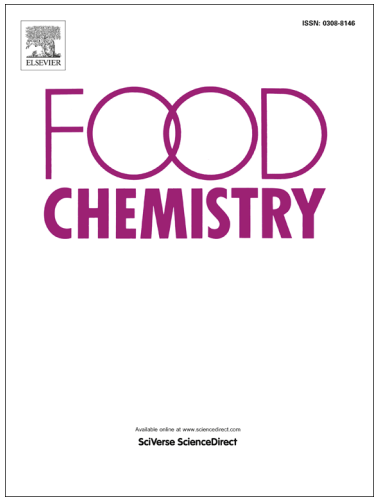
# Purification and characterisation of a protease (tamarillin) from tamarillo fruit




# Purification and characterisation of a protease (tamarillin) from tamarillo fruit

Zhao Li, Ken Scott, Yacine Hemar, Huoming Zhang, Don Otter

PII: S0308-8146(18)30327-3
DOI: https://doi.org/10.1016/j.foodchem.2018.02.091
Reference: FOCH 22475

To appear in: *Food Chemistry*

Received Date: 25 October 2017
Revised Date: 13 February 2018
Accepted Date: 16 February 2018

Please cite this article as: Li, Z., Scott, K., Hemar, Y., Zhang, H., Otter, D., Purification and characterisation of a protease (tamarillin) from tamarillo fruit, *Food Chemistry* (2018), doi: https://doi.org/10.1016/j.foodchem.2018.02.091

# Purification and characterisation of a protease (tamarillin) from tamarillo fruit

Zhao Li [a], Ken Scott [b], Yacine Hemar [a,c], Huoming Zhang[d], Don Otter [e,*]

[a] School of Chemical Sciences, The University of Auckland, Private Bag 92019, Auckland 1142, New Zealand

[b] School of Biological Sciences, The University of Auckland, Private Bag 92019, Auckland 1143, New Zealand

[c] The Riddet Institute, Massey University, Private Bag 11222, Palmerston North 4442, New Zealand

[d] King Abdullah University of Science and Technology (KAUST), Core Labs, Thuwal 23955-6900, Saudi Arabia

[e] Center for Dairy Research, College of Agriculture and Life Sciences, University of Wisconsin-Madison, 1605 Linden Drive, Madison 53706, USA



* Corresponding author: Don Otter. *Email address*: donotter1@gmail.com

## Highlights

• A novel serine protease (tamarillin) purified from tamarillo fruit

• High pH optimum at 11 and wide range of thermostability

• Stable at most of environment condition but $Hg^{2+}$ in present

• A subtilisin-like serine protease was identified using *de novo* sequencing

**Abstract**

A protease from tamarillo fruit (*Cyphomandra betacea* Cav.) was purified by ammonium sulphate precipitation and diethylaminoethyl-Sepharose chromatography. Protease activity was determined on selected peak fractions using a casein substrate. Sodium dodecyl sulphate polyacrylamide gel electrophoresis analysis showed that the peak with the highest protease activity consisted of one protein of molecular mass ca. 70 kDa. The protease showed optimal activity at pH 11 and 60°C. It was sensitive to phenylmethylsulphonyl fluoride while ethylenediaminetetraacetic acid and *p*-chloromercuribenzoic acid had little effect on its activity, indicating that this enzyme was a serine protease. $Hg^{2+}$ strongly inhibited enzyme activity, possibly due to formation of mercaptide bonds with the thiol groups of the protease, suggesting that some cysteine residues may be located close to the active site. *De novo* sequencing strongly indicated that the protease was a subtilisin-like alkaline serine protease. The protease from tamarillo has been named 'tamarillin'.

**Key words:** *Cyphomandra betacea*; Solanaceae; Purification; Characterisation; *De novo* sequencing; Subtilisin-like serine protease; Tamarillin

## 1. Introduction

Proteolytic enzymes account for approximately 60% of the total global commercial enzyme business due to their wide range of substrate specificities, their wide range of pH activity and their relatively high tolerance to thermal treatment (Rajkumar, Jayappriyan, & Rengasamy, 2011). In the food industry, Sumantha, Larroche, and Pandey (2006) stated that proteases contributed to the improvement of functional and nutritional properties of proteins, and to the hydrolysis of gelatin, soy protein, casein and whey proteins. In terms of food processing, especially in the dairy industry, proteases play a significant role in milk coagulation, which is a key stage in cheesemaking.

Calf rennet is a traditional milk-clotting enzyme which has been used predominantly in the manufacture of different cheeses for centuries because of its high specificity to cleave the $Phe_{105}$-$Met_{106}$ peptide bond in κ-casein. The cleavage of κ-casein, which is located on the surface of the casein micelle, results in the initiation of milk coagulation due to the decrease of repulsive forces between the casein micelles (Vishwanatha, Rao, & Singh, 2010). A worldwide shortage of rennet and increased global cheese production has resulted in the

search to find new proteases with rennet-like properties as a substitute in cheese manufacture (Jacob, Jaros, & Rohm, 2011).

Proteases from plant sources have shown a high potential as rennet substitutes and have already been used in cheese production processes (Jacob et al., 2011). Most of the proteases purified from plant sources are aspartic proteases, but milk-clotting activity has also been identified in cysteine and serine proteases. These enzymes are typically active over a wide range of temperature and pH values, and exhibit great potential in the food, biotechnology and medical industries (Shah et al., 2014). Cysteine proteases, such as actinidin purified from kiwifruit (Katsaros, Tavantzis, & Taoukis, 2010), ficin extracted from the latex of ficus (Devaraj, Gowda, & Prakash 2008) and papain from papaya (Saran, Solanki, & Choudhary, 2016), all showed milk-clotting activities and casein hydrolysis ability. However, air oxidation and metal ions are two factors that negatively affect the activity of cysteine proteases, and thus need to be considered during enzymatic reactions. Unfortunately, it is not convenient or economical to add reductants and chelating agents to cysteine proteases (Tomar, Kumar, & Jagannadham, 2008).

Serine proteases, which were once thought to be rare in plants, have been found and extracted from different parts of various plants, including the latex, fruits, seed, root and leaves (Shah et al., 2014). Recently, several serine proteases have been isolated and used for milk coagulation, including dubiumin, which is purified from the seeds of *Solanum dubium* (Ahmed, Morishima, Babiker, & Mori, 2009), streblin, extracted from *Streblus Asper* (Tripathi, Tomar & Jagannadham, 2011), and neriifolin, which is present in the latex of *Euphorbia neriifolia* (Yadav, Patel & Jagannadham, 2011). Compared to cysteine proteases, serine proteases do not require the use of reductants and chelating agents. In addition, serine proteases have the ability to retain their activity and remain stable under relatively high

temperatures, a wide range of pH values and even in the presence of both oxidizing agents and surfactants (Ahmed et al., 2009). Therefore, serine proteases have high potential usage in the food industry and represent economical value as milk-clotting enzymes.

Tamarillo (or tree tomato, *Cyphomandra betacea* (F)) originated in South America. Tamarillo was introduced into New Zealand, and has been planted commercially, since the late 19th century (Bohs, 1989). Tamarillo is closely related to the genus *Solanum* which belongs to the *Cyphomandra* (Solanaceae) species (Bohs, 1989). In the present work, an alkaline serine protease isolated from tamarillo was purified and characterised. To the best of our knowledge, a tamarillo protease and its properties have never been reported before.

## 2. Material and methods

### *2.1. Material*

Tamarillo fruit (Laird's Large cultivar) were obtained from Maungatapere, New Zealand. Rennet was purchased from RENCO New Zealand Laboratory, Eltham, New Zealand. Casein from bovine milk was purchased from Sigma-Aldrich (Auckland, New Zealand). diethylaminoethyl (DEAE) - Sepharose was bought from GE Healthcare, Sweden. All chemicals were of chemical grade and were purchased from Sigma-Aldrich.

### *2.2. Protease purification*

Tamarillo fruit (300 g) were homogenised for 20 min using a Polytron PCU2 laboratory homogeniser (Brinkmann Instruments, Luzern, Switzerland), followed by filtration through two layers of cheesecloth to remove insoluble material. The filtrate was mixed with 80 mL sodium citrate (0.05 M, pH 5.5). The mixture was centrifuged at 15,000 × *g* at 4 °C for 20 min. According to the method of Burgess (2009), 61.69 g of ammonium sulphate was added to 155 mL supernatant to make a 65% saturation ammonium sulphate solution, mixed gently

for 20 min using a magnetic stirrer, followed by standing in an ice bath overnight. The precipitated proteins were collected by centrifugation at 15,000 × *g* at 4 °C for 20 min and re-dissolved in 30 mL 0.05 M citrate buffer (pH 5.5). Ammonium sulphate was removed by dialysing the protein solution in a DEAE cellulose tube (molecular weight cut-off, 12,000 Da) against the same citrate buffer at 4 °C for 24 hrs, with the dialysis buffer being changed three times.

Ion-exchange chromatography on a DEAE-Sepharose resin packed in a 22 × 5 cm column equilibrated with sodium citrate buffer (0.05 M, pH 5.5) overnight was used for protein purification. Tamarillo protein solution (40 mL) was applied to the column. After washing the column with 240 mL of citrate buffer the protein was eluted using a 200 mL linear gradient of 0.0-1.0 M NaCl in citrate buffer at a flow rate of 1.0 mL/min. Eluant fractions (3 mL) were collected in glass tubes and their absorbance measured at 280 nm using an UVmini-1240 spectrophotometer (Shimadzu Corporation, Australia). Consecutive tubes giving similar absorbance readings were pooled together as a single fraction and dialysed against Milli-Q water for 24 hrs. The protein in each fraction was concentrated by freeze-drying and stored at -80 °C for subsequent analysis.

*2.3. Electrophoresis and protein content*

The molecular mass of the purified protease was determined using 4-12% (w/v) Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) (Laemmli, 1970), under denaturing and reducing conditions. The gels were stained with 0.1% (w/v) Coomassie brilliant blue R-250. The molecular mass was calculated by comparison to standard protein markers.

The protein content of the different fractions collected from the ion-exchange chromatography was estimated using the absorbance at 280 nm. Protein was quantitatively measured by the Bradford method (Bradford, 1976) using a calibration curve of bovine serum albumin (BSA).

*2.4. Proteolytic activity assay*

Proteolytic enzyme activity was measured using the method of Kunitz (1947) with some modifications, and with casein as the substrate. Protein preparation (100 μL) was mixed with 1.1 mL of 0.1 M Tris-HCl buffer (pH 7) containing 1% (w/v) casein, vortexed and then incubated at 35 °C in a water bath for 20 min. The reaction was stopped by adding 1.8 mL 10% (w/v) trichloroacetic acid (TCA), and the solution was then allowed to stand for 30 min at room temperature. Precipitated protein was removed by centrifugation at 10,300 × *g* for 20 min, and the absorbance of the supernatant was measured at 280 nm against the blank. The blank was prepared by mixing TCA and enzyme solution, and then adding casein solution. One unit of protease activity was defined as the amount of enzyme that liberated 1 μg of tyrosine per milliliter in 1 min under the assay conditions. A standard curve was made by using 0-100 mg/L tyrosine (Hadj-Ali, Agrebi, Ghorbel-Frikha, SellamiKamoun, Kanoun, & Nasri, 2007), and a standard specific enzyme activity was defined in units/mg of protein.

*2.5. Optimal temperature, optimal pH and stability*

The effect of temperature on the enzyme activity was studied using casein as the substrate. The standard enzyme assay (Section 2.4 above) was performed over a temperature range from 0 to 100 °C for 20 min. The temperature stability of the enzyme was investigated by pre-incubating the enzyme at specific temperatures ranging from 0 to 100 °C for 20 min, followed by measurements using the enzyme assay under optimum pH and temperature conditions. The remaining protease activity was expressed as a percentage of the control.

The optimal pH value for enzyme activity was also studied using 50 mM buffers of different pH values ranging from 6 to 14 following the method of Attri, Khaket, Jodha, Singh, and Dhanda (2015). The buffers used in the standard proteolytic activity assays were: phosphate buffer, pH 6; Tris-HCl buffer, pH 7-8; glycine-NaOH buffer, pH 9-10; phosphate-NaOH buffer, pH 11; KCl-NaOH buffer, pH 12-13 and 1 M NaOH, pH 14. The pH stability of the protease was determined by pre-incubating 100 µL of the purified enzyme with 150 µL of the different pH buffers at 25 °C for 20 min, and the remaining enzyme activity was then measured using the optimum assay conditions. The results of the residual enzyme activities were expressed as a percentage of the control activity.

*2.6. Effect of organic solvents, inhibitors and metal ions on protease activity*

Purified protease (100 µL) was mixed with 25% (v/v) of different organic solvents including isopropanol, methanol, ethanol, glycerol, dimethyl sulfoxide (DMSO), and chloroform, and the mixture was then incubated at 25 °C for 20 min. The enzyme was assayed (Section 2.4) and the activity was compared to the control without organic solvent.

Inhibitors (1 mM and 5 mM) such as phenylmethylsulphonyl fluoride (PMSF), for serine proteases, *p*-chloromercuribenzoic acid (PCMB), for cysteine proteases, and ethylenediaminetetraacetic acid (EDTA), for metalloproteases, were added to the enzyme activity assay. The enzyme was pre-mixed with the inhibitors and incubated at room temperature for 20 min. The residual activity was measured by the standard enzyme assay. The results were expressed as the percentage of the protease activity measured without inhibitors.

The effect of different chloride salts on enzyme activity, which included $Na^{+}$, $Hg^{2+}$, $Zn^{2+}$, $Co^{2+}$, $Ca^{2+}$, $Mn^{2+}$, $Mg^{2+}$ (at concentrations of 1 mM and 5 mM), was studied. The enzyme

was pre-incubated with the individual metal ions for 20 min at room temperature, followed by the standard enzyme assay. The results were expressed as the percentage of the protease activity measured without metal ions.

*2.7. Kinetic parameters determination*

The relationship between increasing substrate concentration and hydrolysis reaction velocity of tamarillin was studied in 50 mM KCl-NaOH buffer, pH 11 at 60 °C. Casein was used as the substrate over the range of 0-800 μM. The measurements were performed using the standard protease assay. Blanks were made by mixing the casein substrates with the enzyme inactivated by TCA. The Michaelis–Menten plot (Lineweaver et at., 1934) and the Hanes plot (Hanes, 1932) were used to calculate the kinetic parameter, $K_m$.

*2.8. Tamarillin characterisation using de novo peptide sequencing*

*De novo* peptide sequencing of the tamarillin was performed based on the method of Shevchenko, Wilm, Vorm, and Mann (1996). Tamarillin, prepared as in Section 2.2, was subjected to electrophorisis using 4-12% (w/v) gradient precast NuPAGE Bis-Tris gels (Invitrogen, Carlsbad, CA, USA) in an XCell SureLock Mini-Cell (Invitrogen) for 1 hr with a constant voltage of 180V. The gels were stained with Coomassie brilliant blue R-250 and then destained twice with 50 mM $NH_4HCO_3$ in 50% methanol and washed three times with Milli-Q water. The protein band corresponding to tamarillin was excised, and then dried and rehydrated twice with 100% acetonitrile and 100 mM $NH_4HCO_3$, respectively. Trypsin digestion of the tamarillin band was performed overnight at 37 °C in 10 μL of 12.5 ng/μL sequencing grade trypsin (Promega, Madison, WI, USA) in 50 mM $NH_4HCO_3$. Peptides were extracted twice using 5% formic acid in 50% acetonitrile. Pooled extracts were dried in a

speed-vacuum (Thermo Electron, Waltham, MA, USA). The extracts were then redissolved in 50 µL 0.1% v/v formic acid, and were desalted with a C18 ZipTip (Millipore, MA, USA).

The extracted peptides were analysed using a NanoLC-MS/MS on-line system which consisted of a nano-pump UltiMate™ 3000 UHPLC binary HPLC system (Dionex, Thermo Fisher) coupled to a Q Exactive HF mass spectrometer (Thermo Fisher, Germany). Peptides were resuspended in 20 µL of sample buffer (3% acetonitrile, 0.1% formic acid) and a 2 µL sample was injected onto a pre-column (Acclaim PepMap, 300 µm × 5 mm, 5 µm particle size). After loading, peptides were eluted onto an Acclaim PepMap100 C18 capillary column (75 µm × 15 cm, 100 Å, 3 µm particle size). Peptides were separated and eluted into the MS, at a flow rate of 300 nL/min, using a 40 min gradient from 5% to 40% mobile phase B. Mobile phase A was 0.1% formic acid in Milli-Q water and mobile phase B was 80% acetonitrile and 0.1% formic acid. The mass spectrometer was operated in positive and data-dependent mode, with a single MS scan (350-1400 *m/z* at 60000 resolution; at 200 *m/z* in a profile mode) followed by MS/MS scans on the 10 most intense ions at 15000 resolution. Ions selected for the MS/MS scan were fragmented using higher energy collision dissociation (HCD) at a normalized collision energy of 28% and using an isolation window of 1.8 *m/z*.

The MS RAW files from the Q Exactive HF were analysed via automated *de novo* sequencing using Peak Studio (v 5.0) software. The peptide sequences of high-quality (score >0.65) were searched using a MS-BLAST strategy. Colour-based criteria (Waridel et al., 2007) were followed after merging the candidate sequences from *in silico de novo* sequencing from the Peak Studio software. The candidate sequences derived from the significant results were merged into a single query string, and submitted to MS-BLAST at http://genetics.bwh.harvard.edu/msblast/ for a search against a nonredundant database (nrdb95). Only hits with a total BLAST score above 100 or with at least one high-scoring

segment pair above 72 were considered as positive (Katz, Waridel, Shevchenko, & Pick, 2007). Average Local Confidence (ALC) scores, which reflect the confidence in peptide sequence, are calculated by adding the positional confidence for each amino acid in the peptide sequence and dividing by the total number of amino acids. Only peptides with an ALC score higher than 50% were considered (Tran et al., 2016, Unsworth et al., 2017).

## 3. Results and discussion

### *3.1. Purification and molecular mass of tamarillo protease*

As shown in Fig. 1, two absorbance peaks eluted from the DEAE-Sepharose column with absorbance maxima for the 150 and 212 fractions. Maxima for the protease activity did not match the protein profile however, with maximum activity in fraction 126, a shoulder at fraction 150 and then a second, smaller maximum at fraction 225. Fractions for two of the different peaks were pooled (fractions 113-137, Peak 1; and fractions 204-221, Peak 2) and the first peak showed a higher protease activity (0.043 ± 0.003 units/mL) compared to the second peak (0.006 ± 0.002 units/mL). Compared with the crude extract, the first peak showed an approximate 5.6-fold increase in protease activity as shown in Table 1, while the second peak contained more protein but only a low level of activity was measured.

The protease in Peak I migrated as a single protein band on a 4-12% polyacrylamide gel under both denaturing and reducing conditions. The molecular mass of the purified enzyme was approximately70 kDa under denaturing conditions (Fig. 1, Insert: Lane, Non-red), while under reducing conditions the main band migrated at approximately 23 kDa (Fig. 1, Insert: Lane, Red). The purified tamarillo protease was named ‘tamarillin’.

The molecular mass of tamarillin was similar to that of the subtilisin-like serine endoproteases from barley (Terp, Thomsen, Svendsen, Davy, & Simpson, 1999) and to a

typical pyrolysin-like subtilase from tomato (*Solanum lycopersicum*) plants (Ottmann et al., 2009). Antão and Malcata (2005) indicated that the molecular mass of serine proteases in plants varied from 19 to 110 kDa with the majority having molecular masses between 60 and 80 kDa.

*3.2. Enzyme pH and temperature optima and stability*

When the assay for tamarillin was conducted using different pH buffers, the enzyme displayed protease activity over a wide range of pH values, from pH 6 to 12, with a maximum hydrolysis activity at pH 11 (Fig. 2A). Although at pH 10 only 60% of the maximum activity was observed, tamarillin displayed broad activity maxima with approximately 90% of this activity being retained at pH 12.

The stability of tamarillin after incubation at different pH values, followed by activity analysis at the optimal pH of 11, showed that the enzyme was pH stable between pH values 7-11, and its stability then started to decrease markedly at pH 12 (Fig. 2A). These two sets of results, regarding enzyme activity and stability, indicated that tamarillin belonged to the plant alkaline protease family.

Other plant proteases with high pH optima include those from the dandelion, *Taraxacum officinale* webb (Rudenskaya et al., 1998) and *Cucurbita ficifolia* (Curotto et al., 1989), which have pH stability ranges of 7 - 9 and 8 - 11, respectively. These are all subtilisin-like serine proteases. A survey of the literature suggested that serine proteases extracted from plant sources that exhibit optimum enzyme activity and stability at high pH values are very rare. The serine proteases, dubiumin, isolated from the seed of fresen (Ahmed et al., 2009) and a protease from *Cucumis trigonus* Roxburghi (Asif-Ullah, Kim, & Yu, 2006) also showed optimal activity at pH 11 (Table S-1). Moreover, a pyrolysin-like subtilase from

tomato plants also exhibited a high level of stability at alkaline pH values (Ottmann et al., 2009). In terms of food processing, pH is an important parameter that affects the usage of most enzymes, especially under alkaline conditions. This is because most enzymes are unstable in alkaline pH, and thus are unable to catalyse hydrolysis reactions under high pH conditions, which might limit their usefulness in food applications, even when used as cheese-making coagulants (Lamas, Barros, Balcao, & Malcata, 2001).

According to the results shown in Fig. 2B, 60 °C was the optimal temperature of tamarillin. The enzyme activity increased exponentially when the temperature was increased from 0 to 60 °C, then decreased nearly linearly to zero activity when the temeprature was further increased to 100 °C, although it still retained nearly 40% of its activity at 80 °C.

The activity of tamarillin also remained stable when it was incubated at different temperatures for 20 min, up to a temperature of 60 °C. Above this temperature the activity decreased nearly linearly as the temperature was increased to 100 °C. This suggested that the optimal tamarillin activity was limited by the stability of the protein at temperatures above 60 °C.

The plant serine protease from *Euphorbia milii* (Yadav et al., 2006) showed a similar optimal temperature. The optimum temperature of some other plant serine proteases were also high, for example, both dubiumin from the seed of *Solanum dubium* Fresen (Ahmed et al., 2009) and the protease from *Cucumis trigonus* Roxburghi (Asif-Ullah et al., 2006) displayed optimum activity at a temperature of 70 °C (Table S-1).

*3.3. Effect of organic solvents, inhibitors and metal ions on tamarillin activity*

The effect of different organic solvents (25% (v/v)) on the activity of tamarillin are reported in Fig. 3A. The enzyme retained ~96% of its original activity in the presence of ethanol, and

up to 89% and 85% of its activity in methanol and glycol, respectively. In the presence of isopropanol, DMSO, and chloroform the enzyme activity retained was ~76%, ~77%, and ~62%, respectively. The activity of enzymes in the presence of organic solvents can be affected by many factors. The hydrogen bonds, the hydrophobic interactions and the charges on the proteins can be disrupted, which in turn can affect the dynamics and conformation of the proteases, leading to changes in their catalytic processes. Under organic solvents, depending on the concentration, enzymes tend to be inactivated and unstable (Ghorbel, Sellami-Kamoun, & Nasri, 2003). While tamarillin remained active in the presence of some of the organic solvents investigated here, unfortunately these results cannot be compared to other plant serine proteases due to the lack of other published data.

Inhibition studies can provide insight into the nature of an enzyme, its cofactor requirement, and the nature of the active center (Sigma & Mooser, 1975). As can be seen in Fig. 3B, EDTA and PCMB (1 mM and 5 mM) had little effect on the activity of tamarillin, indicating that this protease was neither a cysteine protease nor a metalloprotease. However, the typical serine protease inhibitor PMSF inhibited the enzyme activity by 70% at 1 mM PMSF and 80% at 5 mM PMSF. Serine residues in the active site of a serine protease covalently bind PMSF. PMSF can therefore be used as a label to determine the active site in serine proteases (James, 1978). Thus, it is likely that the purified tamarillo protease is a serine protease. Other plant serine proteases, including a subtilisin-like protease from *Cucumis trigonus* Roxburghi (Asif-Ullah et al., 2006), dubiumin from seed of *Solanum dubium* Fresen (Ahmed et al., 2009) and milin from the medicinal plant *Euphorbia milii* (Yadav, Pande, & Jagannadham, 2006), were reported to be strongly inhibited by PMSF.

Metal ions, namely $Na^{+}$, $Zn^{2+}$, $Co^{2+}$, $Ca^{2+}$, and $Mn^{2+}$ (at 1mM or 5mM), showed negligible effect on tamarillin activity (Fig. 3C). These results illustrated that there was no requirement

for the tamarillin to be activated by these metal ions. However, the protease activity was strongly inhibited by the $Hg^{2+}$ metal ion, with the enzyme retaining only ~32% and ~30% of its original activity after 20 min incubation with 1 mM and 5 mM $Hg^{2+}$, respectively. According to Ahmed et al. (2009), dubiumin, a serine protease from the seeds of *Solanum dubium*, was inhibited by $Hg^{2+}$, indicating that dubiumin contained cysteine residues near its active site. Similar findings were reported for a subtilisin-like protease from *Bacillus cereus* (Moriyama, Sugimoto, Zhang, Inoue, & Makino, 1998).

*3.4. Kinetic parameters determination*

The effect of increasing the casein concentration on the hydrolysis reaction velocity can be described using the Michaelis-Menten plot (Fig. 4A) and the Hanes plot (Fig. 4B). The activity of the tamarillin reached saturation at a high substrate concentration (800 μM casein). The value of $K_m$ was 32.3 μM for casein as the substrate. The $K_m$ value gave an indication of the binding affinity of the substrate to the enzyme and was very comparable with the $K_m$ values of some other plant serine proteases. $K_m$ values of ~330, ~33.3 and ~66 μM were reported for the serine proteases benghalensin (Sharma, Kumari, & Jagannadham, 2009), milin (Yadav et al., 2006), and religiosin (Kumari, Sharma, & Jagannadham, 2010) extracted from the latex of the medicinal plant *Ficus benghalensis*, the latex of *Euphorbia milii*, and the latex of *Ficus religiosa* respectively, using casein as the substrate. Note that the $K_m$ values of these plant serine proteases were obtained under different temperature and pH conditions.

*3.5. Protease characterization using de novo protein sequencing*

The functional specificity of tamarillin was further characterised by identifying amino acid sequence similarities between tamarillin and known proteases from other species. Tryptic digested peptides fragments (quality score>0.65) from tamarillin were subject to *de novo*

sequencing using Peak Studio software followed by MS-Blast analysis against the nrdb95 database. The *de novo* sequences of seven peptide fragments are listed in Table 2. The sequences were found to match an identified subtilisin-like protease from two *Solanum tuberosum* database sequences (gi|565386017 and gi|75180715) with BLAST scores greater than 87 (Fig. S-1A and B). These results strongly indicated that tamarillin was a subtilisin-like serine protease.

## 4. Conclusion

Overall, to the best of our knowledge, this is the first study on the purification and characterisation of a protease from tamarillo fruit. This protease exhibited activity over a broad range of pH values, had good temperature stability, and displayed high activity optima for pH (pH 11) and temperature (60 °C). In addition, there was no requirement for tamarillin to be activated by metal ions, $Hg^{2+}$ ions inhibited the protease activity and the typical serine protease inhibitor PMSF strongly inhibited the enzyme activity. Amino acid sequences from tamarillin matched other identified plant subtilisin-like proteases. This subtilisin-like alkaline serine protease might be used in food applications, such as cheese-making, or other applications which require alkaline pH environments.

## Acknowledgements

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

**References:**

Ahmed, I. A. M., Morishima, I., Babiker, E. E., & Mori, N. (2009). Dubiumin, a chymotrypsin-like serine protease from the seeds of *Solanum dubium* Fresen. *Phytochemistry, 70*, 483–491.

Antao, C. M., & Malcata, F. X. (2005). Plant serine proteases: biochemical, physiological and molecular features. *Plant Physiology and Biochemistry*, *43*, 637-650.

Asif-Ullah, M., Kim, K. S., & Yu, Y. G. (2006). Purification and characterization of a serine protease from *Cucumis trigonus* Roxburghi. *Phytochemistry*, *67*, 870-875.

Attri, P., Khaket, T. P., Jodha, D., Singh, J., & Dhanda, S. (2015). Biochemical, kinetic, and *in silico* characterization of DING protein purified from probiotic lactic acid bacteria *Pediococcus acidilactici* NCDC 252. *Applied Biochemistry and Biotechnology*, *175*, 1092-1110.

Bohs, L. (1989). Ethnobotany of the genus *Cyphomandra* (Solanaceae). *Economic Botany*, *43*, 143-163.

Bradford, M. B. (1976). A rapid and sensitive method for the quantitation of micrograms quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry*, *72*, 248-254.

Burgess, R. R. (2009). Protein precipitation techniques. *Methods in Enzymology*, *463*, 331-342.

Devaraj, K. B., Gowda, L. R., & Prakash, V. (2008). An unusual thermostable aspartic proteinase from the latex of *Ficus racemosa* (L.). *Phytochemistry, 69*, 647–655.

Ghorbel, B., Sellami-Kamoun, A., & Nasri, M. (2003). Stability studies of protease from *Bacillus cereus* BG1. *Enzyme Microbial Technology*, *32*, 513-518.

Hadj-Ali, N. E., Agrebi, R., Ghorbel-Frikha, B., SellamiKamoun, A., Kanoun, S., & Nasri, M. (2007). Biochemical and molecular characterization of a detergent stable alkaline serine-protease from a newly isolated *Bacillus licheniformis* NH1, *Enzyme and Microbial Technology*, *40*, 515–523.

Hanes, C. S. (1932). Studies on plant amylases. The effect of starch concentration upon the velocity of hydrolysis by the amylase of germinated barley. *Biochemical Journal*, *26*, 1406-1421.

Jacob, M., Jaros, D., & Rohm, H. (2011). Recent advances in milk clotting enzymes. *International Journal of Dairy Technology, 64*, 14–33.

James, G. T. (1978). Inactivation of the protease inhibitor phenylmethylsulfonyl fluoride in buffers. *Analytical Biochemistry*, *86*, 574–579.

Katsaros, G. I., Tavantzis, G., & Taoukis, P. S. (2010). Production of novel dairy products using actindin and high pressure as enzyme activity regulator. *Innovative Food Science and Emerging Technologies*, *11*, 47-51.

Katz, A., Waridel, P., Shevchenko, A., & Pick, U. (2007). Salt-induced changes in the plasma membrane proteome of the halotolerant alga *Dunaliella salina* as revealed by blue native gel electrophoresis and nano-LC-MS/MS analysis. *Molecular and Cellular Proteomics*, *6*, 1459-1472.

Kumari, M., Sharma, A., & Jagannadham, M. V. (2010). Decolorization of crude latex by activated charcoal, purification and physico-chemical characterization of religiosin, a milk-clotting serine protease from the latex of *Ficus religiosa*. *Journal of Agricultural and Food Chemistry*, *58*, 8027-8034.

Kunitz, M. (1947). Crystalline soybean trypsin inhibitor II. General properties. *The Journal of General Physiology*, *30*, 291-310.

Laemmli, U. K. (1970). Cleavage of structural proteins during assembly of the head of bacteriophage T4. *Nature*, *277*, 680–685.

Lamas, E. M., Barros, R. M., Balcao, V. M., & Malcata, F. X. (2001). Hydrolysis of whey proteins by proteases extracted from *Cynara cardunculus* and immobilized onto highly activated supports. *Enzyme and Microbial Technology*, *28*, 642–652.

Lineweaver, H., Burk, D., & Deming, W. E. (1934). The dissociation constant of nitrogen-nitrogenase in Azotobacter. *Journal of the American Chemical Society*, *56*, 225-230.

Moriyama, R., Sugimoto, K., Zhang, H., Inoue, T., & Makino, S. (1998). A cysteine-dependent serine protease associated with the dormant spores of *Bacillus cereus*: purification of the protein and cloning of the corresponding gene. *Bioscience, Biotechnology, and Biochemistry*, *62*, 268–274.

Ottmann, C., Rose, R., Huttenlocher, F., Cedzich, A., Hauske, P., Kaiser, M., Huber, R., & Schaller, A. (2009). Structural basis for $Ca^{2+}$-independence and activation by homodimerization of tomato subtilase 3. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 17223-17228.

Rajkumar, R., Jayappriyan, K. R., & Rengasamy, R. (2011). Purification and characterization of a protease produced by *Bacillus megaterium* RRM2: Application in detergent and dehairing industries. *Journal of Basic Microbiology*, *51*, 614–624.

Rudenskaya, G. N., Bogacheva, A. M., Preusser, A., Kuznetsova, A. V., Dunaevsky, Y. E., Golovkin, B. N., & Stepanov, V. M. (1998). Taraxalisin–a serine proteinase from dandelion *Taraxacum officinale* Webb sl. *FEBS Letts*, *437*, 237-240.

Rudenskaya, G. N., Bogdanova, E. A., Revina, L. P., Golovkin, B. N., & Stepanov, V. M. (1995). Macluralisin – a serine proteinase from fruit of *Maclura pomifera* Schneid. *Planta*, *196*, 174–179.

Saran, P. L., Solanki, I. S., & Choudhary, R. (2016). Papaya: biology, cultivation, production and uses. New York: CRC Press.

Shah, M. A., Mir, S. A., & Paray, M. A. (2014). Plant proteases as milk-clotting enzymes in cheesemaking: A review. *Dairy Science and Technology*, *94*, 5-16.

Sharma, A., Kumari, M., & Jagannadham, M. V. (2009). Benghalensin, a highly stable serine protease from the latex of medicinal plant *Ficus benghalensis*. *Journal of Agricultural and Food Chemistry*, *57*, 11120-11126.

Shevchenko, A., Wilm, M., Vorm, O., & Mann, M. (1996). Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. *Analytical Chemistry*, *68*, 850-858.

Sigma, D. S., & Mooser, G. (1975). Chemical studies of enzyme active sites. *Annual Review of Biochemistry*, *44*, 889-931.

Sumantha, A., Larroche, C., & Pandey, A. (2006). Microbiology and industrial biotechnology of food grade proteases: a perspective. *Food Technology and Biotechnology*, *44*, 211–220.

Terp, N., Thomsen, K. K., Svendsen, I., Davy, A., & Simpson, D. J. (1999). Purification and characterization of hordolisin, a subtilisin-like serine endoprotease from barley. *Journal of Plant Physiology*, *155*, 468-476.

Tomar, R., Kumar, R., & Jagannadham, M. V. (2008). A staple serine proteinase, wrightin, from the latex of the plant *Wrightia tinctoria* (Roxb) R. Br.: purification and biochemical properties. *Journal of Agricultural and Food Chemistry*, *56*, 1479–1487.

Tran, N. H., Rahman, M. Z., He, L., Xin, L., Shan, B., & Li, M. (2016). Complete de novo assembly of monoclonal antibody sequences. *Scientific Reports*, *6*, 31730.

Tripathi, P., Tomar R., & Jagannadham M. V. (2011). Purification and biochemical characterisation of a novel protease streblin. *Food Chemistry*, *125*, 1005–1012.

Unsworth, J., Loxley, G. M., Davidson, A., Hurst, J. L., Gómez-Baena, G., Mundy, N. I., Beynon, R. J., Zimmermann, E., & Radespiel, U. (2017). Characterisation of urinary WFDC12 in small nocturnal basal primates, mouse lemur (Microcebus spp.). *Scientific Reports*, *7*, 42940.

Vishwanatha, K. S., Rao, A. A., & Singh, S. A. (2010). Production and characterization of a milk-clotting enzyme from *Aspergillus oryzae* MTCC 5341. *Applied Microbiology and Biotechnology*, *85*, 1849-1859.

Waridel, P., Frank, A., Thomas, H., Surendranath, V., Sunyaev, S., Pevzner, P., & Shevchenko, A. (2007). Sequence similarity-driven proteomics in organisms with unknown genomes by LC-MS/MS and automated *de novo* sequencing. *Proteomics*, *7*, 2318-2329.

Yadav, R. P., Patel A. K., & Jagannadham, M. V. (2011). Purification and biochemical characterization of a chymotrypsin-like serine protease from *Euphorbia neriifolia* Linn. *Process Biochemistry, 46,* 1654–1662.

Yadav, S. C., Pande, M., & Jagannadham, M. V. (2006). Highly stable glycosylated serine protease from the medicinal plant *Euphorbia milli*. *Phytochemistry, 67*, 1414-1426.

## Tables

**Table 1** Purification of tamarillo protease (tamarillin).

| Sample | Vol (mL) | Protein | | Activity | | Specific Activity (units/mg protein) | Purification Factor | Recovery (%) |
|---|---|---|---|---|---|---|---|---|
| | | Concentration (mg/mL) | Total (mg) | Concentration (units/mL) | Total (units) | | | |
| Initial sample[a] | 155 | 0.94 ± 0.01 | 145.0 | 0.112 ± 0.005 | 17.3 | 0.119 ± 0.002 | - | 100 |
| Crude extract[b] | 120 | 1.17 ± 0.03 | 140.0 | 0.143 ± 0.009 | 17.1 | 0.121 ± 0.005 | 1.0 | 99 |
| Peak I[c] | 213[d] | 0.06 ± 0.00 | 13.3 | 0.043 ± 0.003 | 9.1 | 0.682 ± 0.072 | 5.6 | 53 |
| Peak II[e] | 147[d] | 0.10 ± 0.02 | 14.7 | 0.006 ± 0.002 | 0.8 | 0.050 ± 0.013 | 0.4 | 5 |

[a] Tamarillo pulp after homogenisation, filtration and centrifugation

[b] After ammonium sulphate precipitation and dialysis

[c] DEAE-Sepharose fractions 113-137

[d] Only 40 mL was loaded onto the DEAE-Sepharose column for each run; this is the accumulation of three runs

[e] DEAE-Sepharose fractions 204-221

**Table 2** *De novo* sequences from tamarillin

| Chain number | Observed *m/z* | Calculated MW[1] (Da) | *De novo* sequence | ALC (%) |
|---|---|---|---|---|
| 1 | 480.897 | 1439.668 | EGAFEGESAYRPK | 54.12 |
| 2 | 509.800 | 1017.586 | AGVVVEGFLK | 52.35 |
| 3 | 543.795 | 1085.576 | LTYQVTFSK | 72.37 |
| 4 | 561.758 | 1121.502 | GVCESDLTNK | 57.05 |
| 5 | 566.809 | 1131.604 | NAPIVAAFSSR | 73.57 |
| 6 | 669.346 | 1336.678 | STHPDWSPAVIK | 66.47 |
| 7 | 671.828 | 1341.641 | TVTNVGDATSSYK | 68.48 |

[1] Calculated molecular weight (MW) of individual sequences based on $z = 3$ for chain number 1, $z = 2$ for chain number 2, 3, 4, 5, 6 and 7. Also, for chain number 4, there is carbamidomethylation of the Cys (+57.021 Da).

**Captions of figures**

**Fig. 1.** Elution curve of tamarillo protease on anion-exchange chromatography resin. DEAE-Sepharose was pre-equilibrated with 0.05 M sodium citrate buffer at pH 5.5. The unbound proteins were washed out with sodium citrate buffer and column was eluted with a linear gradient of 0.0-1.0 M NaCl at the same pH. Fractions of 3 mL volume were collected. Protein content (■) and proteolytic activity (■) were measured. Insert: analysis of Peak I purified tamarillo protease by SDS-PAGE. Lanes: Marker, molecular weight standards (kDa); Non-red, sample prepared under non-reducing conditions; Red, sample prepared under reducing conditions.

**Fig. 2. (**A) Effect of pH on tamarillin activity (■) and stability (●) ranging from pH 6-14. (B) Effect of temperature on tamarillin activity (■) and stability (●) ranging from 0 to 100 °C. Assay conditions are as described in the Methods section.

**Fig. 3. (**A) Effect of different organic solvents on tamarillin activity. (B) Effect of different inhibitors (■ 1 mM and ■ 5 mM) on the proteolytic activity of tamarillin. Inhibitors were pre-incubated with tamarillin for 20 min at room temperature. (C) Effect of different metal ions (■ 1 mM and ■ 5 mM) on proteolytic activity of the tamarillin. The tamarillin was pre-incubated with various metal ions for 20 min.

**Fig. 4.** (A) Michaelis-Menten plot and (B) Hanes plot for the activity of tamarillin as a function of substrate concentration.

## Highlights

- A novel serine protease (tamarillin) purified from tamarillo fruit
- High pH optimum at 11 and wide range of thermostability
- Stable at most of environment condition but $Hg^{2+}$ in present
- A subtilisin-like serine protease was identified using *de novo* sequencing

**Figures**

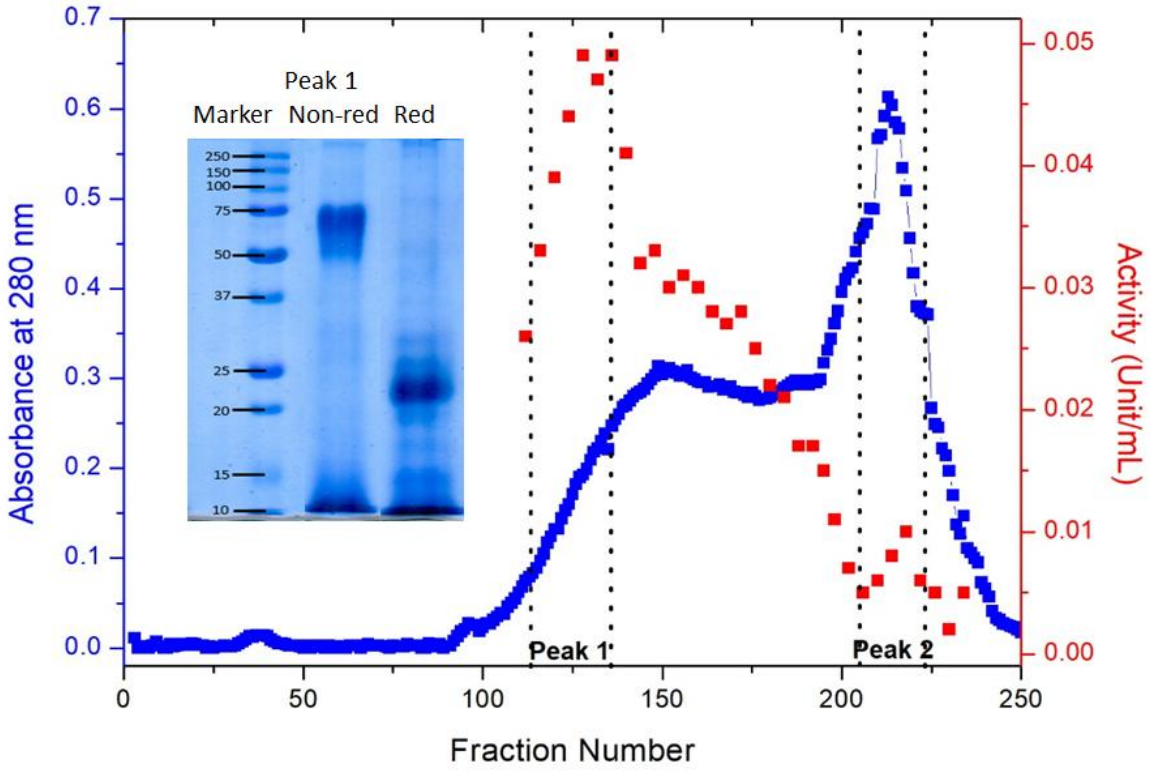



**Fig. 1.**

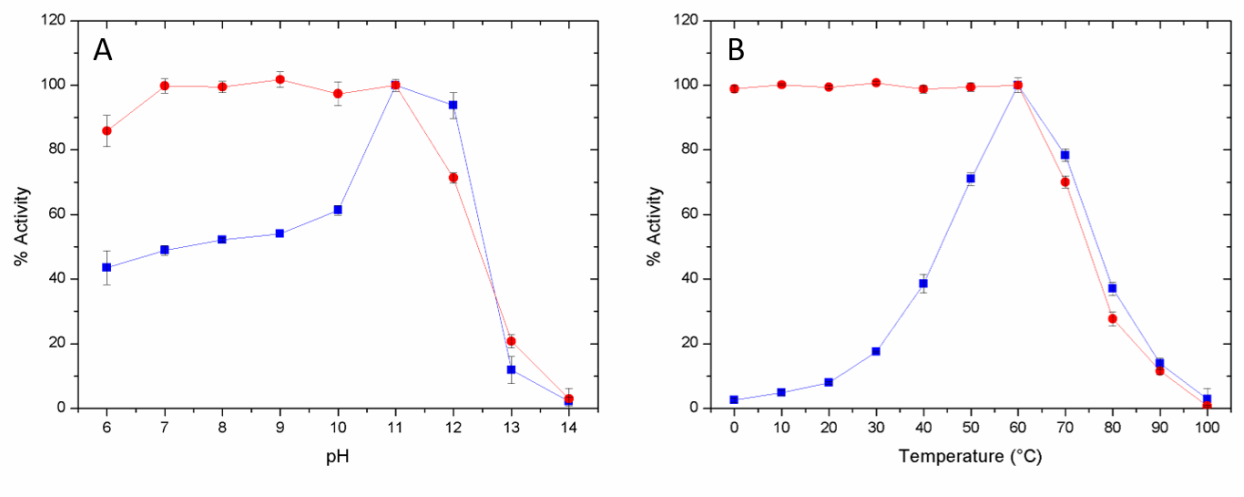



**Fig. 2.**

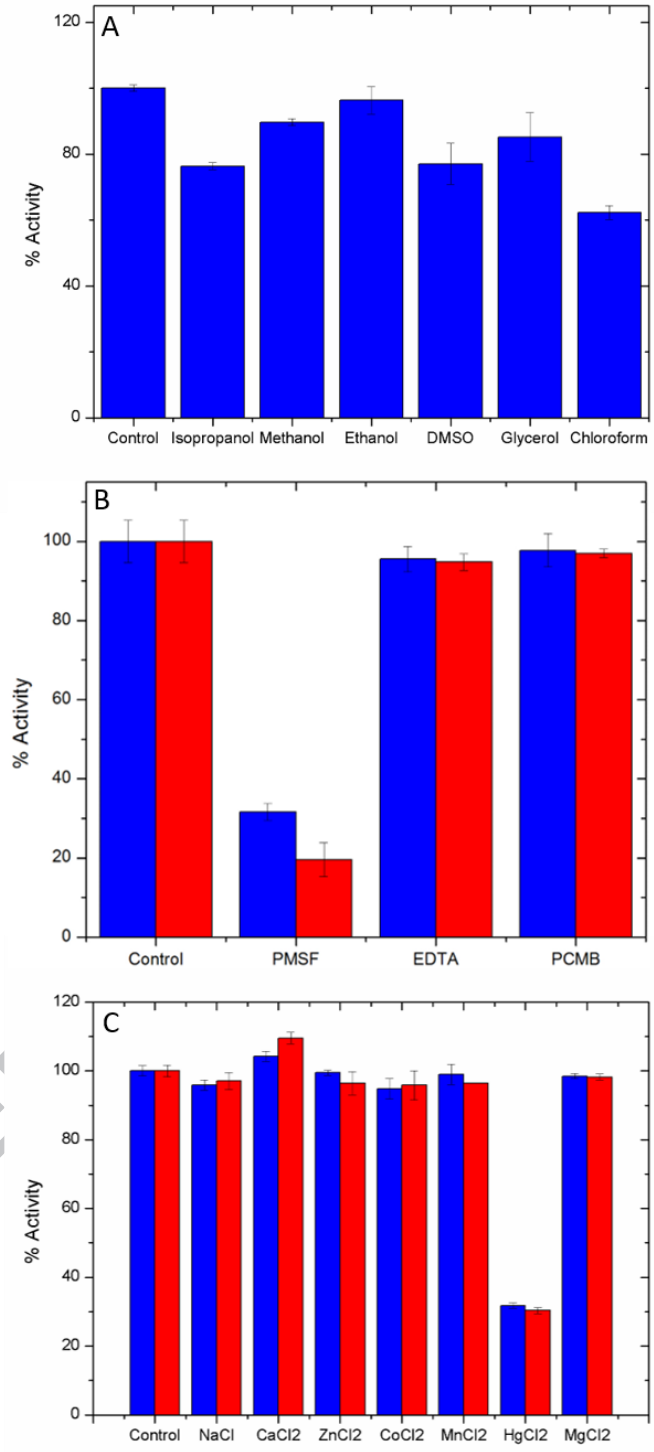



**Fig. 3.**

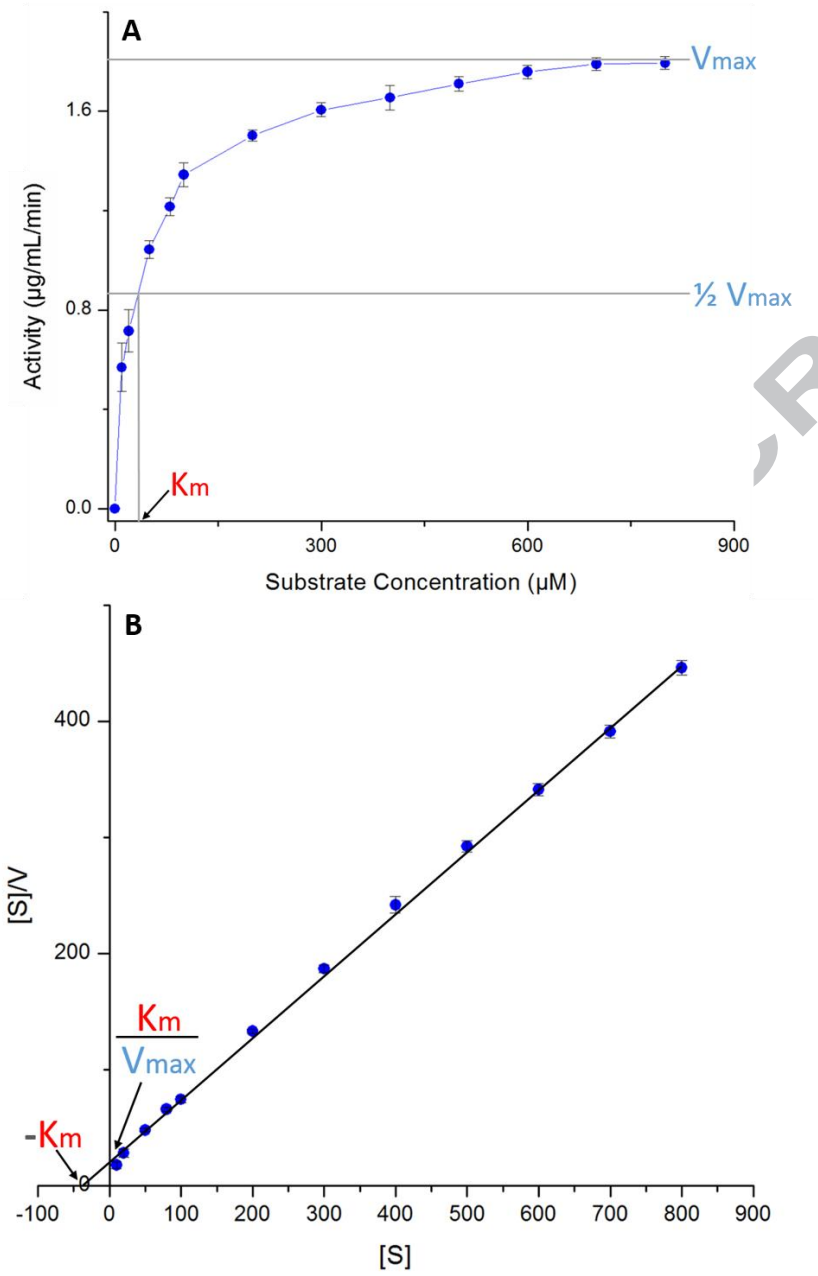



**Fig. 4.**