized in the formula **C[32]**-X_2-**C[35]**-X_3-F-X_{10}-**H[50]**-X_{20}-**H[70]**. C2H2 ZnF´s are the most common DNA-binding motifs found in eukaryotic transcription factors but importantly have also been identified in prokaryotes (Chou *et al.* 1998).

To complete this novel found fact we prepared *in silico* models of entire ΔP1 with proper folding of the C2H2 ZnF cavity.

ZnF was in the Potyviruses predicted for the first time. Also Vali *et al.* (2007) has previously described some phylogenetically conserved Cys and His in the N-terminal part of a few Potyviral P1s without explanation of this phenomenon (*see supplementary figures of Valli et al. 2007*).

We search for homology in the family *Potyviridae* and we found the same conserved motif in next 4 potyviruses (*Pea seed-borne mosaic virus* isolate *DPD1*, *Onion yellow dwarf virus*, *Leek yellow stripe virus* and *Zucchini yellow mosaic virus*) and 2 tritimoviruses (*Brome streak mosaic virus* and *Oat necrotic mottle virus*) (Table 1).

Mention about C2H2 ZnF in P1 peptidase of virus in family Potyviridae was dedicated for the first time. We have found homologous motif in other members of this family. Furthermore we perform *in silico* prediction to confirm secondary structure of the ZnF cavity. Quality of this motif as well as confirmation of the secondary structure of this motif by crystallization requires further exploration.

25	DADGEYRCT	----	QCDMGFD	40	44	MARVNHCCDG	54	66	DPIMHLV	-	DSK	75	PSbMV DPD1
25	DADGEYRCT	----	QCDMGFD	40	44	MARVNHCCDG	54	66	DPIMHLV	-	DSK	75	PSbMV 204
28	DDAVHYHCT	----	KCNFAFE	43	47	MVRVNHDCDG	57	83	DVFRHLANDNA			93	OYDV
77	WDDDVYECT	----	TCSGAFQ	92	95	LDFKEHDCDE	104	161	APIIHECQQEL			171	LYSV
128	IAAQLYMCP	----	KCCSASD	143	147	YFDTNHNDSC	156	207	TVPVHDVVQVY			217	BStMV
78	LAADVFVCG	----	MCRSSCA	93	89	RSSCAHYRYF	98	98	FIEDHFACEKL			108	ONMV
25	SNRVNIVCPGHMATCPPPKT			45	41	PPPKTHTYYR	50	47	TYYRHESKKLM			57	ZYMV

Tab. 1 Alignment of identified conserved C2H2 ZnF motifs (*red box*) in PSbMV isolates 204 and DPD1 and found P1 coding regions of other member of family *Potyviridae*. Numbers in front and behind the sequence are remarking the position of following aa in the sequence of each P1 peptidase.

2.2 VPg and P3-6K1

2.2.1 Sequential analysis of VPg and P3+6K1 coding region of 11 known isolates of PSbMV identified in Czech Republic

We analyzed sequences of coding regions for VPg and P3- 6K1 Czech isolates of PSbMV to class found isolates to the pathotype group. Interestingly, we found that isolates are distributed into all three pathotypes with a slight difference in the sequence of PSB204CZ isolate. We have designed novel primers bordering coding region of interest. Amplified PCR products were sequenced and resulting sequences were aligned to the sequences of reference PSbMV isolates representing each pathotype in detail DPD1for P1 (*GenBank No D10930*), L1 for P2 (*GenBank No AJ252242*) and NY for P4 (*GenBank No X89997*). To confirm the results coming from the molecular experiments we made biological testing.

2.2.2 Diversification of isolates by biological testing on different pea cultivars.

Results reached from sequential analysis were confirmed by biological testing on different pea cultivars. Surprisingly Czech isolates were distributed into two pathotypes, the P-1 (7 isolates) and P-4 (1 isolate) (*for detailes see Chapter 6.1: Research articles and reviews*; Šafářová *et al.* 2008).

3 Methods

Detailed description of methods is listed in each publication. For reached results just standard laboratory methods were used if it is not mentioned below.

4 Conclusion

We have successfully prepared rabbit polyclonal antibody against viral P1 peptidase. We found out that this viral protein is acting during infectious cycle in the truncated form (30 kDa). During infectious cycle it is possibly interacting with viral CI protein and other unknown 48 kDa protein. For its activity found C2H2 ZnF can be crucial. We described this motif for the first time in the viral world. We have designed model of secondary structure confirming building the ZnF cavity. All of the mentioned results are novel found. Reached results are in detail described and prepared in the manuscript for publication Rohožková *et al.* 2011 (*see Chapter 6.1: Research articles and reviews*).

In the second part of the PhD thesis we assessed PSbMV isolates found in the Czech Republic into two groups of pathotypes- P1 and P2; regarding the results from molecular as well as biological experiments.

5 References

Adams MJ, Antoniw JF, Fauquet CM 2005 Molecular criteria for genus and species discrimination within the family *Potyviridae*. Arch Virol 150:459–579

Alconero R, Provvidenti R, Gonsales D 1986 Three Pea seed-borne mosaic virus pathotypes from pea and lentil germ plasma. Plant Dis 70:783–786

Barret AJ 1994 Proteolytic enzymes: serine and cystein peptidases. Method Enzymol 244:1–15

Berger PH, Banett OW, Brunt AA, Colinet D, Edwardson JR, Hammond J, Hill J, Jordan RL, Kashiwazaki S, *et al.* 2000 Family *Potyviridae*; in Virus taxonomy: Seventh report of International Committee on Taxonomy of Viruses (Academic Press) pp 703–724

Carrington JC and Dougherty WG 1988 A viral cleavage site cassette: identification of amino acid sequences required for *Tobacco etch virus* polyprotein processing. Proc Natl Acad Sci USA 85:3391–3395

Chou AY, Archdeacon J, Kado CI 1998 *Agrobacterium* transcriptional regulator Ros is prokaryotic zinc finger protein that regulates the plant oncogene. Proc Natl Acad Sci USA 95:5293–5298

Dougherty WG and Semler BL 1993 Viral proteinases. Microbiol Rev 57:781–822

Edwardson JR 1974 Some properties of the potato virus Y-group. Fla Agric Exp Stn Monogr Ser 4:225

Gao Z, Johansen E, Eyers S, Thomas CL, Noel Elis TH, Maule AJ 2004 The potyvirus recessive resistance gene, *sbm-1*, identifies a novel role for translation initiation factor eIF4E in cell-to-cell trafficking. Plant J 40:376–385

Hampton RO, Phillips S, Knesek JE, Mink GI 1973 Ultrastructural cytology of pea leaves and roots infected by *Pea seed-borne mosaic virus*. Arch Gesamte Virusforsch 42:242-53

Hjulsager ChK, Olsen BS, Jensen DMK, Cordea MI, Krath BN, Johansen E, Lund OS 2006 Multiple determinants in the coding region of *Pea seed-borne mosaic virus* P3 and involved in virulence against *sbm*-2 resistance. Virology 355:52–61

Johansen IE, Lund OS, Hjulsager CK, Laursen J 2001 Recessive resistance in *Pisum sativum* and Potyvirus pathotype resolved in a gene-for cistron correspondence between host and virus. J Virol 75:6614–6669

Kasimor K, Baggett JR, Hampton RO 1997 Pea cultivar susceptibility and inheritance of resistance to the Lentil strain (Pathotype 2) of *Pea seed-borne mosaic virus*. J A Soc Hortic Sci 122:325–328

Kekarainen T, Merits A, Oruetcebarria I, Rajamäki M, Valkonen JPT 1999 Comparison of the complete sequences of five different isolates of *Potato virus A* (PVA), genus Potyvirus. Arch Virol 144:2355–2366

Klein PG, Klein RR, Rodriguez-Cerezo E 1994 Mutational analysis of the *Tobacco vein mottling virus* genome. Virology 204:759–769

Letunic I, Doerks T, Bork P 2009 SMART 6: recent updates and new developments. Nucleic Acids Res 37:229–232

López-Moya JJ and García JA 2008 Potyviruses; in Encyclopedia of virology (eds.) BWJ Mahy and MHV Van Regenmortel (Oxford: Elsevier) pp 313–322

Mavankal G and Rhoads RE 1991 *in vitro* cleavage at or near the n-terminus of the helper component protein in the *Tobacco vein mottling virus* polyprotein. Virology 185:721–731

Oruetxebarria I and Valkonen JPT 2001 Analysis of the P1 gene sequences and the 3´terminal sequences and secondary structures of the single-stranded RNA genome of *Potato Virus V*. Virus Genes 23:335–343

Provvidenti R, Alconero R 1988 Inheritance of resistance to a lentil strain of *Pea seed-borne mosaic virus* in *Pisum sativum*. J Hered 79:45–47

Rajamäki ML, Valkonen JP 2009 Control of nuclear and nucleolar localization of nuclear inclusion protein a of picorna-like Potato virus A in Nicotiana species. Plant Cell 21:2485–2502

Rao MB, Tanksale AM, Ghatge MS, Deshpande VV 1998 Molecular and biotechnological aspects of microbial proteases. Microbiol Mol Biol Rev 62:597–635

Rohožková J, Navrátil M 2011 P1 peptidase-a mysterious protein of family *Potyviridae*. J Biosci 36:189–200

Rawlings NS and Barrett AJ 1993 Evolutionary families of the peptidases. Biochem J 290:205–218

Schultz J, Milpetz F, Bork P, Ponting CP 1998 SMART, a simple modular architecture research tool: identification of signaling domains. Proc Nat Acad Sci USA 95:5857–5864

Sharp PM, Li WH 1987 The codon Adaptation Index-a measure of directional synonymous codon usage bias, and its potential applications. Nucleic Acids Res 15:1281–1295

Susaimuthu J, Tzanetakis IE, Gergerich RC, Martin RR 2008 A member of a new genus in the *Potyviridae* infects Rubus. Virus Res 131:145–151

Torok VA 2006 Discriminating Between Isolates of PSbMV Using Nucleotide Sequence Polymorphisms in the HC-Pro Coding Region. Plant Dis 91:490–496

Tatineni S, Ziems AD, Wegulo SN, French R 2009 *Triticum mosaic virus*: a distinct member of a family *Potyviridae* with an unusually long leader sequence. Phytopathology 99:943–950

Urcuqui-Inchima S, Haenni AL, Bernardi F 2001 Potyvirus proteins: a wealth of functions. Virus Res 74:157–175

Valli A, López-Moya JJ, García JA 2007 Recombination and gene duplication in the evolutionary diversification of P1 proteins in the family *Potyviridae*. J Gen Virol 88:1016–1028

Verchot JM and Carrington JC 1995 Debilitation of plant potyvirus infectivity by P1 proteinase-inactivating mutations and restoration by second-site modifications. J Virol 69:1582–1590

Verchot JM, Herndon KL and Carrington JC 1992 Mutational analysis of the Tobacco Etch Potyviral 35-kDa proteinase: Identification of essential residues and requirements for autoproteolysis. Virology 190:298–306

Verchot JE, Koonin V and Carrington JC 1991 The 35 kDa protein from the N-terminus of the potyviral polyprotein functions as a third virus-encoded proteinase. Virology 185:527–535

Ward CW and Shukla DD 1991 Taxonomy of potyviruses: current problems and some solutions. Intervirology 32:269–296

Xu DL, Zhou DH, Xie YJ, Mock R and Li R 2010 Complete nucleotide sequence and taxonomy of *Sugarcane streak mosaic virus*, member of a novel genus in the family Potyviridae. Virus Genes 40:432–439

6 Publications

6.1 Research articles and reviews

Šafářová D, Navrátil M, **Petrusová J**, Pokorný R, Piaková Z 2008 *Genetic and biological diversification of the Pea seed-borne mosaic virus isolates occurring in Czech Republic.* Acta Virol 52:53-57

Rohožková J, Navrátil M 2011 *P1 peptidase-a mysterious protein of family Potyviridae.* J Biosci 36:189-200.

Rohožková J, Navrátil M, Šebela 2011 Novel biochemical properties of P1 peptidase of Pea seed-borne mosaic virus. Virol J. *Manuscript prepared to submission in:* Virology Journal, c/o BioMed Central, United Kingdom

Šafářová D, Navrátil M, **Petrusová J**, Pokorný R, Piaková Z 2008 *Genetic and biological diversification of the Pea seed-borne mosaic virus isolates occurring in Czech Republic.* Acta Virol 52:53-57

GENETIC AND BIOLOGICAL DIVERSITY OF THE PEA SEED-BORNE MOSAIC VIRUS ISOLATES OCCURRING IN CZECH REPUBLIC

D. ŠAFÁŘOVÁ[1], M. NAVRÁTIL[1], J. PETRUSOVÁ[1], R. POKORNÝ[2], Z. PIÁKOVÁ[3]

[1]Faculty of Sciences, Palacký University, Olomouc, Šlechtitelů 11, 783 71 Olomouc-Holice, Czech Republic; [2]Faculty of Agronomy, Mendel University of Agriculture and Forestry, Brno, Czech republic; [3]Research Institute for Fodder Crops, Troubsko, Czech Republic

Received November 26, 2007; accepted February 28, 2008

Summary. – Eight isolates of the Pea seed-borne mosaic virus (PSbMV) from the Czech Republic were studied regarding their biological and molecular characteristics. Molecular characterization using RT-PCR was done on the 5'(Nter)NIb-CP-UTR3' region amplified using universal CPUP/P9502 primer pair and the newly designed PSB8812/PSB944, and PSB8800/PSB9440 primer pairs, respectively. Sequential and phylogenetic analysis of CP-UTR3' region from all isolates showed that the available Czech and GenBank PSbMV isolates were distributed into 4 clusters in agreement with their diversification and according to their biological characteristics (i.e. pathotype). The molecular data were confirmed by biological testing on different pea cultivars. The Czech isolates were distributed into two pathotypes, the P-1 (7 isolates) and P-4 (1 isolate).

Key words: Pea seed-borne mosaic virus; pathotypes; phylogenetic analysis; sequencing

Introduction

PSbMV is a member of the family *Potyviridae*. This virus causes stunting, downward rolling of leaflets, swelling of leaf veins, and delayed flowering of most cultivars of *Pisum sativum*. Besides the pea, other economically important host plants for PSbMV are lentil, chickpea, and broad bean. The virus is transmitted in a non-persistent manner by aphids and vertically through seeds (Hampton and Mink, 1975). PSbMV was discovered in former Czechoslovakia (Musil, 1966), but probably is distributed worldwide due to a dissemination from germplasm collection, breeding lines, and commercial cultivar seeds (Grünwald *et al.*, 2004). The resistance of pea to transmission of PSbMV was described and four recessive resistance genes were identified allowing the identification and grouping of viral isolates into the pathotypes. Standardized pathotypization was developed based on the reaction of the differential host pea genotypes (Alconero *et al.*, 1986; Hjulsager *et al.*, 2002). The gene *sbm-1* confers resistance to the pathotype P-1 represented by the pea isolates US or DPD1; *sbm-2* and *sbm-3* genes confer resistance to the pathotype P-2 represented by the lentil isolate L1 and *sbm-4* gene confers resistance to the pathotype P-4 represented by the pea isolates NY and S6 (Provvidenti and Alconero, 1988; Johansen *et al.*, 1991). The broad bean isolate NEP-1 is considered as the pathotype P-3 based on its biological characterization (Hjulsager *et al.*, 2002). The P-1 pathotype seems to be prevalent in Europe, North America, Australia, New Zealand, Pakistan and P-4 pathotype is found in Australia, North America, and Pakistan (Alconero *et al.*, 1986; Timmerman *et al.*, 1990; Johansen *et al.*, 1991; Ali and Randles, 1997). The geographical distribution of P2 and P3 pathotypes is not known. Both pathotypes are represented by unique members, e.g. P-3 is represented by NEP-1 isolated from the seed material originating from Nepal and P-2 is represented by isolate L1 from the germplasm accession (Alconero *et al.*, 1986; Hjulsager *et al.*, 2002). Since the PSbMV is

E-mail: dana.safarova@upol.cz; fax: +4205-85634905.

Abbreviations: AMV = Alfalfa mosaic virus; BCMV = Bean common mosaic virus; PEMV-1 = Pea enation mosaic virus 1; PSbMV = Pea seed-borne mosaic virus; CP = coat protein; NIb = nuclear inclusion b protein; UTR = untranslated region

transmissible by seeds, this propagation material represents an important phytosanitary risk.

This study reports molecular and biological characterization of the PSbMV isolates obtained in Czech Republic with recognition of their pathotypes.

Materials and Methods

Virus isolates. The PSbMV isolates were obtained from naturally-infected pea plants from three different localities Smržice, Šumperk, and Troubsko (Table 1) during the year 2003–2005. The presence of the PSbMV in plants as well as the co-infection with Pea enation mosaic virus 1 (PEMV-1), Alfalfa mosaic virus (AMV), and Bean common mosaic virus (BCMV) was confirmed by DAS-ELISA (Loewe Biochemica). The PSbMV isolates were mechanically transmitted and maintained on the pea cultivar Merkur.

RT-PCR. Total RNA was extracted from 50 mg of fresh pea leaves using RNeasy Plant Mini Kit (Qiagen). The reverse transcription was carried-out in two steps, in total reaction volume of 40 µl. The primer annealing mixture contained 5 µl of total RNA, 0.4 µmol/l oligo (T)18 primer and deionized water in total volume of 15 µl. The mixture was denatured at 70°C for 5 mins and chilled on ice. The annealed mixture was completed with AMV RT 5x buffer, 0.4 mmol/l dNTPs, 20 U RNasin® RNase Inhibitor (Promega), 5 U AMV reverse transcriptase (Promega), and deionized water. The RT was run at 42°C for 60 mins.

The two overlapping fragments covering the 5'(Nter) NIb-CP-UTR3' region (1108 bp and 1115 bp lengths, respectively) were obtained with the three different primer pair combinations. The first fragment was amplified using the potyviral universal primer pair P9502 and CPUP (van der Vlugt and Bouwen, 1997). For amplification of the second part of the region, the two combinations of newly designed primers, PSB8812 (5'-TTGAGAAATA CACGGAAGC-3') and PSB9440 (5'-CATTATCTGTCTGAA AGTTGG-3'); or primer PSB8800 (5'-ACAAAGTGAATTAGA AAGGTA-3') and PSB9440, were used. The PCR conditions involved pre-denaturation 94°C for 2 mins, 40 cycles of amplification (94°C for 60 secs, 52°C for 60 secs, and 72°C for 60 secs), and a final extension at 72°C for 10 mins. The PCR amplifications were carried out in mixture that consisted of 5 µl cDNA, RedTaq (1x) buffer, 1.5 mmol/l $MgCl_2$, 0.625 mmol/l dNTPs, 0.5 µmol/l sense and antisense primer, 1 U RedTaq Polymerase (Sigma), and deionized water to the total volume of 25 µl. The 5 µl of PCR products were analyzed by 2% agarose gel electrophoresis in a standard manner.

Cloning and sequencing. The PCR products were purified using a Gel extraction kit (Qiagen), cloned into a pGEM-T plasmid and propagated in JM109 cells (Promega). Three clones of each isolate were sequenced using a BigDye v. 1.1 sequencing terminator kit and an ABI PRISM 3130 sequencer (both Applied Biosystems). The fragments were assembled into the final contigs with the SeqMan program (Lasergene package, DNASTAR, Inc.).

Phylogenetic analysis. Sequences were checked for homologous sequences in GenBank, using the BLAST program (Altschul *et al.*, 1990) and were aligned with corresponding sequences of other available PSbMV isolates. Multiple sequence alignments of nucleotide and deduced amino acid sequences were used for the analysis of variability and the construction of a phylogenetic tree using a neighbor-joining method ClustalW. The bootstrap option was run with 1000 re-samplings. The tree was visualized by TreeView program v. 1.6.1 (Page, 1996).

Biological testing. For the pathotype determination, the pea cultivars Fjord, No. 8720 (PI 193586), Bonneville (PI 471187), and Sankia (PI 269774) were used (Hjulsager *et al.*, 2002). Selected PSbMV isolates were mechanically transmitted on the plants. For mechanical transmission, 1 g of infected leaf tissue was grinded with 4 ml of 10 mmol/l phosphate buffer (pH 8.0), supplemented with 1% activated charcoal and Celite. The obtained homogenate was rubbed onto two bottom leaves of an experimental plant 10 days after planting. The symptoms were observed at one-week interval and the systemic infection was confirmed with DAS-ELISA.

Results and Discussion

The occurrence of PSbMV on the pea plants was found in each of the studied growing locations Smržice, Šumperk, and Troubsko in Czech Republic. The infected pea plants manifested symptoms of leaf rolling, stunting, mild mosaic,

Table 1. Characterization of PSbMV isolates according to their infectivity on the pea cultivars

Isolate designation	Locality	Susceptibility of pea cultivars to infection				Pathotype
		Fjord	No. 8720 PI 193586	Bonneville PI 471187	Sankia PI 269774	
PSB58CZ	Smržice	S	R	S	S	P-4
PSB117CZ	Smržice	S	R	S	R	P-1
PSB118CZ	Smržice	S	R	S	R	P-1
PSB141CZ	Smržice	S	R	S	R	P-1
PSB178CZ	Šumperk	S	R	S	R	P-1
PSB329CZ	Šumperk	S	R	S	R	P-1
PSBDCZ	Troubsko	S	R	S	R	P-1
PSBECZ	Troubsko	S	R	S	R	P-1

S = susceptible, R = resistant

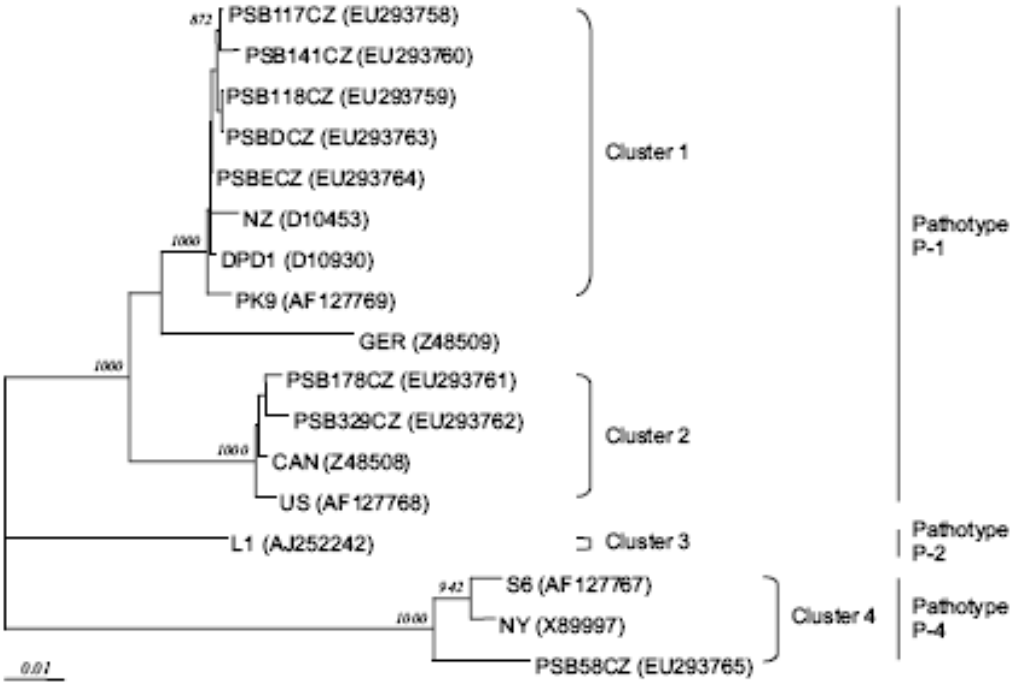


Fig. 1

Phylogenetic tree of PSbMV isolates reconstructed from the 1106 bp of the region 5'(Nter)NIb-CP-UTR3'

The scale bar represents 0.01 nucleotide substitutions per site. Only the bootstrap values >70% are shown.

and vein clearing. The infection with PSbMV was confirmed by DAS-ELISA in all samples. In some cases, mixed infections of PSbMV with PEMV-1 were found.

The eight PSbMV isolates were included in this study. All of them were easily mechanically transmissible and were maintained on the pea cultivar Merkur. The first symptoms of down-rolling leaves were noticed 12 days after inoculation. During the following two weeks the symptoms of vein clearing and leaf yellow mosaic developed. The growth reduction, previously described for NY isolate (Johansen *et al.*, 1996), was typical for the infection by the PSB58CZ isolate.

The 5'(Nter)NIb-CP-UTR3' region of analyzed isolates was amplified using RT-PCR and all fragments were cloned and sequenced. The amplification of the first fragment, using the universal CPUP/P9502 primer pair was successful in all cases. On the other hand, the amplification of the beginning of the CP gene using the newly designed primer pair PSB8812/PSB9440 gave the expected product at a length of 629 bp only for 7 of the 8 isolates. Because the PCR amplification was negative in the case of the isolate PSB58CZ, the successful amplification of a similar fragment of this isolate required to design another primer PSB8800. Afterwards, the amplicon of a length of 641 bp was obtained using the PSB8800/PSB9440 primer pair. The subsequent analysis of the obtained sequence showed that the isolate PSB58CZ lacked the appropriate priming site of primer PSB8812.

The full sequence of chosen segment was obtained for each isolate and their classification as PSbMV was confirmed by comparison with the PSbMV GenBank sequences using the BLAST algorithm. The sequences of 8 tested isolates were deposited in GenBank under Acc. Nos.: EU293758 (PSB117CZ), EU293759 (PSB118CZ), EU293760 (PSB141CZ), EU293761 (PSB178CZ), EU293762 (PSB329CZ), EU293763 (PSBDCZ), EU293764 (PSBECZ), and EU293765 (PSB58CZ).

Phylogenetic analysis based on the 1106 bp long nucleotide sequences of 5'(Nter)NIb-CP-UTR3' region (position 8814–9919 nt according to the isolate DPD1) showed distribution of the Czech isolates into 3 significantly distant clusters (Fig. 1). The five Czech isolates (PSB117CZ, PSB141CZ, PSB118CZ, PSBDCZ, and PSBECZ) clustered with the non-American isolates DPD1, GER, NZ and PK9 in cluster 1; the two Czech isolates (PSB178CZ and PSB329CZ) formed cluster 2 together with the North American isolates. The lentil isolate (L1) formed the separate cluster 3. Only one of the Czech isolates, PSB58CZ belonged to cluster 4 together with the isolates from Pakistan (NY) and Australia (S6). Analysis of the shorter sequence according to Andersen and Johansen

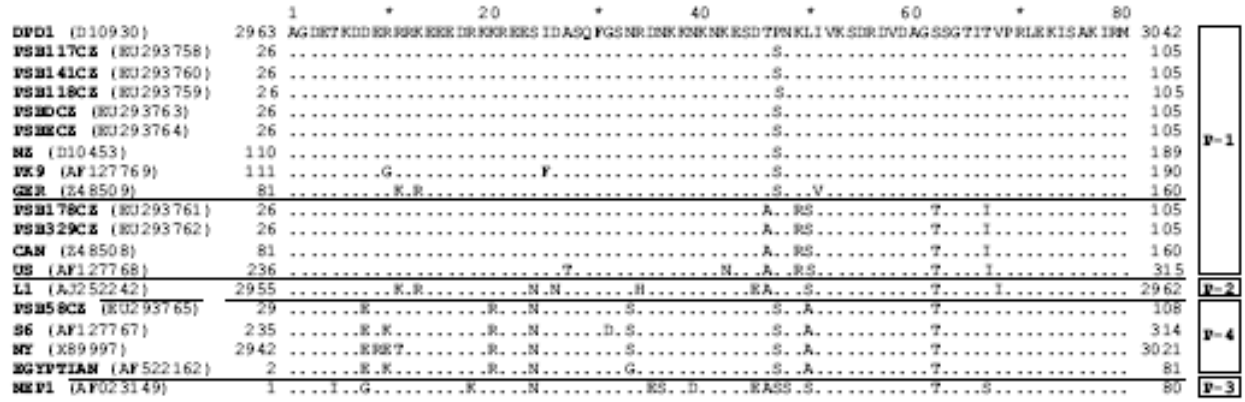


Fig. 2

Partial amino acids sequence alignment of PSbMV isolates

The first 80 deduced amino acids of the coat protein are shown (position aa 2693–3042 according to PSbMV isolate DPD1).

(1998), e.g. the variable beginning of the CP gene with a length of 418 bp (position 8889–9306 nt in the isolate DPD1) allowed the comparison of a larger number of isolates and gave the same topology of the phylogenetic tree with the same distribution of Czech isolates into clusters. The isolate NEP-1, that could have been included in the analysis this way, represented the fifth significant cluster (data not shown).

The analysis of the deduced amino acids sequences of the CP gene showed the presence of typical parsimony informative sites and also supported the previous diversification. Discrete clusters were characterized by the highly conserved amino acids L_{50}, S_{62} for cluster 1, R_{49}, I_{57} for cluster 2, N_{26}, H_{34} for cluster 3, and E_8, R_{20}, A_{59} for cluster 4 (Fig. 2). In all cases, the DAG motif in position 59–61 was conserved, in agreement with the previously described position (Timmerman *et al.*, 1990).

The distribution of the isolates into the 4 clusters corresponded with the division of the PSbMV isolates into pathotypes according to their infection profile on *P. sativum* cultivars. All previously characterized type isolates of pathotypes P-1 (DPD1, US), P-2 (L1), P-3 (NEP1), and P-4 (NY, S6) are members of the formed branches (i.e. groups). Accordingly, the classification of our isolates into pathotypes P-1 and P-4 can be assumed from the biological testing performed on the pea cultivars Fjord, No. 8720 (PI 193586), Bonneville (PI 471187), and Sankia (PI 269774). We noted that the pathotype-specific infection confirmed the classification of the 7 Czech isolates PSbMV (PSB117CZ, PSB141CZ, PSB118CZ, PSBDCZ, PSBECZ, PSB178CZ, and PSB329CZ) as pathotype P-1 and one isolate (PSB58CZ) as pathotype P-4 (Table 1).

Summing up, the molecular analysis of 5'(Nter)NIb-CP-UTR3' region of 8 Czech PSbMV isolates in agreement with the biological testing demonstrated their distribution into two pathotypes P-1 and P-4. According to our information, the classification of the isolate PSB58CZ as the pathotype P-4 represented the first detection of the PSbMV isolate with this characteristic in Europe.

Acknowledgements. This work was supported by the grant QF3072 from the Ministry of Agriculture, Czech Republic and by the grant MSM 6198959215 from the Ministry of Education, Czech Republic. We thank Mr. P. Lemkin for manuscript editing.

References

Alconero R., Provvidenti R., Gonsalves D. (1986): Three pea seedborne mosaic virus pathotypes from pea and lentil germ plasm. *Plant Dis.* **70**, 783–786.

Ali A., Randles J.W. (1997): Early season survey of pea viruses in Pakistan and the detection of two new pathotypes of pea seed-borne mosaic virus. *Plant Dis.* **81**, 343–347.

Altschul S.F., Gish W., Miller W., Myers E.W., Lipman D.J. (1990): Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410.

Andersen K., Johansen I.E. (1998): A single conserved amino acid in the coat protein gene of pea seed-borne mosaic potyvirus modulates the ability of the virus to move systemically in Chenopodium quinoa. *Virology* **241**, 304–311.

Hampton R.O., Mink G.I. (1975): Pea seed-borne mosaic virus. *CMI/AAB Descr. Pl. Viruses* No. **146**.

Grünwald N.J., Chen W., Larsen R.C. (2004): Pea diseases and their management. In Naqvi S.A.M.H. (Ed): *Diseases of Fruits and Vegetables*. Vol. II. Kluwer Academic Publishers, the Netherlands, pp. 301–331.

Hjulsager CK, Lund OS, Johansen E (2002): A new pathotype of Pea seedborne mosaic virus explained by properties if the P3-6k1 and viral genome linked protein (VPg)-coding regions. *Mol. Plant-Microbe Interact.* **15**, 169–171.

Johansen E, Rasmussen OF, Heide M, Borkhardt B (1991): The complete nucleotide sequence of pea seed-borne mosaic virus RNA. *J. Gen. Virol.* **72**, 2625–2632.

Johansen IE, Keller KE, Dougherty WG, Hampton RO (1996): Biological and molecular properties of a pathotype P-1 and a pathotype P-4 isolate of pea seed-borne mosaic virus. *J. Gen. Virol.* **77**, 1329–1333.

Musil M (1966): Über das vorkommen des Blattrollens der Erbse in der Slowakei (Vorrlaufige mitteilung). *Biologia* **21**, 133–138.

Provvidenti R, Alconero R (1988): Inheritance of resistance to a third pathotype of pea seed-borne mosaic virus in Pisum sativum. *J. Hered.* **79**, 76–77.

Page RDM (1996): TREEVIEW: An application to display phylogenetic trees on personal computers. *CABIOS* **12**, 357–358.

Timmerman GM, Calder VL, Bolger LEA (1990): Nucleotide sequence of coat protein gene of pea seed-borne mosaic potyvirus. *J. Gen. Virol.* **71**, 1869–1872.

van der Vlugt RAA, Bouwen I (1997): Identification of potyviruses infecting Alstroemeria. In Dehne HW, Adam G, Diekmann M, Frahm J, Mauler-Machnik A, Van Halteren P (Eds): *Diagnosis and Identification of Plant Pathogens.* Kluwer Academic Publishers, Dordrecht, the Netherlands, pp. 465–467.

Rohožková J, Navrátil M 2011 *P1 peptidase-a mysterious protein of family Potyviridae.* J Biosci 36:189-200.

Review

P1 peptidase – a mysterious protein of family Potyviridae

JANA ROHOŽKOVÁ* and MILAN NAVRÁTIL

Faculty of Science, Palacký University in Olomouc, Šlechtitelů 11, 783 71, Olomouc-Holice, Czech Republic

**Corresponding author (Fax, +420-241-062219; Email, jana.rohozkova@img.cas.cz)*

The Potyviridae family, named after its type member, *Potato virus Y* (PVY), is the largest of the 65 plant virus groups and families currently recognized. The coding region for P1 peptidase is located at the very beginning of the viral genome of the family Potyviridae. Until recently P1 was thought of as serine peptidase with RNA-binding activity and with possible influence in cell-to-cell viral spreading. This N-terminal protein, among all of the potyviruses, is the most divergent protein: varying in length and in its amino acid sequence. Nevertheless, P1 peptidase in many ways is still a mysterious viral protein. In this review, we would like to offer a comprehensive overview, discussing the proteomic, biochemical and phylogenetic views of the P1 protein.

[Rohožková J and Navrátil M 2011 P1 peptidase – a mysterious protein of family Potyviridae. *J. Biosci.* 36 189–200] DOI 10.1007/s12038-011-9020-6

1. Introduction

The Potyviridae family [named after its type member, *Potato virus Y* (PVY)] is the largest of the 65 plant virus groups and families currently recognized (Virus taxonomy list in 2009, ICTV). This family contains over 100 different species in six genera (Berger *et al*. 2000; López-Moya and García 2008). The viruses of this family cause significant losses in agricultural, pastoral, horticultural and ornamental crops (Ward and Shukla 1991). Within the Potyviridae family, the following genera are recognized: *Brambyvirus*, *Potyvirus*, *Rymovirus*, *Macluravirus*, *Ipomovirus* and *Tritimovirus*. A seventh genus, *Bymovirus*, contains viruses with bipartite genomes. Recent attempts to expand the genera include *Blackyvirus* (Susaimuthu *et al*. 2008), *Poacevirus* (Tatineni *et al*. 2009) and *Susmovirus* (Xu *et al*. 2010). Virions are flexuous and rod-shaped, 80 to 900 nm long and 11 to 15 nm wide. The nucleocapsid is composed of around 2000 units of a single structural protein, surrounding one molecule of nucleic acid (Carrington and Dougherty 1988). A feature shared by all of the potyviruses is the induction of characteristic pinwheel or scroll-shaped inclusion bodies in the cytoplasm of infected cells (Edwardson 1974). The predominant way of transmission of potyviruses is through aphids. Other possible means of the virus' spread are by mites and whiteflies. The genome is created by single-stranded RNA of positive sense, around 10 000 bp long. In most potyvirid species with a monopartite genome, their genomic sequence encodes nine multifunctional proteins plus the capsid protein. Each viral protein matures after its proteolytic cleavage from the polyprotein by nuclear inclusion-a peptidase (NIa), helper component proteinase (HC-Pro) and P1 peptidase. NIa and HC-Pro are multifunctional proteins, containing a C-terminal proteolytic domain as well as an N-terminal domain, each of which plays different roles in the replication cycle. Whether the topological quality of P1 protease should be compared to these is still not clear (Verchot *et al*. 1991).

The coding region for P1 proteinase is located at the very beginning of the viral genome. This terminal protein, among all

Keywords. Genetic diversity; P1 serine peptidase; Potyviridae; viral infection

Abbreviations used: BCMNV, *Bean common mosaic necrosis virus*; CI, cylindrical inclusion; CMV, *Cucumber mosaic virus*; CVYV, *Cucumber vein yellowing virus*; HC-Pro, helper component proteinase; MP, movement protein; NIa, nuclear inclusion-a peptidase; NBRF, National Biomedical Research Foundation; PPV, *Plum pox virus*; PSbMV, *Pea seed-borne mosaic virus*; PVA, *Potato virus A*; PVV, *Potato virus V*; PVX, *Potato virus X*; PVY, *Potato virus Y*; SMV, *Soybean mosaic virus*; TEV, Tobacco etch virus; TuMV, *Turnip mosaic virus*; TVMV, *Tobacco vein mottling virus*; WMV, *Watermelon mosaic virus*; YMV, *Yam mosaic virus*; YTHS, yeast two-hybrid system; ZYMV, *Zucchini yellow mosaic virus*

of the potyviruses, is the most divergent protein: varying in length (39–85) and in its amino acid sequence (NCBI Protein Database). This mostly non-conservative potyviral protein is thought to have contributed to the successful adaptation of the potyviruses into a wide range of host species (Valli *et al.* 2007).

With the accumulation of new information that has been assembled about potyviral P1, it is now an appropriate time to review the research progress on this protein. The aim of this review is to offer a comprehensive overview, discussing the proteomic, biochemical, and phylogenetic views of the P1 protein. Additionally, this review attempts to offer the latest research directions, following current theories on the function of the P1 protein.

2. Potyviral P1 protein: function and genetic diversity

It was not very long ago when the molecular properties of potyviruses were a mystery. The first breakthrough was achieved in 1986, when the complete genomic sequences of two members of this group were reported – for *Tobacco etch virus* (TEV) (Allison *et al.* 1986) and the *Tobacco vein mottling virus* (TVMV) (Domier *et al.* 1986). A year later, a step closer to uncovering the mystery occurred. In order to understand the functions of the potyviral proteins, the working group of Domier *et al.* (1986) made use of the then newly determined nucleotide sequences of the RNAs of TEV and TVMV. Subsequently, the predicted amino acid sequences of TVMV- and TEV-encoded proteins were compared with the protein sequence bank of the National Biomedical Research Foundation (NBRF). This study regarded as one of the significant studies that dealt with protein similarities among the potyviruses. After that came the useful achievement of identifying the functional properties of each genome's coding region, which brought cDNA clones of three potyviruses: TVMV (Domier *et al.* 1989), *Plum pox virus* (PPV) (Riechmann *et al.* 1990) and *Zucchini yellow mosaic virus* (ZYMV) (Gal-On *et al.* 1991). The ability to generate a virus infection from the cloned cDNA unlocked the possibility of applying the techniques of genetic engineering, at the molecular level, into their biology. The experiments initially performed on the previously mentioned group of viruses allowed for the beginnings of direct mutational tests and analyses at specific coding regions of the potyviruses, which had direct impact on both its phenotype as well as the process of its pathogenesis.

At this time, we can recognize two different groups of experiments dealing with the function of the P1 protein. On the one hand, there are those trying to solve the real function of P1 protease in several potyviruses: *Plum pox virus* (PPV) (Salvador *et al.* 2008), *Tobacco vein mottling virus* (TVMV) (Brantley and Hunt 1993; Klein *et al.* 1994; Salvador *et al.* 2008), *Cucumber vein yellowing virus* (CVYV) (Valli *et al.* 2006, 2008), *Soybean mosaic virus* (SMV) (Gagarinova *et al.* 2008), *Pea seed-borne mosaic virus* (PSbMV) (Aparicio *et al.* 2005), *Potato virus A* (PVA) (Merits *et al.* 1999; Kekarainen *et al.* 2002), *Potato virus V* (PVV) (Oruetxebarria and Valkonen 2001), *Potato virus Y* (PVY) (Arbatova *et al.* 1998; Pehu *et al.* 1995), *Potato virus X* (PVX) (Vance *et al.* 1995; Pruss *et al.* 1997), *Yam mosaic virus* (YMV) (Aleman-Verdaguer *et al.* 1997), *Tobacco etch virus* (TEV) (Carrington and Dougherty 1988; Carrington and Freed 1990; Verchot *et al.* 1992; Verchot and Carrington 1995a, b) and *Turnip mosaic virus* (TuMV) (Soumounou and Laliberté 1994). On the other hand, there are those experiments concerned with the sequence variability of the P1 coding region in YMV (Aleman-Verdaguer *et al.* 1997), PVA (Kekarainen *et al.* 1999), PVV (Oruetxebarria *et al.* 2000), PVY (Tordo *et al.* 1995; Mukherjee *et al.* 2004), ZYMV (Lee and Wong 1998; Wisler *et al.* 1995), TVMV and TEV (Domier *et al.* 1987), within the group of Potyviruses (Riechmann *et al.* 1992), within the Potyviridae family (Adams *et al.* 2005b; Valli *et al.* 2008), or among several viral families, including Potyviridae (Chare and Holmes 2006). There has been a dedicated effort on the phylogenetic studies of P1. Intensive research during the subsequent years led to a better understanding of the genomic structure of the potyvirus expression, quality, as well as the activity of specific viral proteins. Nevertheless, the P1 protein in many ways is still a mysterious viral protein, showing direct evidence of having some role in host range definition; however, this is yet to be demonstrated.

3. P1 serine peptidase

P1 peptidase was the last of the three peptidases to have been identified as processing the potyviral polyprotein (Verchot *et al.* 1991). In the microbial world, peptidases are not uncommon. Serine, aspartic and cystein peptidases (although not metallopeptidases) have been found in various viruses (Rawlings and Barrett 1993). They have acquired their importance because of their functional involvement in the processing of proteins of viruses that cause certain fatal diseases such as AIDS and cancer. Consequently, we have taken this opportunity to consolidate information about the progress in plant virus research as one of the crucial steps in understanding the activity of RNA viruses, not only in the realm of plants but also with the wider application of those viruses concerning and affecting humans.

All virus-encoded proteases are endopeptidases. The mature protease is released by the autolysis of the precursor (Rao *et al.* 1998). Serine proteases are characterized by the presence of a serine group within their active site. The numerous examples of serine proteases among viruses suggest that they are vital to the organisms (Barret 1994). With the intent of preventing serious losses to several major crop plants, strong efforts continue into research on the viral

proteases, precisely because this portion of the viral infection still offers likely prospects for answers leading to improvements in the defence mechanisms of the host plants.

The first research works dealing with the processing of viral polyproteins with proteases were published between 1979 and 1990 (Korant *et al.* 1979; Ghysdael *et al.* 1981; Franssen *et al.* 1982; Nicklin *et al.* 1986; Krausslich and Wimmer 1988; Wellink and van Kammen 1988; Skalka 1989; Oroszlan and Luftig 1990; Palmberg 1990; Strauss and Strass 1990). The reader is referred to these articles for additional specific information, and a summary of the earlier work on this subject. Of the known potyviral proteinases, NIa and HC-Pro were discovered in their multifunctional characteristics. The first N-terminal peptidase taking part in polyprotein processing (P1 protein) remains the last discovered, without any clear elucidation as to its function during viral infection and propagation. As listed earlier, there have been several interesting works that indirectly brought results which might ultimately be linked to the explanation of P1's true activities.

3.1 *Serine peptidase domain is located at the C-terminus of P1*

Verchot *et al.* (1991) were the first to mention P1 as a third proteinase in the processing of the viral polyprotein. In a wheat germ system, they had synthesized a polyprotein containing the 35 kDa protein and HC-Pro. Proteolysis generates products that resemble fully processed proteins. Furthermore, this protease was classified as a serine-like protease. All then-known facts on Potyviridae were compiled in an extensive review (Dougherty and Semler 1993).

Serine and serine-like peptidases are found in both prokaryotic and eukaryotic organisms, and participate in a wide spectrum of biological reactions. As a group, these enzymes are characterized by the presence of an active site domain that contains a reactive Ser. The active site also comprises two additional amino acids, Asp and His. These three residues make up a catalytic triad (for reviews, *see* Kraut 1977; Bond and Butler 1987; Craik *et al.* 1987a, b; Higaki *et al.* 1987; Sprang *et al.* 1987). The serine-like peptidase domain in the P1 coding region was identified in the C-terminal, highly conserved portion for all serine peptidases, associated with proteolytic activity and containing the Gly-Xaa-Ser-Gly motif (Barret 1986). It has been identified in all of the Potyviridae family members examined (Adams *et al.* 2005a; Valli *et al.* 2007). The conserved His and Asp residues are present upstream of the putative active-site Ser residue (for precise positioning in each virus, *see* Adams *et al.* 2005a; Valli *et al.* 2007); further, their spacing is not typical of those observed in other cellular serine proteases. In experiments involving the substitution of His or Ser in TEV (within positions His-215 and Ser-256), the results of Verchot *et al.* (1991) achieved the elimination of proteolytic activity.

3.2 *Processing of potyviral P1*

The processing of the serine protease, out of HC-Pro and polyprotein, is secured by a specific motif on the border of P1/HC-Pro. Mavankal and Rhoads (1991), with research on *Tobacco vein mottling virus* (TVMV), brought the first approaches in this research field of P1, which released itself from HC-Pro as the N-terminal 34 kDa protein. These facts have since been proven by the immunoprecipitation of proteins with antibody against HC protein. With the specific identification of the cleavage location by the automated Edman radiochemical degradation at positions Phe-256 and Ser-257 (of the TMVM polyprotein) confirmed its necessity for the proper functioning of HC-Pro in TMVM (Mavankal and Rhoads 1991). Later, Verchot *et al.* (1992) confirmed this theory with the specification of Tyr-304 and Ser-305 motif (of the TEV polyprotein) undergoing proteolysis with supplementary His-303. In deletion tests with this specific motif, they concluded that the expressed protein adjacent to His-303 was unable to undergo proteolysis, suggesting that Tyr-304 was not essential for protein processing but only for substrate and/or protein activity. From this finding arose the hypothesis that the cleavage occurs between the Tyr304 and Ser305, because the P1 cleavage site position (Tyr-304) is crucial for substrate recognition by several types of viral and cellular proteases (Verchot *et al.* 1992).

As mentioned earlier, potyviral P1 proteinase is an N-terminal product, derived from a genome-length polyprotein cleaving itself from HC-Pro. Verchot and Carrington (1995b) added a further contribution that this separation, but not P1's own, appears to be essential for viral infectivity. The relevance of P1 integrity for virus amplification and movement from cell to cell was assessed by mutagenesis. Using the coding region of *Tobacco etch virus* (TEV), they produced clones with the entire P1 coding region (DP1) deleted, further having investigated the reporter gene's expression of β-glucuronidase replication as well as the movement of the virus in both tobacco protoplasts and plants. In the protoplasts, this DP1 mutant accumulated to approximately 2% to 3% of the level of parental amplification, with both cell-to-cell and systemic movement, registering decreased values. On the contrary, the clone with a point mutation affecting the Ser of the proteolytic triad (S256A), which abolished the proteolytic activity of the protein and prevented its separation from the rest of polyprotein, yielded a non-viable mutant. A comparison of the infection phenotypes of DP1 and S256A mutants further suggests that the lack of proteolytic separation of P1 and HC-Pro to be more detrimental than the deletion of the P1 sequence altogether.

Experiments in 1992 with a specific motif of the P1 cleavage activity, resulting in it releasing itself from the rest of viral polyprotein, led to the specific information that these experiments had been successful in wheat germ extracts and transgenic plants, but not in rabbit reticulocyte lysate (Verchot *et al.* 1992). Furthermore, there had been previous efforts (Carrington and Freed 1990; Mavankal and Rhoads 1991; Verchot *et al.* 1991) that gave puzzling results. Based on the processing characteristics of P1-proteinase-containing polyprotein, within a mixture of reticulocyte lysate and wheat germ extract, Verchot *et al.* (1991) concluded that some positive-acting factor resides within the wheat germ system. Confirmation of this idea came from an experiment that involved adding small amounts of wheat germ extract to the rabbit reticulocyte cell-free expression system, which led to a revival of P1's protease activity. On the other hand, the addition of reticulocyte to wheat germ extract had no significant effect. Furthermore, there was an experiment using denatured wheat germ extract, which resulted in the assertion that the essential protein for proteinase activity had to be a heat-labile protein. The explanation of this phenomenon does not seem to reside solely in an essential co-factor in the P1 pathway but rather in a factor required for protein synthesis, as well as the proteins involved in the proper co-translational folding of the proteinase in the arrangement of the cleavage site. It is not the first time that a theory suggested this as a possible explanation of P1's ability to process itself under specific conditions. The involvement of accessory factors in polyprotein processing at specific cleavage sites has precedents in several other viral groups, including the comoviruses (Goldbach 1990), the picornaviruses (Harris *et al.* 1990) and flaviviruses (Rice and Strauss 1990).

3.3 *P1's N-terminal portion – the most variable part of the genome in all potyviruses*

From the first announcements regarding the sequence of the P1 coding region, this protein has been mentioned as the least conserved protein among the potyviruses (Domier *et al.* 1987; Vance *et al.* 1992), ranging in size from 30 to 63 kDa. Owing to this great divergence, and the discovery of the conserved motif in the C-terminal portion, it is obvious that the increase in the size is created by the N-terminal portion of P1 of each potyvirus. In a study by Verchot and Carrington (1995a), it was shown to be non-essential for most functions such as viral propagation, cell-to-cell movement, and viral replication. By use of the GUS reporter system, which is appropriate for quantitative measurements of the parental and mutant TEV genome amplifications in inoculated protoplasts (Carrington *et al.* 1993; Dolja *et al.* 1992; Restrepo-Hartwig and Carrington 1994), Verchot and Carrington (1995a) characterized three truncated forms of P1 protein with the deletion in the N-terminal portion. The P1 coding region of the TEV lacked the region from 1 to 139 nt in the Δ5 mutant, and the region of 1–157 nt in the Δ6 mutant, i.e. in almost half of the P1 coding region, but not affecting the catalytic triad of proteolytic activity. Interestingly, most dramatic mutations in the Δ5 and Δ6 constructs actually had only small effects on viral replication and cell-to-cell movement. Experiments by Kekarainen *et al.* (2002) and Rajamäki *et al.* (2005) gave results inconsistent with those of deletions. Both of these two working groups have confirmed the high tolerance of P1 to insertions. The aim of the work by Kekarainen *et al.* (2002) was to map the genome regions (both essential and non-essential) for PVA propagation.

To that end, they generated 15-bp-insertion mutants. The C-terminal region of P1 tolerated most of the insertions. Based upon those results of tolerance to short insertions, they constructed a map of the identified essential, as well as non-essential, regions of P1. From this viewpoint, we can note a conflict in the categorization of the importance of P1's N-terminal portion. Seventy percent of the regions described as essential, resulting in non-viable mutants, were located in the N-terminal half of P1. Rajamäki *et al.* (2005) also demonstrated a similar tolerance. They confirmed this ability by the insertion (insertion of up to 783 nucleotides) and expression of heterologous protein in PVA's N-terminal portion of the P1-coding region.

The next efforts to identify an even-lesser-conserved motif in the P1 N-terminal portion (which could help in understanding its evolution), and at first sight a non-conservative portion of the potyviral genome, were by Gagarinova *et al.* (2008) and Valli *et al.* (2008). They both focused on the sequencing analysis of the whole coding region of the P1 protein in batches of potyviruses. Their interesting results and conclusions are analysed in section 4 of this article.

3.4 *P1 is able to bind RNA*

The net charge of the P1 protein was computationally evaluated to be +30 (TVMV; Brantley and Hunt 1993) and +38 (TuMV; Soumounou and Laliberté 1994), which are much higher than those of any other potyviral protein. This charge has been characterized as a z value. It has been proposed that the binding of a ligand of charge $+z$ to a linear nucleic acid would neutralize z phosphates, resulting in the release of the counter-ions that are thermodynamically associated with those z phosphates (Record *et al.* 1976). The existence of a net charge for P1, of the magnitude mentioned above, led to the suggestion that it represents nucleic-acid-binding activity. Brantley and Hunt (1993) were the first to assert that P1 generally binds RNA, and they subsequently decided to characterize the possible RNA-binding properties. They had shown that P1 is able to bind 35-nt-long RNA, and that the size of the binding site is consistent with the number of phosphate groups that

would be neutralized by the binding of a protein of the +30 charge previously mentioned (Mascotti and Lohman 1990). The next set of experiments were led by Soumounou and Laliberté (1994), who had analysed P1 of TuMV, and contributed further with pronouncements of the binding capabilities to ss- and ds-RNA. Under *in vitro* conditions, it was confirmed that P1 binds to ss- but not to ds-DNA. The potential interaction domain of P1 in TuMV was characterized as the very basic domain, RSSRAMKQKRARERR RAQQ (spanning residues 150 to 168), which has the potential to interact with nucleic acids. The basic residues (Lys and Arg) can form ionic bonds with the negatively charged phosphate groups; whereas amide (Q), acidic (E) or hydroxylated (S) amino-acids could interact, via hydrogen bonds, with the nucleic acid bases (Soumounou and Laliberté 1994). Contributing factors that have led to the speculation of P1's silencing activity during the viral infection cycle should be mentioned. Concluding with the finding that the depletion of the entire P1 protein had not disrupted the infection cycle, and that even P1's ability to bind RNA, could suggest that this protein acts as a helper protein, improving the expression of the viral genome; yet, on the other hand, P1 could subdue the defense activities of the host plant.

In order to prove P1's activity in overcoming host plant defense mechanisms, several experiments have been performed that dealt with transgenic plants expressing the short-sequence homologue to P1 (Mäkki-Valkama *et al.* 2000a, b; Tavert-Roudet *et al.* 1998). However, more extensive research has attempted to define the extent of P1's involvement in the processes of interactions between virus and its host. Pruss *et al.* (1997) reported that the potyviral sequence P1/HC-Pro enhanced the pathogenicity and accumulation of two heterologous viruses [*Cucumber mosaic virus* (CMV) and TMV]. Kasschau and Carrington (1998) dealt with P1 and its influence on transcription activity, performing the nuclear transcription assay. These experiments indicated that the silencing suppression activity of P1/HC-Pro could possibly be on the post-transcriptional level. These two results are in agreement, and could reveal that this combination of molecules can condition an enhanced susceptibility within a host through the interdiction of the defense response. In this connection, Valli *et al.* (2006) proved the crucial role of P1 in the HC-Pro gene's silencing activity during the infectious, providing further evidence that P1 enhances the activity of HC-Pro in the *Potyvirus* genus.

One very special situation within the P1 proteinase coding region and the activity of this protein has been described in the ipomovirus (family Potyviridae) *Cucumber vein yellowing virus* (CVYV). It was recently reported that this virus lacks HC-Pro, but has a duplicated P1 coding sequence (Janssen *et al.* 2005). After cleavage, two mature proteins are produced, P1a and P1b. Both of these are serine proteases (Valli *et al.* 2007). HC-Pro in potyviruses is the typical silencing suppressor of viruses of the family Potyviridae (Anandalakshmi *et al.* 1998; Brigneti *et al.* 1998; Kasschau and Carrington 1998). The crucial point of this P1 study was a demonstration of P1b protease activity, as well as the RNA-silencing suppression ability, which is consistent with that of HC-Pro (Valli *et al.* 2008). According to the subsequent discoveries for the RNA-silencing quality, the specific LXKA motif is crucial – it has been suggested as the zinc finger motif (Cox and McLendon 2000; Klug and Rhodes 1987); after its point mutation, P1b's silencing activity was abolished. Nevertheless, disruption of protease activity has no acute effects on this silencing quality. Additionally, there were informational data sets with complementary findings about P1 ability to bind siRNA in crude plant extracts. This piece of information supports the suggestion that CVYV P1b uses a strategy of siRNA sequestration, with the interference of viral RNA degradation. Within the context of this interesting discovery, it is very important to mention that the positive charge of the P1 protein is universal in all of the members within the genus *Potyvirus*. In contrast, the ipomoviral P1bs, or P1b-like P1 protein from ipomoviruses (as well as tritimoviruses), either has a neutral or a negative charge.

3.5 *Atypical P1s – Tritimovirus and Ipomovirus*

Because of phylogenetic discoveries, we now know that rymoviruses are more closely related to potyviruses than to ipomoviruses. Tritimoviral P1 protein constitutes an independent branch in the tree (Adams *et al.* 2005b). The most conspicuous difference between the two types of P1 is represented by their pI difference, which probably reflects not only a great phylogenetic distance but also some functional divergence. As has been mentioned before, CVYV lacks the HC-Pro gene and its tritimo-like P1b protein appears to compensate for this defect. This could be conclusive evidence for the occurrence of the functional diversification between poty-like and tritimo-like P1s', P1a's, as well as P1b's own protease activities. If these two putative P1 protease domains are included in the phylogenetic analysis, P1a clustered with the potyviral and rymovila P1s and P1b was more closely related to the tritimoviral P1s (Valli *et al.* 2007; data not shown)

3.6 *P1's localization in infected plant cell*

Arbatova *et al.* (1998) completed one of the first attempts to localize the P1 protein on the ultra-structural level of the infected cell. They produced polyclonal antibody against the P1 protein of PVY-O. Fusing this antibody to gold particles, they conducted their experiments with the immunogold localization of the P1 protein in the cells of infected plants.

This protein has been found in association with cytoplasmic inclusion bodies, characteristic for several potyviruses. Particles marking the P1 protein have also been found freely diffused in the cytoplasm. However, no significant P1 antibody binding with other plant cell organelles, or with the cell wall and plasmodesmata, was detected with the immunogold labelling. Arbatova *et al.* (1998) described their findings of P1 associated with pinwheel-shaped inclusion bodies with two possible explanations: (1) P1 could stay associated with this complex due to its simultaneous synthesis there as well, as it has been shown by the P3 protein (Ammar *et al.* 1993, 1994; Rodríguez-Cerezo *et al.* 1991, 1993). (2) As suggested by Ammar *et al.* (1994), all potyviral proteins might by synthesized in, on or near the inclusions, which are associated with a rough endoplasmic reticulum. Speculative, yet conclusive, seems to be the inference stemming from P1's ability to bind RNA, and the fact that several point mutations have led to a decrease of viral particles in infected plants. The piece of evidence behind the development of this theory is the newly found helicase activity of cylindrical inclusion protein (CI); and related to this matter, replication of viral RNA is likely to also be associated with the connections found with cytoplasmic inclusions. This might be the missing piece accounting for the P1 cytoplasmic localization. In regard to this information, it would be appropriate for someone to supplement such conclusions with experiments that confirm any biochemical pathways that might help account for an explanation of the localization, i.e. the finding of molecular interaction partners of P1 protein during the infection cycle.

3.7 *P1 and its activity in viral cell-to-cell movements and systemic spreading*

P1 activity, influencing virus cell-to-cell movement, has been hypothetically suggested (Domier *et al.* 1987; Hull 1989; Maiss *et al.* 1989; Atabekov and Taliansky 1990; Ward and Shukla 1991; Riechmann *et al.* 1992). Later, such P1 activity was refuted by Verchot and Carrington (1995a, b). As previously mentioned, experiments with point-mutated P1 protein of TEV (Verchot and Carrington 1995b) resulted in a group of nonviable viruses (three mutations termed S256A, F and Δ304). Nevertheless, a fourth proteinase-debilitating mutation (termed C) caused a slow-infection phenotype. According to these findings, the speculation about a direct movement function of the P1 protein now appears to have little basis; however, a movement-enhancing function cannot be totally excluded. Several independent researchers (Klein *et al.* 1994; Cronin *et al.* 1995; Verchot and Carrington 1995b; Kasschau *et al.* 1997) have found that P1 and HC-Pro are both active enhancers of genome amplification at the single-cell level, and have additionally demonstrated the long-distance movement. However, parallel to this information, it has also been asserted that neither of these two proteins appears to be required for cell-to-cell movement. Arbatova *et al.* (1998) later indirectly confirmed this assertion. Immunogold localization of PVY P1 protein in infected plants showed an absence of this protein in the cell wall and plasmodesmata (Arbatova *et al.* 1998). On the other hand, we need to mention the situation in the sobemoviruses. Two open reading frames (ORF1 and ORF2) create their genome, while P1 is located in ORF1. Due to the theory about the non-essentiality of P1 in virus spread, there have been conflicts with previous findings. Bonneau *et al.* (1998) proclaimed, working on *Rice yellow mottle virus* (RYMV), that the creation of deleted and frame-shifted mutants in the P1 coding region (resulting in truncations of 83 amino acids from the C-terminus of P1) were incapable of replicating in protoplasts. Furthermore, a mutant not expressing P1 at all replicated in a protoplast at a reduced level (0.5–2× less), when compared with the replication of the wild-type RNA. In the case of cell-to-cell movement as a system infection, none of the mutants caused systemic infection in the host rice plants. Their results demonstrated that in the case of the sobemoviral infection, P1 of RYMV is indispensable for virus replication; however, nucleotide deletions or additions in ORF1 are lethal for virus replication. Furthermore, P1 of RYMV is required for the infection of plants and is important for the spread of the virus.

3.8 *P1's interaction partners in the infected cell*

Further efforts to study the P1 molecular interaction partners were led by Merits *et al.* (1999). However, they tried to analyse the possible P1 pathway during an infection cycle; thus, their work involved an interaction study of only the viral proteins. Consistent with the balance between P1's affinity for one or more viral proteins, they had expected P1's activity as a helper factor of the presently identified mechanisms during the infection cycle. Using *Escherichia coli*–expressed recombinant proteins in two *in vitro* interaction assays, and a genetic yeast two-hybrid system (YTHS), they tried to analyse the protein–protein interactions of the P1 and P3 proteins of PVA with six other viral proteins, creating presumed replication complexes. In these *in vitro* experiments, they concluded that P1 and P3 interacted with each other, as well as with proteins of the putative replication complex of potyvirus: RNA helicase (CI), viral protein genome linked (VPg), NIa peptidase part (NIa-pro) and RNA-dependent RNA-polymerase (NIb). P1 also interacted with itself and with HC-Pro. In this situation, it seems to be reasonably confirmed that the interaction between the P1 and CI proteins form the main superficial component of cytoplasmic inclusion bodies, as shown by Arbatova *et al.* (1998) and Rodríguez-Cerezo *et al.* (1993, 1997).

Experiments that resulted in the interaction of P1 with HC-Pro have endeavoured to suggest that these interactions are associated with events that help proteinase domains to recognize their cleavage sites. The creation of a self-dimer of P1, and its possible interaction with P3, has contributed to the proposed role of the protein's interaction with the replication complex of the potyvirus. Other interactions explored in regard to viral proteins are yet to be explained. However, it is necessary to add that Merits *et al.* (1999) have been cautious in the interpretation of their results reached, because protein expression in bacterial cells possibly lacking post-translational modification (phosphorylation and glycosylation) could have a large influence on real protein activity.

A novel identification of the interaction partner of P1 was discovered using yeast two-hybrid screen coupling P1 from SMV-P (*Pinellia* isolate) and a c-DNA expression library of its host, the aroid *Pinellia ternate*. Shi *et al.* (2007) identified an interesting 23.7-kDa-sized molecule that was interacting with P1 and was closely related to the cytochrome b6/f complex Rieske Fe/S genes of plants. The topology of the interaction of P1 to Rieske Fe/S was found to be located at the N-terminal part of the protein (1–33 amino acids), interacting with the transitional form of the Rieske Fe/S protein; whereas, a 34- to 82-amino-acid-long motif was able to interact with the entire Rieske Fe/S protein. Owing to these results, some authors have suggested that the P1–Rieske Fe/S pair is likely to be involved in symptom development, and that the very variable N-terminus of P1 may play an important role in host adaptation.

A specific feature about P1, recently discovered, is its ability to induce production of the HSP70 heat-shock protein. Experiments dealing with this activity had been performed by Aparicio *et al.* (2005), using *GUS* as a reporter gene fused to the HSP70 coding region of *Arabidopsis*. This construct, infiltrated into the leaf of *Nicotiana benthamiana*, allowed them to see the different responses of the promoter–reporter system. A few of the tested constructs have been used for testing the hypothesis that viral induction of pHSP70-GUS represents a broad response to the expression of the virus' proteins. For their experiment, Aparicio *et al.* selected *Tobacco mosaic virus* (TMV, genus *Tobamovirus*) movement protein (MP), the P1 protein of *Pea seed borne mosaic virus* (PSbMV, genus *Potyvirus*) and a 6 kDa protein (6K1) fused to the cylindrical inclusion (CI) protein from PSbMV. All of these proteins showed a significant induction of pHSP70-GUS as increased *GUS* activity; the P1 protein was the most activating one. Increase in the activity of heat-shock protein have been demonstrated by a wide range of viruses in diverse host plants including pea (*Pisum sativum*), *Nicotiana benthamiana*, squash (*Cucurbita pepo*), tobacco (*Nicotiana tabacum*) and *Arabidopsis thaliana* (Aranda *et al.* 1996, 1999; Escaler *et al.* 2000; Havelda and Maule 2000; Jockusch *et al.* 2001; Whitham *et al.* 2003). These authors have conceding that viral proteins crucial to the viral infection cycle may have a specific induction mechanism that would meet the particular needs of viral genome replication and expression. In any case, there remains a possible alternative interpretation, where the P1 activity is just like the response of pHSP70 in the ability of cytosol to sense the individual properties of particular proteins when expressed at a high level. This phenomenon is reminiscent of the unfolded protein response observed when the induced accumulation of proteins in the endoplasmic reticulum also induces a specific suite of chaperons.

4. Evolution of P1 serine proteinase

According to the MEROPS databases, P1 proteinase has been classified within the Clan SA, S30 family of serine proteases. It is possible to attempt an overview of the evolution of P1 proteinase. Although there are publications that have concluded that the evolution of P1 in Potyviridae emerged from the sequences of the coding region, with their differences in evaluating possible conserved motifs (listed further ahead), we can speculate about phylogenesis inside the protease family, leading to a theory about its selection under the functional properties of proteinase and its adaptation in the different metabolic pathways of new host organisms (Barrett and Rawlings 2007). Initially, the first announcements about this phenomenon were described by Rawlings and Barrett (1993), who tried to identify specific groups of peptidase, according to scored similarities in a sequence of over 600 peptidases. They suggested that the classification by peptidase families could be used as an extension of the present classification by catalytic type. No studies dealing with the evolutionary origins of family S30 have been done; however, several families of serine proteinases present in viral organisms have been found, for example, the S3-*Togavirus* endopeptidase with the catalytic triad His/Asp/Ser as well as the family S14-ClpP, identified by *Potato leafroll luteovirus* genomic RNA (Luteoviridae, the genome is created in a similar manner to Potyviridae, by single-stranded RNA of positive sense) with the catalytic motif Ser/His (Asp not known). Rawlings and Barrett (1993) concluded that due to the highly conserved catalytic triad motif of serine peptidase activity, these proteins likely share a common evolutionary origin with family S1 (most of the eukaryotic and bacterial serine peptidases), despite the differences in their sequences. Due to the very small conserved portions, there might be some doubts about the actual relationships between the first site-distinct viral species. Structural/functional studies in many of the peptidase families have highlighted particular amino acid residues that contribute to the specificity sites of the enzymes. Variations of these, between sequences, can be taken as strong indications of the

differences in substrate specificity, and therefore in their function. Such differences would surely require the assignment of the peptidases concerned to be used by the distinct species (Barrett and Rawlings 2007).

4.1 *Phylogenesis of potyviral P1 protein*

Recombination is one of the main forces driving plant virus evolution (García-Arenal *et al.* 2003; Roossinck 2003). Both interspecies and intraspecies recombination events are involved in potyviral evolution, and some of these affect the P1 sequence (Glais *et al.* 2002; Desbiez and Lecoq 2004; Tan *et al.* 2004; Larsen *et al.* 2005; Valli *et al.* 2006). Up to now, there have been several attempts at phylogenetic analysis of the most variable coding region of P1. Each of these has been made achievable from the RNA or amino acid viral sequences, which were taken from theoretical cDNA translations. Despite these functional analyses, phylogenetic analyses have been performed on TVMV and TEV (Domier *et al.* 1987), WMV (Desbiez and Lecoq 2004; Ali *et al.* 2006) and *Soybean mosaic virus* (SMV) (Gagarinova *et al.* 2008). Additionally, Valli *et al.* (2007) attempted the widest overview and most extensive endeavours to explain the origin and phylogeny path of potyviral P1. They performed extensive computational analysis of P1 proteins from 53 virus species in the family Potyviridae. In their last, most extensive assay (Valli *et al.* 2007), they have tried to explain recombination and gene duplication in the evolutionary diversification of P1 proteins.

The first effort to specify some homology or conserved motifs in the P1 coding region, shared by two potyviruses (TVMV and TEV), was described by Domier *et al.* (1987). However, after the TVMV and TEV polyprotein sequence comparison, they announced that the regions of the N-terminal and the 42 kDa region (also involving the P1 coding region) clearly showed less homology. They justified these findings by suggesting that the apparent lack of sequence conservation found in these two regions was due to the involvement of these proteins in specific virus–host interactions.

An extension of the theory regarding the presence of specific motifs in correlation with the host–virus interaction came from the work of Desbiez and Lecoq (2004). They had focused their attention on the N-terminal region of P1, and had assumed it was especially relevant in regard to the previously mentioned interaction between the virus and host organism. This, together with the theory about the interspecific recombination between two related potyviruses WMV and SMV, revealed the specific nucleotide sequence variability of WMV in the 5′ part of the genome (Desbiez and Lecoq 2004). Using that paradigm, Larsen *et al.* (2005) working on insertion mutants of *Bean common mosaic necrosis virus* (BCMNV) demonstrated by the differences found within the P1 N-terminus that although P1 played some significant role in the pathogenicity and virulence, it did not appear to affect the virus–host range (Larsen *et al.* 2005).

The results from comparison analysis suggested that not only intraspecies and intragenus but also intergenus recombination within the P1 gene contributed to potyvirus evolution. Well-conserved motifs were identified within the C-proximal protease domain of all *Potyvirus* P1s (H-D/E-G-x-S-G-I/V-I/V-R-G). The cleavage site between P1 and the next protein, HC-Pro, is also well conserved. It is 22 to 28 amino acids downstream of the strictly conserved RG dipeptide, and it has as its consensus sequence I/V/L/M-x-H/E/Q-F/Y ↓ S. Short, but also frequently present, conserved motifs have been demonstrated to be spread all over the P1 coding region. Motifs IXFG and VELI have been shown to be represented in most potyviruses and rymoviruses. Conversely, there are also motifs only represented within a group of potyviruses, for example, ISIXGG, TPS and FLXG in the small group of potyviruses (for a definitive review, *see* Valli *et al.* 2007). All of those are located approximately in the middle of the P1 coding region. Finally, the identified cystein-rich domain is represented by 13 potyviral viruses usually found in the N-terminal portion of P1. Despite their certain identification, these short motifs have no simple phylogenetic relationships justifying either their presence or absence. Interestingly, all of these results suggest that the potyviral P1 gene has undergone extensive and uneven evolutionary diversification, which has not always paralleled the evolution of the complete genome (Valli *et al.* 2007).

Gagarinova *et al.* (2008) conducted the latest experiments dedicated to gathering evidence of recombination breakage points in P1 on selected potyviruses (which could be helpful in an explanation of P1 phylogenesis). In contrast to previous studies, they used the Recombination Detection Programme 3.3.1 (RDP3), which provided automated analysis using the RDP, GENECONV, Bootscan, MaxiChi, Chimera and SiScan methods (Martin *et al.* 2005) instead of the Genetic Algorithms for Recombination Detection (GARD, Kosakovsky Pond *et al.* 2006) used by Valli *et al.* (2007) for mapping phylogenetic detection of recombination using a genetic algorithm, and Chare and Holmes's (2006) application of Sawyer's Runs Test for the detection of recombination events. Gagarinova *et al.* (2008) demonstrated for the first time that recombination occurred during SMV evolution among distinct viral isolates, thus providing evidence that at least two distinct viral SMV pathotypes could simultaneously infect a host cell and exchange genetic materials through RNA recombination.

5. Concluding remarks

The understanding of P1 peptidase has expanded rapidly in the genomic as well as in proteomic area. The brief overview of known facts about this peptidase presented herein demonstrates

that much is known. However, P1 remains the mysterious protein. In our opinion, a possible way to uncover P1 properties is to identify proteins interacting with it during infection cycle and to try to co-localize them on the level of light microscopy as well as electron microscopy. From the proteomic point of view, it is also important to identify the concrete structure of this protein. In spite of its high variability in the coding sequence all over Potyviridae, there have to be conserved structural motif (even several amino acids) responsible for P1's interaction with other active molecules of virus or host plant. This identification seems to be crucial for uncovering the real activity of P1. Biochemical and proteomic studies of P1 peptidase are essential for this. Such work will offer us a deeper understanding of the involvement of P1 in viral process, spreading and its involvement during viral infection.

Acknowledgements

This work was supported by grant MSM 6198959215 (Ministry of Education, Youth and Sports of the Czech Republic). The excellent critical review by Juan Antonio Garcia Alvarez is gratefully acknowledged. We thank Peter Lemkin for editing the manuscript.

References

Adams MJ, Antoniw JF and Fauquet CM 2005a Molecular criteria for genus and species discrimination within the family Potyviridae. *Arch. Virol.* **150** 459–579

Adams MJ, Antoniw JF and Beaudoin F 2005b Overview and analysis of the polyprotein cleavage sites in the family Potyviridae. *Mol. Plant Pathol.* **6** 471–487

Aleman-Verdaguer ME, Goudou-Urbino C, Dubern J, Beachy RN and Fauquet C 1997 Analysis of sequence diversity of the P1, HC, P3 NIb and CP genomic regions of several yam mosaic potyvirus isolates: implications for the intra-species molecular diversity of potyviruses. *J. Gen. Virol.* **78** 1253–1264

Ali A, Natsuaki T and Okuda S 2006 The complete nucleotide sequence of a Pakistani isolate of *Watermelon mosaic virus* provides further insights into the taxonomic status in the *Bean common mosaic virus* subgroup. *Virus Genes* **32** 307–311

Allison R, Johnston RE and Dougherty WG 1986 The nucleotide sequence of the coding region of tobacco etch virus genomic RNA: evidence of the synthesis of a single polyprotein. *Virology* **154** 9–20

Ammar ED, Rodriguez-Cerezo E, Shaw JG and Pirone TP 1993 Association of virions and coat protein of tobacco vein mottling potyvirus with cylindrical inclusions in tobacco leaves and protoplasts. *Phytopathology* **83** 1374–1375

Ammar ED, Rodriguez-Cerezo E, Shaw JG and Pirone TP 1994 Association of virions and coat protein of tobacco vein mottling potyvirus with cylindrical inclusion in tobacco cells. *Phytopathology* **84** 520–524

Anandalakshmi R, Pruss GJ, Ge X, Marathe R, Mallory AC, Smith TH and Vance VB 1998 A viral suppressor of gene silencing in plants. *Proc. Natl. Acad. Sci. USA* **95** 13079–13084

Aparicio F, Thomas CL, Lederer C, Niu Y, Wang D and Maule AJ 2005 Virus induction of Heat shock protein 70 reflects a general response to protein accumulation in the plant cytosol. *Plant Physiol.* **138** 529–536

Aranda MA, Escaler M, Thomas CL and Maule AJ 1996 Induction of HSP70 and polyubiquitin expression associated with plant virus replication. *Proc. Natl. Acad. Sci. USA* **93** 15289–15293

Aranda MA, Escaler M, Thomas CL and Maule AJ 1999 A heat shock transcription factor in pea is differentially controlled by heat and virus replication. *Plant J.* **20** 153–161

Arbatova J, Lehto K, Pehu E and Pehu T 1998 Localization of the P1 protein of *potato Y potyvirus* in association with cytoplasmic inclusion bodies and in the cytoplasm of infected cells. *J. Gen. Virol.* **79** 2319–2323

Atabekov JG and Taliansky ME 1990 Expression of plant virus-coded transport function by different viral genomes. *Adv. Virus Res.* **38** 201–248

Barret AJ 1986 An introduction to the proteinases; in *Proteinase inhibitors* (eds.) AJ Barrett and G Salvesen (New York: Elsevier) pp 3–22

Barret AJ 1994 Proteolytic enzymes: serine and cystein peptidases. *Method. Enzymol.* **244** 1–15

Barrett AJ and Rawlings ND 2007 'Species' of peptidases. *Biol. Chem.* **388** 1151–1157

Berger PH, Banett OW, Brunt AA, Colinet D, Edwardson JR, Hammond J, Hill J, Jordan RL, Kashiwazaki S, *et al.* 2000 Family Potyviridae; in *Virus taxonomy: Seventh report of International Committee on Taxonomy of Viruses* (Academic Press) pp 703–724

Bond JS and Butler PE 1987 Intracellular proteases. *Annu. Rev. Biochem.* **56** 333–364

Bonneau C, Brugidou C, Chen L, Beachy R and Fuaquet C 1998 Expression of the Rice yellow mottling virus P1 protein *in vitro* and in vivo and its involvement in virus spread. *Virology* **244** 79–86

Brantley JD and Hunt AG 1993 The N-terminal protein of the polyprotein encoded by the potyvirus tobacco vein mottling virus is an RNA-binding protein. *J. Gen. Virol.* **74** 1157–1162

Brigneti G, Voinnet O, Li WX, Ji L H, Ding SW and Baulcombe DL 1998 Viral pathogenicity determinants are suppressors of transgene silencing in *Nicotiana bethamiana*. *EMBO J.* **17** 6739–6746

Carrington JC and Dougherty WG 1988 A viral cleavage site cassette: identification of amino acid sequences required for tobacco etch virus polyprotein processing. *Proc. Natl. Acad. Sci. USA* **85** 3391–3395

Carrington JC and Freed DD 1990 Cap-independent enhancement of translation by a plant potyvirus 5′ nontranslated region. *J. Virol.* **64** 1590–1597

Carrington JC, Haldeman R, Dolja VV and Restrepo-Hartwig MA 1993 Internal cleavage and trans-proteolytic activities of the VPg-proteinase (NIa) of tobacco etch potyvirus in vivo. *J. Virol.* **67** 6995–7000

Chare ER and Holmes EC 2006 A phylogenetic survey of recombinant frequency in plant RNA viruses. *Arch. Virol.* **151** 933–946

Cox EH and McLendon GL 2000 Zinc-dependent protein folding. *Curr. Opin. Chem. Biol.* 4 162–165

Craik CS, Roczniak S, Lagman C and Rutter RJ 1987a The catalytic role of the active site aspartic acid in serine proteases. *Science* 237 909–913

Craik CS, Roczniak S, Sprang S, Fletterick R and Rutter W 1987b Redesigning trypsin via genetic engineering. *J. Cell Biochem.* 33 199–211

Cronin S, Verchot J, Haldeman-Cahill R, Shaad MC and Carrington JC 1995 Long-distance movement factor: a transport functions of the potyvirus helper component-proteinase. *Plant Cell* 7 549–559

Desbiez C and Lecoq H 2004 The nucleotide sequence of *Watermelon mosaic virus* (WMV, Potyvirus) reveals interspecific recombination between two related potyviruses in the 5′ part of the genome. *Arch. Virol.* 149 1619–1632

Dolja VV, McBride HJ and Carrington JC 1992 Tagging of plant potyvirus replication and movement by insertion of β-glucuronidase into the viral polyprotein. *Proc. Natl. Acad. Sci. USA* 89 10208–10212

Domier LL, Franklin KM, Shahabuddin M, Hellmann GM, Overmeyer JH, Hiremath ST, Siaw MFE, Lomonossoff GP, *et al.* 1986 The nucleotide sequence of tobacco vein mottling virus RNA. *Nucleic Acids Res.* 14 5417–5430

Domier LL, Shaw JG and Rhoads RE 1987 Potyviral proteins share amino acid sequence homology with *picorna-*, *como-* and *caulimoviral* proteins. *Virology* 158 20–27

Domier LL, Franklin KM, Hunt AG, Rhoads RE and Shaw JG 1989 Infectious *in vitro* transcripts from cloned cDNA of a potyvirus, tobacco vein mottling virus. *Proc. Natl. Acad. Sci. USA* 86 3509–3513

Dougherty WG and Semler BL 1993 Viral proteinases; *Microbiol. Rev.* 57 781–822

Edwardson JR 1974 Some properties of the potato virus Y-group. *Fla. Agric. Exp. Stn. Monogr. Ser.* 4 225

Escaler M, Aranda MA, Thomas CL and Maule AJ 2000 Pea embryonic tissues show common responses to the replication of a wide range of viruses. *Virology* 267 318–325

Franssen H, Goldbach R, Broekhuijsen M, Moerman M and van Kammen A 1982 Expression of middle-component RNA of *Cowpea mosaic virus*: *in vitro* generation of a precursor to both capsid proteins by a bottom-component RNA-encoded protease from infected cells. *J. Virol.* 41 8–17

Gagarinova AG, Babu M, Strömvik M and Wang A 2008 Recombination analysis of Soybean mosaic virus sequences reveals evidence of RNA recombination between distinct pathotypes. *Virol. J.* 5 143–151

Gal-On A, Antignus Y, Rosner A and Raccah B 1991 Infectious *in vitro* RNA transcripts derived from cloned cDNA of the cucurbit potyvirus *Zucchini yellow mosaic virus*. *J. Gen. Virol.* 72 2639–2643

García-Arenal F, Fraile A and Malpica JM 2003 Variation and evolution of plant virus populations. *Int. Microbiol.* 6 225–232

Ghysdael J, Neil JC and Vogt PK 1981 Cleavage of four avian sarcoma virus polyproteins with virion protease p15 removes gag sequences and yields large fragments that function as tyrosine phosphoacceptors *in vitro*. *Proc. Natl. Acad. Sci. USA* 78 5847–5851

Glais L, Tribodet M and Kerlan C 2002 Genomic variability in *Potato potyvirus Y* (PVY): evidence that PVY^{NW} and PVY^{NTN} variants are single to multiple recombinants between PVY^{O} and PVY^{N} isolates. *Arch. Virol.* 147 363–378

Goldbach H 1990 Plant viral proteinases; *Semin. Virol.* 1 335–346

Harris KS, Hellen CUT and Wimmer E 1990 Proteolytic processing in the replication of picornaviruses. *Semin. Virol.* 1 323–333

Havelda Z and Maule AJ 2000 Complex spatial responses to cucumber mosaic virus infection in susceptible *Cucurbita pepo* cotyledons. *Plant Cell* 12 1975–1985

Higaki JN, Gibson BW and Craik CS 1987 Evolution of catalysis in the serine proteases. *Cold Spring Harb. Symp.* 52 615–621

Hull R 1989 The movement of viruses in plants; *Annu. Rev. Phytopathol.* 27 213–240

Janssen D, Martín G, Velasco L, Gómez P, Segundo E, Ruiz L and Cuadrado IM 2005 Absence of coding region for the helper component proteinase in the genome of a cucumber vein yellowing virus, a whitefly transmitted member of Potyviridae. *Arch. Virol.* 150 1439–1447

Jockusch H, Wiegand C, Mersch B and Rajes D 2001 Mutants of tobacco mosaic virus with temperature-sensitive coat proteins induce heat shock response in tobacco leaves. *Mol. Plant-Microb. Interact.* 14 914–917

Kasschau KD and Carrington JC 1998 A counterdefensive strategy of plant viruses: suppression of posttranscriptional gene silencing. *Cell* 95 461–470

Kasschau KD, Cronin S and Carrington JC 1997 Genome amplification and long-distance movement functions associated with the central domain of tobacco etch potyvirus helper component-proteinase. *Virology* 228 251–262

Kekarainen T, Merits A, Oruetcebarria I, Rajamäki M and Valkonen JPT 1999 Comparison of the complete sequences of five different isolates of *Potato virus A* (PVA), genus Potyvirus. *Arch. Virol.* 144 2355–2366

Kekarainen T, Savilahti H and Valkonen JPT 2002 Functional genomics on *Potato virus A*: Virus genome-wide map of sites essential for virus propagation. *Genome Res.* 12 584–594

Klein PG, Klein RR, Rodríguez-Cerezo E, Hunt AG and Shaw JG 1994 Mutational analysis of tobacco vein mottling virus genome. *Virology* 204 759–769

Klug A and Rhodes D 1987 'Zinc fingers': a novel protein motif for nucleic acid recognition. *Trends Biochem. Sci.* 12 464–469

Korant B, Chow N, Lively M and Powers J 1979 Virus-specified protease in poliovirus infected HeLa cells. *Proc. Natl. Acad. Sci. USA* 76 2992–2995

Kosakovsky Pond SL, Posada D, Gravenor MB, Woelk CH and Frost SD 2006 GARD: a genetic algorithm for recombination detection. *Bioinformatics* 22 3096–3098

Krausslich H and Wimmer E 1988 Viral proteinases. *Annu. Rev. Biochem.* 57 701–754

Kraut J 1977 Serine proteases: structure and mechanism of catalysis; *Annu. Rev. Biochem.* 46 331–358

Larsen RC, Miklas PN, Druffel KL and Wyatt SD 2005 NL-3 K strain is s stable and naturally occurring interspecific recombinant derived from *Bean common mosaic necrosis virus* and *Bean common mosaic virus*. *Phytopathology* 95 1037–1042

Lee KC and Wong SM 1998 Variability of P1 protein of *Zucchini yellow mosaic virus* for strains differentiation and phylogenetic analysis with other potyviruses. *DNA Sequence* **9** 275–293

López-Moya JJ and García JA 2008 Potyviruses; in *Encyclopedia of virology* (eds.) BWJ Mahy and MHV Van Regenmortel (Oxford: Elsevier) pp 313–322

Maiss E, Timpe U, Brisske A, Jelkmann W, Casper R, Himmler D, Mattanovich D and Katinger HWD 1989 The complete nucleotide sequence of *Plum pox virus* RNA. *J. Gen. Virol.* **70** 513–524

Mäkki-Valkama T, Valkonen JPT, Kreuze JF and Pehu E 2000a Transgenic resistance to PVY(O) associated with post-transcriptional silencing of P1 transgene is overcome by PVY(N) strains that carry highly homologous P1 sequences and recover transgene expression at infection. *Mol. Plant-Microb. Interact.* **13** 366–373

Mäkki-Valkama T, Pehu T, Santala A, Valkonen JPT, Koivu K, Lehto K and Pehu E 2000b High level of resistance to *Potato virus Y* by expressing P1 sequence in antisense orientation in transgenic potato. *Mol. Breed.* **6** 95–104

Martin DP, Williamson C and Posada D 2005 RDP2: recombination detection and analysis from sequence alignments; *Bioinformatics* **21** 260–262

Mascotti DP and Lohman TM 1990 Thermodynamic extent of counterion release upon binding oligolysines to single-stranded nucleic acids. *Proc. Natl. Acad. Sci. USA* **87** 3142–3146

Mavankal G and Rhoads RE 1991 *in vitro* cleavage at or near the n-terminus of the helper component protein in the *Tobacco vein mottling virus* polyprotein; *Virology* **185** 721–731

Merits A, Guo D, Järvekülg L and Saarma M 1999 Biochemical and genetic evidence of interactions between *Potato A Potyvirus*-encoded proteins P1 and P3 and proteins of the putative replication complex. *Virology* **263** 15–22

Mukherjee K, Verma Y, Chakrabarti SK and Khurana SM P 2004 Phylogenetic analysis of 5′-UTR and P1 protein of Indian common strain of *Potato virus Y* reveals its possible introduction in India; *Virus Genes* **29** 229–237

Nicklin MJH, Toyoda H, Murray MG and Wimmer E 1986 Proteolytic processing in the replication of polio and related viruses. *Nat. Biotechnol.* **4** 33–34

Oroszlan S and Luftig RB 1990 Retroviral proteinases; *Curr. Top. Microbiol.* **157** 153–185

Oruetxebarria I and Valkonen JPT 2001 Analysis of the P1 gene sequences and the 3′ terminal sequences and secondary structures of the single-stranded RNA genome of *Potato Virus V. Virus Genes* **23** 335–343

Oruetxebarria I, Kekarainen T, Spetz C and Valkonen JPT 2000 Molecular chracterization of *Potato virus V* genomes from Europe indicates limited spatiotemporal strain differentiation. *Phytopathology* **90** 437–444

Palmberg AC 1990 Proteolytic processing of picornaviral polyprotein. *Annu. Rev. Microbiol.* **44** 603–623

Pehu TM, Mäki-Valkama TK, Valkonen JPT, Koivu KT, Lehto KM and Pehu EP 1995 Potato plants transformed with potato virus Y P1 gene sequence are resistant to PVY^{0}. *Am. Potato. J.* **72** 523–532

Pruss G, Ge X, Shi XM, Carrington JC and Vance VB 1997 Plant viral synergism: The Potyviral genome encodes a broad-range pathogenicity enhancer that transactivates replication of heterologous viruses. *Plant Cell* **9** 859–868

Rajamäki ML, Kelloniemi J, Alminaite A, Kekarainen T, Rabenstein F and Valkonen JP 2005 A novel insertion site inside the potyvirus P1 cistron allows expression of heterologous proteins and suggests some P1 functions. *Virology* **342** 88–101

Rao MB, Tanksale AM, Ghatge MS and Deshpande VV 1998 Molecular and biotechnological aspects of microbial proteases. *Microbiol. Mol. Biol. Rev.* **62** 597–635

Rawlings NS and Barrett AJ 1993 Evolutionary families of the peptidases; *Biochem. J.* **290** 205–218

Record MT, Lohman TM and deHaseth PL 1976 Ions effects on ligand-nucleic acid interactions. *J. Mol. Biol.* **107** 145–158

Restrepo-Hartwig MA and Carrington JC 1994 The Tobacco etch potyvirus 6-kilodalton protein is a membrane associated and involved in viral replication. *J. Virol.* **68** 2388–2397

Rice CM and Strauss JH 1990 Production of flavivirus polypeptides by proteolytic processing; *Semin. Virol.* **1** 357–367

Riechmann JL, Laín S and García JA 1990 Infectious *in vitro* transcripts from a plum pox potyvirus cDNA clone. *Virology* **177** 710–716

Riechmann JL, Laín S, García JA 1992 Highlights and prospects of potyvirus molecular biology. *J. Gen. Virol.* **73** 1–16

Rodriguez-Cezero E and Shaw JG 1991 Two newly-detected non-structural viral proteins in poty virus infected cells. *Virology* **185** 572–579

Rodríguez-Cerezo E, Ammar ED, Pirone TP and Shaw JG 1993 Association of the non-structural P3 viral protein with cylindrical inclusions in potyvirus-infected cells. *J. Gen. Virol.* **74** 1845–1949

Rodríguez-Cerezo E, Findlay K, Shaw JG, Lomonossoff GP, Qiu SG, Linstead P, Shanks M and Risco C 1997 The coat and cylindrical inclusion proteins of a potyvirus are associated with connections between plant cells. *Virology* **236** 296–306

Roossinck MJ 2003 Plant RNA virus evolution. *Curr. Opin. Microbiol.* **6** 406–409

Salvador B, Saénz P, Yangüez E, Quiot JB, Quiot L, Delgadillo MO, García JA and Simón-Mateo C 2008 Host-specific effect of P1 exchange between two potyviruses. *Mol. Plant Pathol.* **9** 147–155

Shi Y, Chen J, Hong X, Chen J and Adams MJ 2007 A potyvirus P1 protein interacts with the Rieske Fe/S protein of its host. *Mol. Plant Pathol.* **8** 785–790

Skalka AM 1989 Retroviral proteases: first glimpses at the anatomy of a processing machine. *Cell* **56** 911–913

Soumounou Y and Laliberté JF 1994 Nucleic acid-binding properties of the P1 protein of *Turnip mosaic potyvirus* produced in *Escherichia coli*. *J. Gen. Virol.* **75** 2567–2573

Sprang S, Standing T, Fletterick RJ, Stroud RM, Finer-Moore J, Xoung N, Hamlin R, Rutter WJ, and Craik CS 1987 The three-dimensional structure of Asn^{102} mutant of trypsin: role of Asp^{102} in serine protease catalysis. *Science* **237** 905–909

Strauss JH and Strass EG 1990 Alphavirus proteases. *Semin. Virol.* **1** 347–356

Susaimuthu J, Tzanetakis IE, Gergerich RC and Martin RR 2008 A member of a new genus in the Potyviridae infects Rubus. *Virus Res.* **131** 145–151

Tan Z, Wada Y, Chen J and Ohshima K 2004 Inter-and intralineage recombinants are common in natural population of *Turnip mosaic virus*. *J. Gen. Virol.* **85** 2683–2696

Tatineni S, Ziems AD, Wegulo SN and French R 2009 *Triticum mosaic virus*: a distinct member of a family potyviridae with an unusually long leader sequence. *Phytoplathology* **99** 943–950

Tavert-Roudet G, Ravelonandro M, Bachelier JC and Dunez J 1998 Transgenic *Nicotiana benthamiana* plants containing the P1 gene of *Plum pox virus* are resistant to virus challenge. *Eur. J. Plant Pathol.* **104** 103–107

Tordo VMJ, Chachulska AM, Fakhfakh H, Romancer ML, Robaglia C and Astier-Manifacier S 1995 Sequence polymorphism in the 5′NTR and n the P1 coding region of *Potato virus Y* genomic RNA. *J. Gen. Virol.* **76** 939–949

Valli A, Martin-Hernández AM, López-Moya JJ and García JA 2006 RNA silencing suppression by second copy of the P1 serine protease of *Cucumber vein yellow ipomovirus* (CVYV), a member of the family Potyviridae that lacks the cysteine protease HC-Pro. *J. Virol.* **80** 10055–10063

Valli A, López-Moya JJ and García JA 2007 Recombination and gene duplication in the evolutionary diversification of P1 proteins in the family Potyviridae. *J. Gen. Virol.* **88** 1016–1028

Valli A, Dujovny G and García JA 2008 Protease activity, self interaction and small interfering RNA binding of the silencing suppressor P1b from *Cucumber vein yellowing ipomovirus*. *J. Virol.* **82** 974–986

Vance VB, Moore D, Turpen TH, Bracker A and Hollowell VC 1992 The complete nucleotide sequence of pepper mottle virus genomic RNA: comparison of the encoded polyprotein with those of other sequences potyviruses. *Virology* **191** 19–30

Vance VB, Berger PH, Carrington JC, Hunt AG and Shi XM 1995 5′ proximal potyviral sequences mediate *Potato virus X* potyviral synergistic disease in transgenic tobacco. *Virology* **206** 583–590

Verchot JM and Carrington JC 1995a Debilitation of plant potyvirus infectivity by P1 proteinase-inactivating mutations and restoration by second-site modifications. *J. Virol.* **69** 1582–1590

Verchot JM and Carrington JC 1995b Evidence that the Potyvirus P1 proteinase functions in trans as an accessory factor for genome amplification. *J. Virol.* **69** 3668–3674

Verchot JE, Koonin V and Carrington JC 1991 The 35 kDa protein from the N-terminus of the potyviral polyprotein functions as a third virus-encoded proteinase. *Virology* **185** 527–535

Verchot JM, Herndon KL and Carrington JC 1992 Mutational analysis of the Tobacco Etch Potyviral 35-kDa proteinase: Identification of essential residues and requirements for autoproteolysis. *Virology* **190** 298–306

Ward CW and Shukla DD 1991 Taxonomy of potyviruses: current problems and some solutions. *Intervirology* **32** 269–296

Wellink J and van Kammen A 1988 Proteases involved in the processing of viral polyproteins. *Arch. Virol.* **98** 1–26

Whitham S, Quan S, Chang HS, Cooper B, Estes B, Zhu T, Wang X and Hou YM 2003 Diverse RNA viruses elicit the expression of common rests of genes in susceptible *Arabidopsis thaliana* plants. *Plant J.* **33** 271–283

Wisler GC, Porciful DE and Hiebert E 1995 Characterization of the P1 protein and coding region of the *Zucchini yellow mosaic virus*. *J. Gen. Virol.* **76** 37–45

Xu DL, Zhou DH, Xie YJ, Mock R and Li R 2010 Complete nucleotide sequence and taxonomy of *Sugarcane streak mosaic virus*, member of a novel genus in the family Potyviridae. *Virus Genes* **40** 432–439

MS received 28 September 2010; accepted 10 Febuary 2011

*e*Publication: 14 March 2011

Corresponding editor: INDRANIL DASGUPTA

Rohožková J, Navrátil M, Šebela 2011 Novel biochemical properties of P1 peptidase of Pea seed-borne mosaic virus. Virol J. *Manuscript prepared to submission in:* Virology Journal, c/o BioMed Central, United Kingdom

A novel biochemical properties of P1 peptidase of Pee seed-born mosaic virus

Jana Rohožková[1§], Milan Navrátil[1], Marek Šebela[2]

[1]Faculty of Science, Palacký University in Olomouc, Dept. of Molecular biology and Genetics, Šlechtitelů 11, 783 71, Olomouc-Holice, Czech Republic

[2]Faculty of Science, Palacký University in Olomouc, Dept. Biochemistry, Šlechtitelů 11, 783 71, Olomouc-Holice, Czech Republic

[§]Corresponding author

Email addresses:

JAR: rohozkova@img.cas.cz

MIN: milan.navratil@upol.cz

MAS: marek.sebela@upol.cz

Abstract

Background

Pea seed-borne mosaic virus (PSbMV) a member of the genus Potyvirus, can be transmitted by seeds and causes serious lost in the legume crop yields. Its genome is consisted of 3 peptidases – P1, HC-Pro and NIa. Interestingly only P1 peptidase localized on the N terminal part of the viral polyprotein still remains without any concrete function influencing viral infection. However P1 peptidase is characterized as the most variable protein (differing in the size as well as in the conserved motifs) all over the Potyviriadae proteins. We have found that this protein can act in the truncated form during viral infection and is consisting of originally described C2H2 zinc finger motif. This novel found qualities can help to clarify its true activity during viral infection.

Results

We have expressed truncated part of P1 (Δ33-143) for immunizing of rabbit to produce polyclonal antibody against P1. Using this antibody we have found, that P1 (42 kDa) can act in the truncated form (30 kDa) during viral infection. Its interacting partner can be CI and Rubisco. Immunohistochemical localization of P1 in the protoplasts obtained from infected pea plant demonstrated localization of P1 in the cytoplasm as well as in the nucleus. We have described C2H2 zinc finger motif (ZnF), which is non-standard of that presented in the eukaryotic world. We demonstrate *in silico* predicted protein models of found ZnF.

Conclusions

P1 peptidase is acting during infectious cycle in the truncated form in till present not described pathway. In its activity originally found C2H2 ZnF can be crucial. We described this motif for the first time in the viral world. After identification of this ZnF we found P1 of other 6 viruses (4 Potyvirus, 2 Tritimovirus) which are containing it too. Predicted protein models are confirming canonical secondary structure of found ZnF.

Background

Pea seed born mosaic virus (PSbMV) is a member of genus Potyvirus, family *Potyviridae*. Infection by PSbMV belongs to the group of important legume diseases transmitted per seeds and causes mosaics, stunting and serious crop lost varying between 11-82% [1]. P1 protein of family *Potyviridae* is a serine peptidase localized at the N terminal part of viral genome cleaving HC-Pro and itself from the rest of polyprotein during the viral infection. More of the P1 peptidase quality was its RNA binding activity and enhancing quality of HC-Pro activity in silencing suppression [2, 3].

Peptidases are no uncommon in the microbial world. Their importance is in their involvement in the processing of viral proteins causing certain fatal diseases such as AIDS and cancer. There are serine, aspartic and cystein peptidase found in various viruses [4]. In this manner of studied problem focusing the study on the plant viral peptidases helps us not just safely investigate the problematic of viral peptidase affecting host plants but consequently apply found homology to understand the activity of RNA viruses; not only in the realm of plants, but also with the wider application of those viruses concerning and affecting humans.

Till present real function of potyviral P1 peptidase still remains unclear without any significance of its influence to the infectious cycle of the virus. Here we suggest a novel biochemical qualities, sub-cellular localization and probable function dedicated from *in silico* analysis and modeling of this mysterious potyviral protein.

Methods

Ethic statement

All animals used in this study were housed in accordance with guidelines from legislative rules of the European Union, supervised by the control authority of the Czech Republic.

Viral isolate and P1 peptidase

P1 coding region has been used form PSbMV isolate PSB204CZ (GenBank No JN246531). This isolate was mechanically inoculated on host plant *Pisum sativum* cultivar Merkur. Inoculated plants have been kept in phyto-chamber with standard photoperiod (16h day/8h night)and temperature (22°C during the day and 18°C during the night). Positivity of viral transfer has been checked by ELISA test using monoclonal antibody against capsid protein (LOEWE Biochemica) after 14 days after inoculation. Isolation of total RNA has been provided by RNeasy Plant mini Kit (QIAGen) according to the manufacturer's recommendations. We synthesized the cDNA using Robust I RT-PCR Kit (Finnzyme) due to manufacturer's recommendations by primers bordering P1 coding region designed on the DPD1 sequence (*Gene bank NP734419*). Total P1 coding region was amplified with forward (5´-ATGTCAACACTAGTTTGC-3´) and reverse (5´-ATAGTGGTCAATTTGAAA-3´) primers using Go*Taq* DNA polymerase (Sigma). Reached P1 coding region sequence in length of 1141 bp was cloned into the pGem-T vector (Promega) and JM109 (Promega) competent cells have been transformed with this construct consequently. Coding region was checked by sequencing. Cycle sequencing reaction was provided using pUC/M13 forward and reverse primers (Promega) with annealing site in vector. For cycle sequencing reaction we used ABI BigDye v. 1.1 sequencing terminator kit (Applied Biosystems). Reached sequences was analyzed and completed by SeqMan^{TM} Software (DNA Star).

Bacterial expression of P1 Δ33-143 peptide

Short peptide P1 Δ33-143 was designed with software with surface display prediction with the highest probability of antigenicity within the whole P1 amino acid sequence. P1 Δ33-143 coding region was amplified by forward

(5´-GAGCTCTGTACTCAGTGTGACATGG-3´) and revers (5´-GGTACCAGCTTTCCGCATGAAAA-3´) primers with linked restriction places SacI (forward) and KpnI (revers; both underlined in the sequence) for re-cloning into the bacterial expression vector pQE30 (QIAGen). Competent cells *E. coli* M15[pRep4] (QIAGen) was transformed with reached constructs. Cell culture cultivated in Luria-Bertain medium with addition of antibiotics ampiciline (100 $mg.ml^{-1}$) and kanamycine (25 $mg.ml^{-1}$) in exponential phase of growth was induced with isopropyl-1-thio-β-D-galactopyranozide to final concentration 1 $mmol.l^{-1}$ and growth for additional 4 hours in 37°C. Then we prepared total cell lyzate by collecting the total amount of cells from medium by centrifugation (4000 x g; 20 min). We disturbed the cells by repetitive freezing (-20°C) and thawing (room temperature) the pellet. We re-suspended the cells in lyzis buffer (50 $mmol.l^{-1}$ $NaH_2PO_4.2H_2O$, 300 $mmol.l^{-1}$ NaCl, 10 $mmol.l^{-1}$ imidazole of pH 8.0). Into this suspension we added lyzozyme to final concentration 1 $mg.ml^{-1}$ briefly shake and incubate the tubes in 4°C for 1 hour. Then we sediment the cell debris by centrifugation (10 000 x g; 30 min) and use the supernatant.

Protein purification and antibody production

We have loaded 2 ml of total cell lyzate onto 0.5 cm thick 10% concentrated SDS-polyacrylamide gel [5]. After separation we stained the gel (GelCode® Blue Stain Reagent, Thermo Scientific) and protein of the size 14 kDa was dissected. Gel slice was minced on the small peaces and inserted into dialysis tube (Spectra/Por, Spectrum Laboratorys) with molecular cut-off 6000-8000 Da. Electro elution environment was reached using dialysis buffer (50 $mmol.l^{-1}$ NaH_2PO_4, 1% SDS) (protocol from Technical resource: *Purify proteins from polyacrylamide gels.* Pierce Biotechnology). Electro elution ran by 100 mA during 8 hours in 4°C. We concentrated the protein from dialysis buffer using Microcon columns (Milipore) with molecular cut-off 10 000 Da. This method allows also de-salting the samples and decontaminating the SDS from samples further used for rabbit immunization. Purified protein solution was homogenized in Freud's complete adjuvant. Initiating volume 0.5 mg was applied intramuscularly two times in period of 7 days. Next 3 immunizing applications were given to rabbits as 0.2 mg of protein homogenate in Freud's adjuvant intradermally in 14 and 28 days. Last immunizing injection containing 0.5mg of purified protein was applied 42nd day from the start. Seven days past last immunization rabbits were bled. The produced rabbit polyclonal serum was preservated by sodium azide added to 0.02%. Isolation of rabbit IgG

was provided using Montage® Antibody Purification Kit (Milipore) due to the manufacturer's recommendations.

Protein amino acid sequencing

Bands containing proteins of rights size to be analyzed were excised from Coomassie-stained SDS-PAGE gels. In-gel digestion by raffinose-modified trypsin was conducted according to a published protocol [6] with an additional reduction/alkylation (2-mercaptoethanol/iodoacetamide) prior to the addition of protease [7]. MALDI matrix has been composed from α-Cyano-4-hydroxycinnamic acid (2 $mg.ml^{-1}$) in 2.5% acetonitrile (v/v)/ trifluoroacetic acid (2:1, v/v) [8]. Volume of 5 μl digest and equal volume of matrix solution were premixed in a test tube. Then 1 μl of the mixture was pipetted on an MSP AnchorChip 600/96 target (Bruker Daltonik) and allowed to dry at ambient temperature. Positive-ion mass spectra were measured in the reflectron mode on a Microflex LRF20 MALDI-TOF mass spectrometer (Bruker Daltonik) equipped with a microScout ion source and a 337-nm nitrogen laser. Mass spectra were accumulated from 100-200 shots at a laser repetition rate of 10 Hz; the examined *m/z* range was 500-5,000. The instrument was calibrated externally using a mixture of peptide standards.

The acquired spectra were processed by flexAnalysis 2.4 and Biotools 3.0 software (Bruker Daltonik). Database searches were done against protein sequence databases Swiss-Prot (ver. 56.6) and NCBInr using the program Mascot Server 2.2 (Matrix Science, London, UK) installed on a local PC. Oxidation of methionine as an optional modification and one enzyme missed cleavage were chosen for all searches performed without taxonomic restriction; a mass tolerance of 150 ppm was allowed.

Co-immunoprecipitation

Leafs from negative and infected pea plant was minced in liquid nitrogen. Then we immediately add denaturating lysis buffer (50 mM Tris-HCl pH 8, 150 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS) containing inhibitor of endogenous proteases (cOmplete, Mini, EDTA-free, Protease Inhibitor Cocktail, Roche) into the powder sample. Furthermore we sonicated the sample. We used rabbit immune serum or purified IgG as well for protein immunoprecipitation. Than we incubated the leaf lysate containg the antibody by

rotation overnight in 4°C. Consequently we add protein-A Sepharose beads in amount regarded by manufacturer and incubated the sample next 1 hour in 4°C. Afterward we discard the supernatant, washed twice the beads with the same lysis buffer and elute the precipitated protein by sample buffer (50 mM Tris-HCl pH 6.8, 2% SDS, 10% glycerol, 1% β-mercaptoethanol, 12.5 mM EDTA, 0.02 % bromophenol blue) and boiling for 5 min. Isolated proteins were loaded on 10% concentrated SDS-polyacrylamide gel [5] and analyzed by MALDI-TOF. For immunodetection of precipitated P1 peptidase we transfer the proteins on nitrocellulose membrane (Amersham Bioscience). P1 peptidase was detected by immune serum or isolated IgG.

In silico secondary structure modeling

Secondary structure was computationally calculated from the 277 amino acid long P1 coding region (ΔP1). Using I-TASSER Online [9, 10], an internet service for protein structure and function predictions. 3D models are built based on multiple-threading alignments by LOMETS and iterative TASSER assembly simulations. Model display was provided by PyMOL (DeLano Scientific LLC 2006).

Protoplast isolation and Immunocytochemistry

Imunocytochemistry was provided on isolated and purified protoplasts from health and PSbMV virus infected pea plant. 100mg of leafs were chopped on small pieces and macerated in 27°C over night in enzyme solution (1% Cellulase Omozuka R-10 , 0.25% Macerozyme R-10 (both Duchefa) diluted in protoplast washing medium) with gentle agitation. Separate have separated the protoplast by filtration using 0.45 µM nylon membrane. We left to sediment the protoplast by centrifugation (1000 x g; 10 min at 4°C). After removing the supernatant we resuspended the protoplasts in washing medium (0.2 $mmol.l^{-1}$ KH_2PO_4, 1 $mmol.l^{-1}$ KNO_3, 10 $mmol.l^{-1}$ $CaCl_2$ $.2H_2O$, 1 $mmol.l^{-1}$ $MgCO_4$ $.7H_2O$, 0.5 $mol.l^{-1}$ mannitol pH 5,6). 60µl of fixing solution (1% paraformaldehyde, 0.2 $mol.l^{-1}$ HEPES, 0.15% Triton X-100, protease inhibitors) was dropped directly on the cover slips. 20 µl aliquot of protoplasts was places on each slide and incubated for 2 hours in humid chamber. We rinsed the slides gently in injection water and let air dried. Cover slips were stored for 1 month in -80°C without lost of antigenicity. For immunocytochemical detection of P1 peptidase the protoplasts were blocked in 1%

bovine serum albumin. The samples were incubated with anti-P1 Δ33-143 polyclona antibody (1:100 dilution in PBS-T) and afterwards with secondary antibody Alexa Fluor® 555 donkey anti-rabbit IgG (Invitrogen). For blocking the signal coming from un-specific interaction of antibody with Rubisco we blocked the slides with 1% bovine serum albumin with diluted chicken polyclonal antibody against Rubisco (AbCam). Acquisition of pictures was done by the Leica DMI 6000 CS TCS SP5 Spectral Scanning Confocal System.

Results and Discussions

P1 33-143 bacterial expression and antibody production

We have expressed *in vitro* P1 33-143 and after cutting the right band from the gel we electro-eluted the protein into the phosphate buffer. P1 33-143 was purified from the phosphate saline buffer and afterwards used for immunization of the animal. We used immune rabbit serum as well as purified rabbit IgG against Δ33-143. We immunoprecipitate viral P1 from the lyzate of infected pea plant to check the specificity of produced antibody. Reached bands were analyzed by MALDI-TOF with surprising result. Antibody P1 Δ33-143 co-precipitated following proteins: CI viral protein (65 kDa), Rubisco (55 kDa), non-identified protein (45 kDa) and 30 kDa protein (Fig.1). Sequence of this 30 kDa protein was found to cover coding region between 80-277 amiono acids (aa) of viral P1 peptidase (Fig.2). Found fragments are covering 49.8% of total P1 of PSbMV. Surprisingly 80-277 aa region size was predicted as 22 kDa due to the amino acid sequence. For further *in silico* analyzis we used slightly expanded region 1-280 aa. Arbatova *et al.* reached similar result [11]. They co-localized P1 and CI via immunogold labeling on ultra-thin sections of tobacco plant infected by PVY. Rubisco was further identified as non-specifically bound protein precipitated also in the negative lyzate from healthy pea plant (Fig.1).

P1 localization in infected pea plant protoplasts

We performed the immunofluorescence using polyclonal rabbit antibody against P1 Δ33-143 on the isolated protoplasts taken from the infected pea plant. Non-specific interaction of rabbit

IgG with plant Rubisco was detected as diffused signal spread all over the cytoplasm (Fig.3A). This diffused signal avoids the identification of P1 specific localization. For confirmation of P1 peptidase localization pattern we blocked fixed coverslips with plant protoplasts with diluted antibody against plant Rubisco. Blocking of non-specific interaction of P1 Δ33-143 to Rubisco was proven by western blot (Fig.3C) and used afterwards for immunofluorescence. Missing signal in compare to the non-blocked IgG immunofluorescence was identified as P1 specific localization pattern (Fig.3B). P1 obviously occupy cell cytoplasm excluding membranous structures (as Endoplasmic reticulum or Golgi apparatus) as well as cytoplasmic membrane. However we have detected clear signal in the cell nucleus (Fig.3B). We haven't found in the P1 coding region any nuclear localization signal (*unpublished data*). To confirm concrete P1 nuclear localization or trafficking possibly via host protein requires more experiments.

Alternative splicing and identification of ZnF

P1 peptidase of PSbMV is 42 kDa protein according to its amino acid sequence. Size of the found protein presents the homology with this protein is about 12 kDa shorter. However precipitated 30 kDa protein has high homology within found short fragments spread between 80-277 aa of P1 coding region. This fact led us to suggestion of alternative splicing of P1 peptidase after its auto-cleavage from the viral polypeptide. However we did not find any alternative restriction places which could explain truncation of P1 regarding to found fragments or respecting final size 30 kDa.

Moreover we used SMART: Sequence analysis database [12, 13] for prediction of active domains in truncated P1: ΔP1 (1-277) avoiding detection of serine peptidase domain specific for cleavage of P1 itself from the rest of viral polyprotein. In N-terminal part of defined amino acid sequence was surprisingly localized Cys2His2 Zinc finger motif (ZnF-C2H2; SMART accession No SM00355). Identification was dedicated according to the E-value near to the 1.22e+03 (Hidden Markov Model). ZnF in P1 was predicted in formula **C[32]**-X_2-**C[35]**-X_3-F-X_{10}-**H[50]**-X_{20}-**H[70]** regarding the positioning in P1 aa sequence. This motif mostly resembles the eukaryotic form of C2H2 ZnF. After identification of this P1 ZnF motif, we screened available potyviral P1 sequences. Interestingly we found a number of homologue in other potyviral P1s (Table.1). They are sharing aa sequence homology and each of them localizes this motif into the proximity of N-terminus.

In silico 3D molecule modeling

Topology of possible alpha helices, beta-sheets and coils regions was in the whole amino acid sequence predicted by Protean (Lasergene DNASTAR®). This software has indicated probable regions with presence of alpha helix or beta sheet with intercalated coils. Due to this prediction we choose the sequence correlating with the MALDI covered motif, final size of *in vivo* found protein (30 kDA) and topology of the polypeptide. Consequently for the 3D structuring of P1 30 kDA we used aa sequence in region 1-277 with predicted size of polyprotein 30,5 kDa (Fig.4A). I-TASSER software [9, 10] modulated 5 possible models. We have picked Model No.4 with the lowest C-score (-3.59). C-score is a confidence score for estimating the quality of predicted models created by I-TASSER (calculating due to the significance of threading template alignments and the convergence parameters of the structure assembly simulations). Total molecule model has also confirmed position of the ZnF C2H2 motif (Fig.4B).

Conclusions

P1 localizes to cytoplasm and nucleus too

We detected P1 peptidase spread all over the cytoplasm using P1 Δ33-143 antibody. Confocal microscopy allows us to detect the presence of this protein very clearly in the nucleus too. However explanation of this presence has to be proved by further experiment, mainly precisely prepared co-immunoprecipitations from the nuclear fractions. We can suggest that this localization can be connected with P1s possible activity in overcoming of host defense mechanisms (P1/HC-Pro) or other function affecting expression of host genes during viral early infection and spreading as indirect stimulation of replication of viral RNA by suppression of host transcriptional activity [14]. Till present there was done localization experiment just once [11]. Immunogold localization confirmed that P1 is spread in the cytoplasm but is excluded from cell wall and plasmodesmata. Surprisingly they did not detect any significant signal in the cell nucleus.

P1 interacts with CI

Possible interaction between P1 and CI was described also previously [11]. Explanation of this finding was a suggestion, that 1) all viral proteins are synthesized near cytoplasmic inclusions, which are usually associated with rough endoplasmic reticulum; 2) cytoplasmic inclusions are indeed sites of potyviral RNA replication. Mentioned interaction can be consistent with the hypothesis of Verchot and Carrington [14] that P1 peptidase could participate in virus replication acting in *trans* accessory or regulatory factor stimulating genome amplification - directly in the RNA replication process by interacting with enzymatic components of replication complexes or with RNA sequence as well as indirectly in genome amplification by stimulating translation of viral RNA [14]. Protocol used for preparation of protein lyzate, which sufficiently disrupts cytoplasmic inclusion and disaggregate most membranous structures because denaturizing conditions are supporting mostly the theory about RNA replication. Found ZnF motif, which possibly binds to RNA is complementing this opinion. Moreover P1 peptidase was found to bind ss and ds RNA (under *in vitro* conditions) as well as ss- but not dsDNA [15].

P1 acts in truncated form in vivo

Identification of shortened P1 peptidase was newly described. What is the phenomenon of the truncation can be explained by the phylogeneticaly most variable part of P1 peptidase all over the potyviruses presented in the N terminal part of this protein. This theory support also fact, that 30 kDa is the size of the smallest P1 peptidase in this family.

Viral proteins are multifunctional. P1 peptidase however till present not described in details is not an exception. Its peptidase activity cleaving itself form the rest of polyprotein is abundant for HC-Pro activity [16]. Mutation analysis and truncations of P1 peptidase showed that affecting the C-terminus is abolishing infectious activity of the virus whereas total deletion of most variable N terminal part has just little affect on the viral quality [17]. Crucial step in this point of view is releasing the P1 from the HC-Pro independently of the proteolytically cleaved site [16].

P1 contains non-canonical C2H2 ZnF

Recently, the first putative prokaryotic C2H2 ZnF domain has been identified in a transcriptional regulator (Ros protein) from *Agrobacterium tumefaciens* [18]. This suggests

that classical ZnF domain, previously thought to be exclusively confined to the eukaryotic organisms, could be widespread to the living kingdom from eukaryotic, both animal, plant, to prokaryotic. A single ZnF domain itself is not sufficient for high affinity binding to a specific DNA target sequence. Nevertheless there are some hints confirming that the single ZnF domain is capable of sequence-specific DNA binding when flanked by basic region [18, 19, 20]. Bacterial ZnF contains three histidine residues and the 9 aa region between the second cystein and the first histidine is shorter than the canonical 12 aa spacer observed in eukaryotic ZnF. After identification of this P1 ZnF motif, we screened available potyviral P1 sequences. There has been found a number of homologue in other potyviral P1s (Table.1). They are sharing aa sequence homology and each of them localizes this motif into the proximity of N-terminus. Find conserved Cys and His residues were previously mentioned by Valli *et al.* ([21]; *supplement figures S5*).

However it is not possible to dedicate strict formula for localization of novel viral ZnF because the spacer between Cys and His varies in different P1. We suggest that P1 peptidase is containing some form of ZnF possibly detectable also in other viral proteins. To proclaim the real activity of this motif and its involvement in some regulative processes during viral infection, it has to be confirmed by further experiments.

References

1. Torok VA: **Discriminating between isolates of PSbMV using nucleotide sequence polymorphisms in the HC-Pro coding region.** *Plant Dis* 2006, **91:**490-496.
2. Pruss G, Ge X, Shi XM, Carrington JC, Vance VB: **Plant viral synergism: The Potyviral genome encodes a broad-range pathogenicity enhancer that transactivates replication of heterologous viruses.** *Plant Cell* 1997, **9:**859–868
3. Valli A, Martin-Hernández AM, López-Moya JJ, García JA: **RNA silencing suppression by second copy of the P1 serine protease of Cucumber vein yellow ipomovirus (CVYV), a member of the family Potyviridae that lacks the cysteine protease HC-Pro.** *J. Virol.* 2006, **80:** 10055–10063.
4. Rawlings NS, Barrett AJ: **Evolutionary families of the peptidases**. *Biochem. J.* 1993, **290:** 205–218.
5. Laemmli UK: **Cleavage of structural proteins during the assembly of the head of bacteriophage T4.** *Nature* 1970, 227: 680-685.
6. Šebela M, Stosova T, Havlis J, Wielsch N, Thomas H, Zdráhal Z, Shevchenko A. **Thermostable trypsin conjugates for high-throughput proteomics: synthesis and performance evaluation.** *Proteomics* 2006, **6:**2959-2963.
7. Shevchenko A, Tomas H, Havlis J, Olsen JV, Mann M. **In-gel digestion for mass spectrometric characterization of proteins and proteomes.** *Nat Protoc* 2006, **1:**2856-2860.

8. Thomas H, Havliš J, Peychl J, Shevchenko A. **Dried-droplet probe preparation on AnchorChip™ targets for navigating the acquisition of matrix-assisted laser desorption/ionization time-of-flight spectra by fluorescence of matrix/analyte crystals.** *Rapid Commun Mass Spctrom* 2004, **18:**923-930.
9. Roy A, Kucukural A, Zhang Y: **I-TASSER: a unified platform for automated protein structure and function prediction.** *Nar Protoc* 2010, **5**:725-738.
10. Zhang Y: **I-TASSER server for protein 3D structure prediction.** *BMC Bioinformatics* 2008, **9:**40.
11. Arbatova J, Lehto K, Pehu E and Pehu T: **Localization of the P1 protein of potato Y potyvirus in association with cytoplasmic inclusion bodies and in the cytoplasm of infected cells.** *J. Gen. Virol.* 1998, **79:**2319–2323.
12. Schultz J, Milpetz F, Bork P, Ponting CP: **SMART, a simple modular architecture research tool: identification of signaling domains.** *Proc Natl Acad Sci USA* 1998 **95:**5857-5864.
13. Letunic I, Doerks T, Bork P: **SMART 6: recent updates and new developments.** *Nucleic Acids Res.* 2009, **37:** 229-232.
14. Verchot JM and Carrington JC: **Evidence that the Potyvirus P1 proteinase functions in trans as an accessory factor for genome amplification.** *J. Virol.* 1995b, **69:**3668–3674.
15. Soumounou Y and Laliberté JF: **Nucleic acid-binding properties of the P1 protein of Turnip mosaic potyvirus produced in Escherichia coli.** *J. Gen. Virol.* 1994, **75:**2567–2573.
16. Verchot JM and Carrington JC: **Debilitation of plant potyvirus infectivity by P1 proteinase-inactivating mutations and restoration by second-site modifications.** *J. Virol.* 1995a, **69:**1582–1590.
17. Rajamäki ML, Kelloniemi J, Alminaite A, Kekarainen T, Rabenstein F and Valkonen JP: **A novel insertion site inside the potyvirus P1 cistron allows expression of heterologous proteins and suggests some P1 functions.** *Virology* 2005, **342:**88–101.
18. Chou AY, Archdeacon J, Kado CI: **Agrobacterium transcriptional regulator Ros is a prokaryotic zinc finger protein that regulates the plant oncogene ipt.** *Proc Natl Acad Sci U S A* 1998 **95:**5293-5298.
19. Dathan N, Zaccaro L, Esposito S, Isernia C, Omichinski JG, Riccio A, Pedone C, Di Blasio B, Fattorusso R, Pedone PV: **The Arabidopsis SUPERMAN protein is able to specifically bind DNA through its single Cys2-His2 zinc finger motif.** *Nucleic Acids Res* 2002 **30:**4945-4951.
20. Malgieri G, Russo L, Esposito S, Baglivo I, Zaccaro L, Pedone EM, Di Blasio B, Isernia C, Pedone PV, Fattorusso R: **The prokaryotic Cys2His2 zinc-finger adopts a novel fold as revealed by the NMR structure of Agrobacterium tumefaciens Ros DNA-binding domain.** *Proc Natl Acad Sci USA* 2007 104:17341-17346.
21. Valli A., López-Moya JJ, García JA: **Recombination and gene duplication in the evolutionary diversification of P1 proteins in the family.** *Potyviridae*. *J. Gen. Virol.* 2007, 88: 1016-1028.

Figures

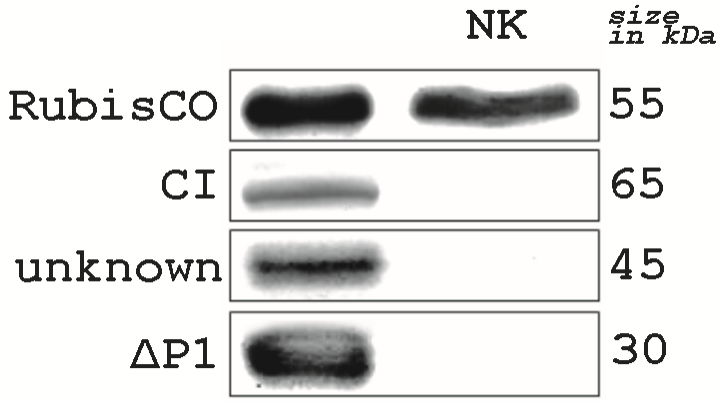


Figure 1 - Co-immunoprecipitation using P1 Δ33-143 from total cell lysate

Figure legend text We co-immunoprecipitated ΔP1 (30 kDa), CI (45 kDa), Rubisco (55 kDa) and another un-known protein (45 kDa). We identified Rubisco also in the lyzate prepared from the healthy pea plant. We dedicate this interaction as non-specific.

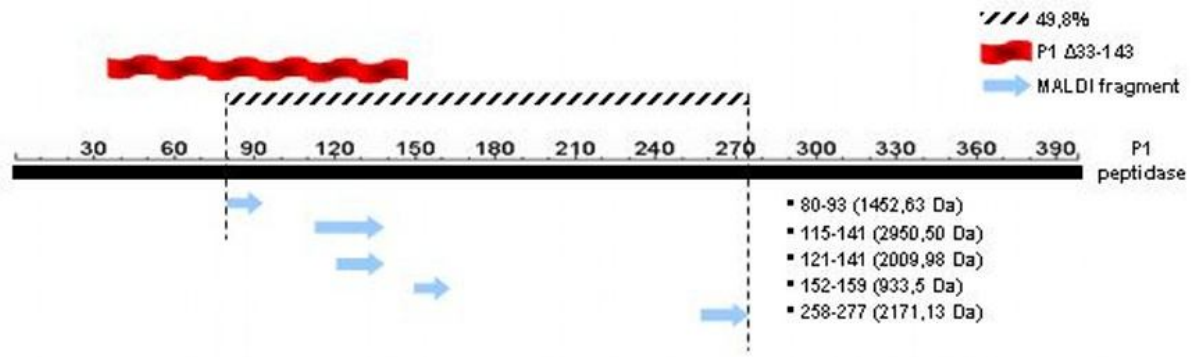


Figure 2 - Scheme of MALDI-TOF identified P1 fragments

Co-immunoprecipitated protein of 30 kDa was analyzed by MALDI-TOF. We identified fragments covering P1 coding sequence in region 80-277, which covers 49,8% of P1. Red tilde represents peptide used for preparation of P1 Δ33-143 rabbit polyclonal antibody.

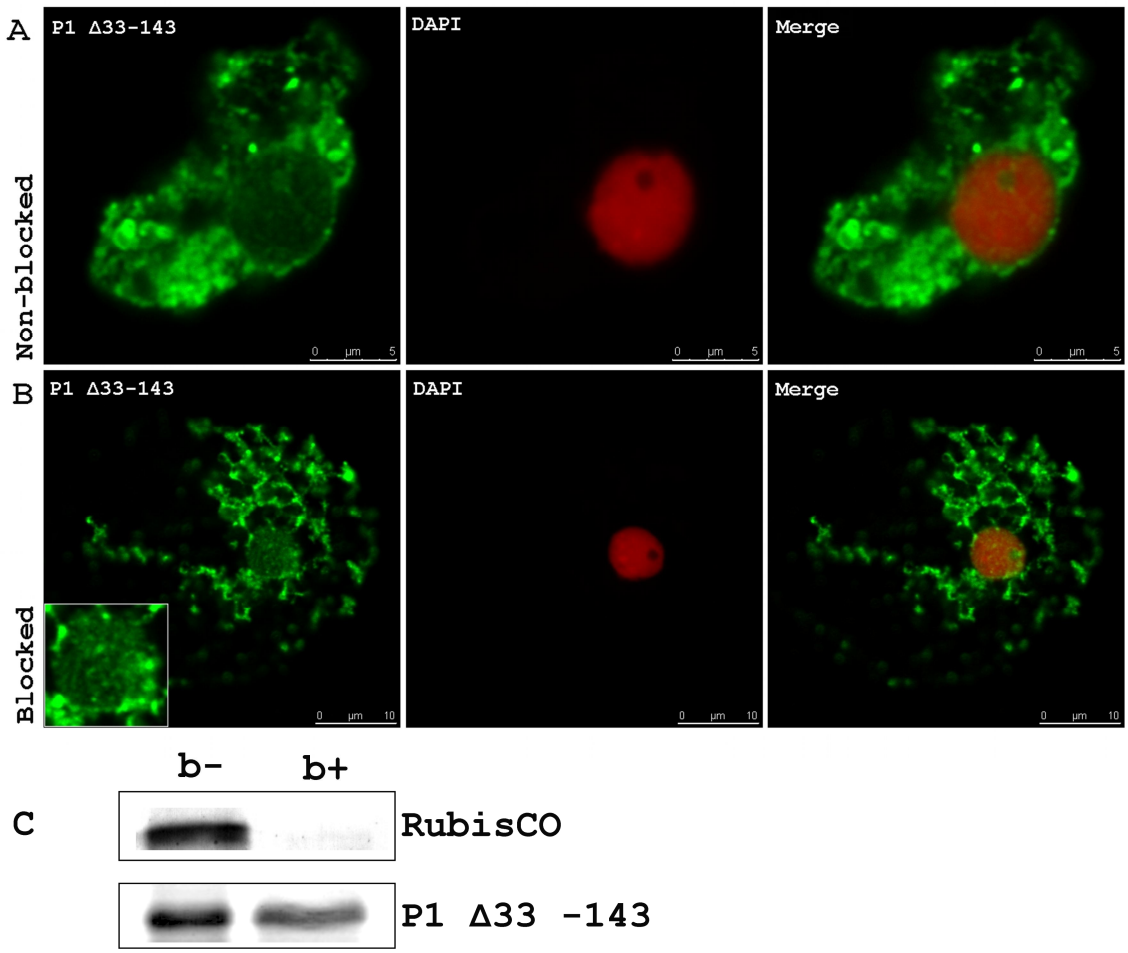


Figure 3 - P1 localization in infected protoplasts

We localized P1 peptidase in protoplasts isolated from infected pea plant. We found Rubisco non-specifically interacting with rabbit polyclonal immunoglobulin (A). For blocking of this non-specificity we used polyclonal antibody against Rubisco. Further localization of P1 peptidase (*green*) was dedicated specific (B). P1 was localized spread over the cytoplasm and clearly in the cell nucleus (B, *detail*). Blocking of non-specific interaction was proved also by western blot (C). Nuclear staining was performed by DAPI (*red*).

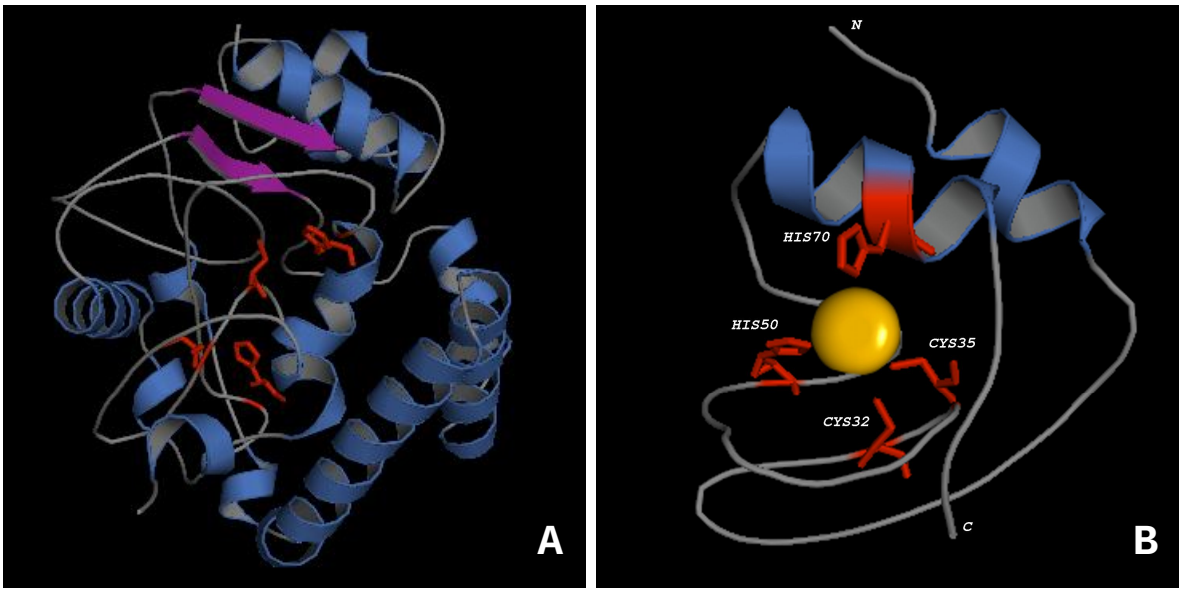


Figure 4 - In silico model of ΔP1 (30 kDa) and C2H2 ZnF

We designed model of ΔP1 (30 kDa) from the sequence 1-277 aa of P1 coding region in *in silico* conditions showing presence of C2H2 ZnF aa (*red*) (A). Detailed view on the ZnF cavity is showing Cys2 (32, 35) and His2 (50, 70) orientation into the Zn (*yellow*) (B).

Tables

25	DADGEYRCT----QCDMGFD	40	44	MARVNHCCDG	54	66	DPIMHLV-DSK	75	PSbMV	DPD1
25	DADGEYRCT----QCDMGFD	40	44	MARVNHCCDG	54	66	DPIMHLV-DSK	75	PSbMV	204
28	DDAVHYHCT----KCNFAFE	43	47	MVRVNHDCDG	57	83	DVFRHLANDNA	93	OYDV	
77	WDDDVYECT----TCSGAFQ	92	95	LDFKEHDCDE	104	161	APIIHECQQEL	171	LYSV	
128	IAAQLYMCP----KCCSASD	143	147	YFDTNHNDSC	156	207	TVPVHDVVQVY	217	BStMV	
78	LAADVFVCG----MCRSSCA	93	89	RSSCAHYRYF	98	98	FIEDHFACEKL	108	ONMV	
25	SNRVNIVCPGHMATCPPPKT	45	41	PPPKTHTYYR	50	47	TYYRHESKKLM	57	ZYMV	

Table 1 - Alignment of C2H2 ZnF motif found in P1 coding region of different Potyviruses

We identified conserved motif C2H2 ZnF in the P1 of PSbMV (isolate 204). We search for homology in the family *Potyviridae* and have found the same conserved motif in next 4 potyviruses (*Pea seed-borne mosaic virus* isolate *DPD1*, *Onion yellow dwarf virus*, *Leek yellow stripe virus* and *Zucchini yellow mosaic virus*) and 2 tritimoviruses (*Brome streak mosaic virus* and *Oat necrotic mottle virus*). Numbers in front and behind the sequence are remarking the position of following aa in the sequence of each P1 peptidase.

6.2 Conferences

Petrusová J 2007 *Molekulárna variabilita oblastí genómu PSbMV zodpovedných za interakciu s génmi rezistencie hrachu.* Sborník příspěvků, Konference doktorandů oboru ochrana rostlin Česká zemědělská Univerzita v Praze. pp 79-83. Praha. ISBN978-80-213-1710-9

Petrusová J and Navrátil M 2008 *Molecular variability of PSbMV genome regions responsible for interactions with the genes of resistance in the pea plant.* Book of Abstracts, The 3rd Conference of the International Working Group on Legume and Vegetables Viruses (IWGLVV), Ljubljana, Slovenia. pp 39. National Institute of Biology, Ljubljana, Slovenia.

Petrusová J 2007 Molekulárna variabilita oblastí genómu PSbMV zodpovedných za interakciu s génmi rezistencie hrachu. Sborník příspěvků z konference doktorandů oboru ochrana rostlin Česká zemědělská Univerzita v Praze. pp 79-83. Praha. ISBN978-80-213-1710-9

Molekulárna variabilita oblastí genómu PSbMV zodpovedných za interakciu s génmi rezistencie hrachu

Petrusová, J.

Katedra buněčné biologie a genetiky, Přírodovědecká fakulta Univerzity Palackého, Šlechtitelů 11, Olomouc 783 71, Česká republika, e-mail: jana_petrus@hotmail.com

Abstrakt

PSbMV je škodlivý, ekonomicky významný vírus (rod *Potyvirus*; čeľaď: *Potyviridae*) infikujúci rastliny hrachu (*P. sativum L*). Izoláty tohto vírusu boli na základe infekčného profilu na selekčnej rade rastlín hrachu zaradené do 4 skupín patotypov: P-1, P-2, P-3 a P-4. U *P.sativum* boli identifikované gény rezistencie označené ako *sbm*-1 a *sbm*-2, ktoré podľa doterajších zistení interagujú pomocou mediátorov s oblasťami VPg a P3+6K1 genómu PSbMV. Táto práca pojednáva o štúdiu variability oblasti VPg, interagujúcou s *sbm*-1 u vybraných izolátov a následné porovnanie nukleotidových sekvencií s referenčnými sekvenciami izolátov DPD1, L1, NEP1 a NY, ako zástupcov jednotlivých patotypov.

Kľúčové slová

Pea Seed-borne Mosaic Virus, VPg, *sbm*-1

Úvod

Pea seed-borne mosaic virus (PSbMV) je ekonomicky významný, semenom prenosný vírus (rod *Potyvirus*; čeľaď *Potyviridae*) infikujúci rastliny hrachu (*Pisum sativum* L.). Nositeľom genetickej informácie je pozitívna jednovláknová RNA o veľkosti 9,8 kb, zložená z kódujúcich oblastí pre 9 neštruktúrnych proteínov a 1 kapsidový protein (Schéma č. 1). Príznakmi ochorenia je zvinovanie listov (Obrázok č. 1), mozaika, deformácie úponkov a oneskorené kvitnutie. U tohto ochorenia sú závažné najmä straty na úrode u infikovaných rastlín, varírujúce od 11% do 82% v závislosti od kultivaru a izolátu vírusu (Torok, 2006).

Izoláty PSbMV boli rozdelené do 4 patotypov P-1, P-2, P-3 a P-4. Toto rozdelenie prebiehalo na základe priebehu ochorenia na líniách hrachu nesúcich gény rezistencie označené ako *sbm*-1 a *sbm*-2 (Alconero, 1986). Oba gény rezistencie fungujú na úrovni špecifického zamedzenia lokálnej a systémovej infekcie vírusom PSbMV. Gén rezistencie *sbm-1* sa javí ako homológ eukaryotického translačného iniciačného faktora nesúceho niekoľko mutácií, označeného aj ako eIF(iso)4E (Gao et al., 2004). Za odpovedajúci gén virulencie u PSbMV bola označená kódujúca oblasť VPg (Hjulsager et al., 2006).

VPg proteín je pri vírusovej infekcii multifunkčný. Jeho prítomnosť je nevyhnutná pri amplifikácii vírusového genómu, ako aj pri pohybe vírusovej častice medzi bunkami a v systémovej infekcii. VPg proteín je kovalentne naviazaný na 5´ konci genómovej RNA, kde slúži ako primer pre naviazanie vírusovej RNA replikázy a bunkových faktorov translácie RNA (Rajamäky and Valkonen, 2001).

Identifikácie a štúdium interakcií VPg - *sbm*-1 a P3+6K1 - *sbm*-2 ako genetických na sebe nezávislých interakcií medzi PSbMV a *P. sativum* nám týmto umožňuje podať bližšie informácie o priebehu vírusovej infekcie vzhľadom k ich zaradeniu do patotypu (Hjulsager et al., 2006).

Keďže experimenty týkajúce sa biologických vlastností izolátov, ale aj molekulovej charakteristiky oblasti P3+6K1 naďalej prebiehajú, uvádzame len čiastkové výsledky vyhodnotenia štúdia variability oblasti VPg genómu *PSbMV* u vybraných izolátov.

Materiál a metódy

Izoláty PSbMV

Analizované boli izoláty vírusu PSbMV: PSB58CZ, PSB117CZ, PSB118CZ, PSB141CZ, PSB178CZ, PSB194CZ, PSB204CZ, PSB262CZ, PSB329CZ, PSBDCZ a PSBECZ, ktoré boli inokulované na hostiteľskej rastline *P. sativum* kultivar Merkur.

PCR detekcia vírusu

Totálna RNA z rastlinného materiálu bola vyizolovaná pomocou RNseasy Plant Mini Kit (Qiagen) podľa pokynov výrobcu. U všetkých vzoriek bola z RNA zrealizovaná RT-PCR pomocou Robust II (Finnzyme). Použité primery boli navrhnuté podľa sekvencie totálneho genómu vírusu PSbMV izolátu DPD1 (D10930). Ich sekvencia a lokalizácia v genóme bola nasledovná: PSbMV VPg For: 5´ TTTATGGCTAAGTGGAACGAACC 3´ lokalizovaný v oblasti 5973-5995; PSbMV VPg Rev: 5´ GACCCCTGTGTTTGTTATTCTAC 3´ lokalizovaný v oblasti 6668-6646. Dĺžka amplyfikátu bola potom 696 bp, pričom táto zahŕňala kódujúcu oblasť pre VPg protein o veľkosti 582 bp.

Klonovanie a sekvenovanie

PCR klonovaného úseku s použitím presahov bola zrealizovaná pomocou primerov M13, s anelačným miestom v klonovacom vektore. PCR produkty boli purifikované z agarózového gélu pomocou QIA Gel Extraction Kit (Qiagen). Produkt bol následne ligovaný do klonovacieho

vektora pGem-T (Promega) podľa inštrukcií výrobcu. Týmto konštruktom boli následne transformované kompetentné bunky *E.coli* JM109 (Promega). Klony boli sekvencované pomocou ABI BigDye v. 1.1 sequencing terminator kit (Applied Biosystems).

Sekvenčná a fylogenetická analýza

Získané sekvecie boli zložené pomocou SeqMan™ softvéru (DNASTAR) a následne porovnané s odpovedajúcimi sekvenciami PSbMV izolátov dostupnými v GenBank pomocou MegAlign (DNASTAR) softvéru. Pre analýzu bola vybraná vlastná kódujúca oblasť pre VPg proteín o veľkosti 582 bp (nukleotidová pozícia 6009-6590 podľa DPD1). Za referenčné izoláty pre jednotlivé patotypy PSbMV boli použité nasledujúce nukleotidové sekvencie: P1 - DPD1 (D10930), P2 – L1 (AJ252242), P3 – NEP1 (AJ311843) a P4 – NY (X89997). Výsledný multiple-alignement bol použitý pre konštrukciu fylogenetického stromu pomocou Neighbor-joining metódy, v programe ClustalW. Získaný strom bol vizualizovaný pomocou programu TreeView (Page *et al.*, 1996).

Výsledky a diskusia

U všetkých vybraných izolátov sme za použitia navrhnutých primerov získali PCR produkt o odpovedajúcej veľkosti 696 bp. Tieto boli následne úspešne osekvenované. Porovnaním nukleotidových sekvencií VPg kódujúcich oblastí vybraných izolátov PSbMV a referenčních izolátov, ako zástupcov jednotlivých biologických patotypov, došlo k ich preukáznemu rozdeleniu do troch skupín (Schéma č. 2).

České izoláty sa na základe podobnosti rozdelili do troch skupín. Prvú skupinu tvoria izoláty PSB262CZ, PSBECZ, PSBDCZ, PSB118CZ, PSB117CZ a PSB141CZ, ktoré vykazujú podobnosť nukleotidových sekvencií s izolátom PSbMV DPD1 ako zástupcom patotypu P1. Izoláty PSB178CZ, PSB194CZ, PSB329CZ a PSB204CZ vykazujú príbuznosť so skupinou 1 a tvoria preukázne samostatnú skupinu 2. Izolát PSB58CZ je výrazne odlišný od ostatných českých izolátov a vykazuje podobnosť izolátom NY, ktorý predstavuje zástupcu patotypu P4, a spoločne tvora 3. skupinu. Typové izoláty ďalších patotypov P2 (L1) a P3 (NEP1) tvoria samostatné vetvy.

VPg je chápaná ako oblasť virulencie v genómu PSbMV (Hjulsager et al, 2006) jej variabilita je v priamom vzťahu k priebehu vírusovej infekcie, má teda vplyv na biologické vlastnosti izolátu/patotypu. Preto na základe získaných výsledkov fylogenetickej analýzy môžme poukazovať na príslušnosť vybraných českých izolátov k patotypom P1, respektíve P4. Doteraz uskutočnené biologické testy toto zistenie potvrdzujú (Tabuľka č.1).

Tieto informácie pre nás môžu v budúcnosti predstavovať užitočný zdroj informácií v procese šľachtenia rezistentných odrôd rastlín hrachu voči PSbMV a ďalším vírusovým ochoreniam tejto úžitkovej rastliny.

Táto práca bola financovaná z projektu MŠMT, MSM 6198959215.

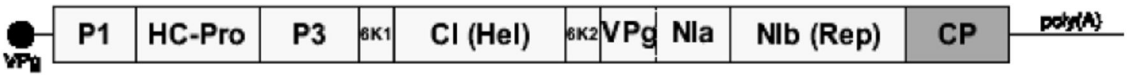
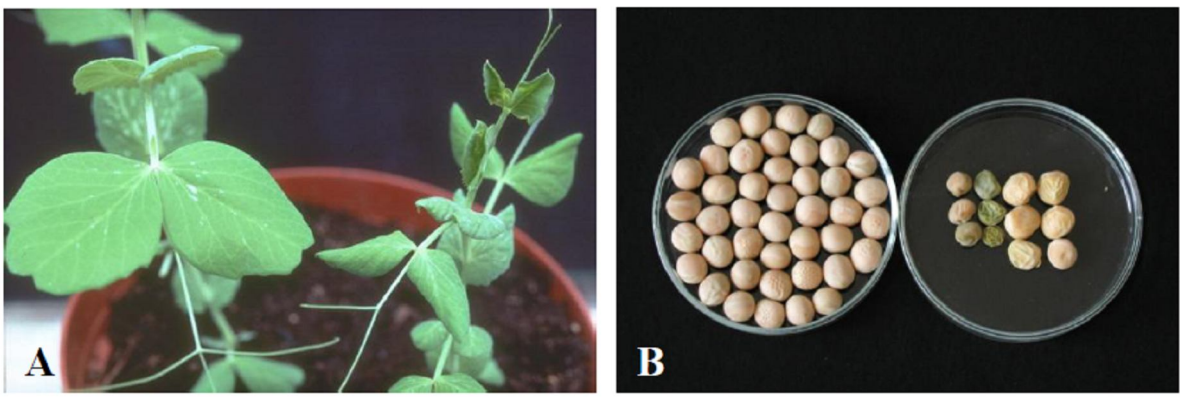


Schéma č. 1 Schematické znázornenie zloženia kódujúcich oblastí pozitívnej jednovláknovej RNA vírusu PSbMV.



Obrázok č. 1A - porovnanie zdravej (ľavá) a infikovanej (pravá) rastliny hrachu s typickým príznakom zvinovania listov; B - porovnanie zdravého osiva (vľavo) a osiva infikovanej rastliny (vpravo)

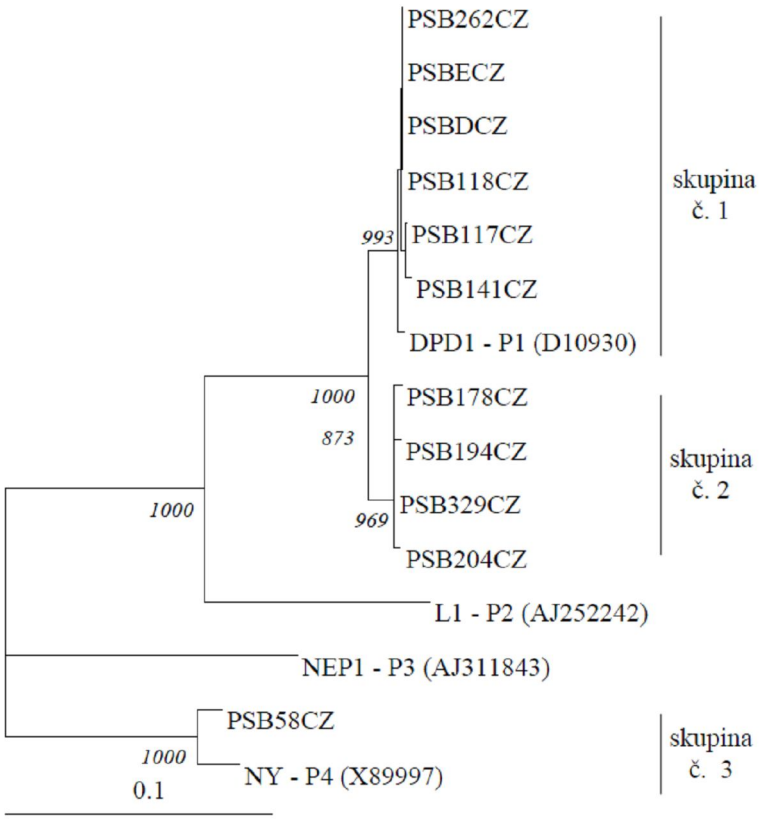


Schéma č.2 Fylogenetický strom príbuznosti vybraných izolátov vírusu nukleotidových sekvencií kódujúcej oblasti VPg (Znázornené hodnoty bootstrap vyššie než ≥ 70%).

Izolát	VPg	Biologický test
PSB58CD	P4	P4
PSB117CZ	P1	P1
PSB118CZ	P1	P1
PSB141CZ	P1	P1
PSB178CZ	P1	P1
PSB194CZ	P1	*?*
PSB204CZ	P1	*?*
PSB262CZ	P1	*?*
PSB329CZ	P1	P1
PSBDCZ	P1	P1
PSBECZ	P1	P1

Tabuľka č.1 Zaradenie českých izolátov PSbMV do patotypov podľa doteraz známych výsledkov porovnania sekvenčnej variability oblasti VPg a biologických testov. (P1 - patotyp P1; P4 - patotyp P4; ?-prebiehajúci experiment)

Použitá literatúra

Alconero, R., Providenti R., Gonsales, D. (1986): Three pea seedborne nosaic virus pathotypes from pea and lentil germ plasm. Plant Disease, 70: 783 - 786.

Gao, Z., Johansen, E., Eyers,S., Thomas,C.L., Noel Elis, T.H., MAule, A.J. (2004): Tghe potyvirus recessive resistance gene, sbm1, identifies a novel role for translation initiation factor eIF4E in cell-to-cell trafficking. Plant Journal, 40: 376-385.

Hjulsager, CH.K., Olsen, B.S., Jensen, D.M.K., Cordea, M.I., Krath, B.N., Johansen, E., Lund, O.S. (2006): Multiple determinatns in the coding region of Pea seed-borne mosaic virus P3 and unvolved in virulence against *sbm*-2 resistance. Virology, 355: 52-61.

Page, R.D.M. (1996): TREEVIEW: An application to display phylogenetic trees on personal computers. Computer Applications in the Biosciences, 12: 357-358.

Rajamäky, M.L., Valkonen, J.P.T. (2001): Viral Genome-Linked Protein (VPg) Controls Accumulation and Phloem –Loading of Potyvirus in Inoculated Potatoe Leaves. Molecular Plant-Microbe Interactions, 15(2): 138-149.

Torok, V.A. (2006): Discriminating Between Isolates of PSbMV Using Nucleotide Sequence Polymorphisms in the HC-Pro Coding Region. Plant Disease, 91(5): 490-496.

Petrusová J and Navrátil M 2008 *Molecular variability of PSbMV genome regions responsible for interactions with the genes of resistance in the pea plant.* Book of Abstracts, The 3rd Conference of the International Working Group on Legume and Vegetables Viruses (IWGLVV), Ljubljana, Sovenia. pp 39. National Institute of Biology, Ljubljana, Slovenia.

The 3rd Conference of the International Working Group on Legume and Vegetable Viruses (IWGLVV)

August 20th – 23rd, 2008
Ljubljana, Slovenia

BOOK OF ABSTRACTS

National Institute of Biology
Department of Biotechnology and Systems Biology

Working group on Legume and Vegetable Viruses

Molecular variability of PSbMV genome regions responsible for interactions with the genes of resistance in the pea plant

J. Petrusová, M. Navrátil

Department of Cell Biology and Genetics, Faculty of Science, Palacky University in Olomouc, Czech Republic

Abstract

Pea seed-borne mosaic virus (PSbMV) is an economically important seed transmitted virus (family *Potyviridae*: genus *Potyvirus*) infecting pea (*Pisum sativum* L.). Viral genome is positive ssRNA 9,8 kb long, consist of 9 regions coding 8 non structural proteins and 1 capside protein. The main symptoms of infection by PSbMV are mosaic, leaf swirling, bine malformation and blooming delay. There are also serious lost of production which varies between 11-82% according to cultivar and virus isolate (Torok, 2006).

PSbMV isolates were divided into 4 pathotypes P-1, P-2, P-3 and P-4 according to the symptoms and progression of infection on the pea lines caring genes of resistance sbm-1, sbm-2, sbm-3, and sbm-4 (Alconero et al., 1986; Providenti and Alconero, 1988). Genes sbm-1 and sbm-2 are effective on the level of specific prohibition of the local and systemic infection by PSbMV (Leonard et al., 2004). In this system of plant resistance they interact with VPg and P3+6K1 in PSbMV genome.

For analyze 11 isolates were selected according to the previous biological tests (Safarova, 2008). In each isolate coding regions VPg and P3+6K1 were sequenced, and nucleotide sequences were compared in phylogenetic analysis. As representatives of each pathotype were used isolates: P-1 - DPD1 (D10930), P-2 – L1 (AJ252242), P-3 – NEP1 (AJ311843) and P-4 – NY (X89997).

In phylogenetic analysis 11 isolates were divided into 2 main clusters - pathotype P1 and P4. Interesting results were noticed in pathotype P1, where 4 isolates created special group. These were selected for next biological tests, which are in progress.

Acknowledgements

This work was partially supported by the Ministry of Education of the Czech Republic grant number MSM 6198959215.

References

Alconero R., Providenti R. and Gonsales D. (1986). *Plant Disease* 70, 783-786.

Loenard S., Viel C., Beauchemin C., Daigneaut N., Fortin M. G., and Laliberte J. F. (2004). *Journal of General Virology* 85, 1055-1063.

Providenti R. and Alconero R. (1988). *Journal of Heredity* 79, 76-77.

Šafářová D., Navrátil M., Petrusová J., Pokorný R. and Piáková Z. (2008). *Acta virologica* 52, 53-57.

Torok V. A. (2006). *Plant Disease* 91 (5), 490-496.

Presenting author information

Jana Petrusová, Dept. of Cell Biology and Genetics, Faculty of Science, Palacky University in Olomouc, 783 71 Olomouc-Holice, Czech Republic

Phone: +420 585634901, Fax: +420 585634905, E-mail: jana_petrus@hotmail.com