Planta (2017) 245:343–353
DOI 10.1007/s00425-016-2611-6

ORIGINAL ARTICLE

A Bowman–Birk protease inhibitor purified, cloned, sequenced and characterized from the seeds of *Maclura pomifera* (Raf.) Schneid

Martín Indarte[1] · Cristian M. Lazza[2] · Diego Assis[3] · Néstor O. Caffini[2] · María A. Juliano[3] · Francesc X. Avilés[4] · Xavier Daura[4] · Laura M. I. López[5,6] · Sebastián A. Trejo[7]

Received: 8 July 2016 / Accepted: 17 October 2016 / Published online: 24 October 2016

Abstract

Main conclusion **A new BBI-type protease inhibitor with remarkable structural characteristics was purified, cloned, and sequenced from seeds of *Maclura pomifera*, a dicotyledonous plant belonging to the Moraceae family.**

In this work, we report a Bowman–Birk inhibitor (BBI) isolated, purified, cloned, and characterized from *Maclura pomifera* seeds (MpBBI), the first of this type from a species belonging to Moraceae family. MpBBI was purified to homogeneity by RP-HPLC, total RNA was extracted from seeds of *M. pomifera*, and the 3′RACE-PCR method was applied to obtain the cDNA, which was cloned and sequenced. Peptide mass fingerprinting (PMF) analysis showed correspondence between the in silico-translated protein and MpBBI, confirming that it corresponds to a new plant protease inhibitor. The obtained cDNA encoded a polypeptide of 65 residues and possesses 10 cysteine residues, with molecular mass of 7379.27, pI 6.10, and extinction molar coefficient of 9105 M^{-1} cm^{-1}. MpBBI inhibits strongly trypsin with K_i in the 10^{-10} M range and was stable in a wide array of pH and extreme temperatures. MpBBI comparative modeling was applied to gain insight into its 3D structure and highlighted some distinguishing features: (1) two non-identical loops, (2) loop 1 (CEEESRC) is completely different from any known BBI, and (3) the amount of disulphide bonds is also different from any reported BBI from dicot plants.

Keywords BBI-type protease inhibitor · Cloning · Homology modeling · Loop · Three-dimensional structure · Trypsin inhibition

Abbreviations

BBI	Bowman–Birk inhibitor
MpBBI	*Maclura pomifera* Bowman–Birk inhibitor
PMF	Peptide mass fingerprinting

L. M. I. López and S. A. Trejo contributed equally.

Electronic supplementary material The online version of this article (doi:10.1007/s00425-016-2611-6) contains supplementary material, which is available to authorized users.

✉ Laura M. I. López
lmilopez@biol.unlp.edu.ar

✉ Sebastián A. Trejo
strejo@imbice.gov.ar

[1] PHusisTherapeutics, 3210 Merryfield Row, San Diego, CA 92121, USA

[2] Centro de Investigación de Proteínas Vegetales, Facultad de Ciencias Exactas, Universidad Nacional de La Plata, 1900 La Plata, Argentina

[3] Departamento de Biofisica, Universidade Federal de São Paulo, Vila Clementino, São Paulo 04044-020, Brazil

[4] Institut de Biotecnologia i de Biomedicina (IBB), Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain

[5] Instituto de Ciencias de la Salud, Universidad Nacional Arturo Jauretche, 1888 Florencio Varela, Buenos Aires, Argentina

[6] CITEC, Gonnet, B1897 Buenos Aires, Argentina

[7] Laboratorio de Neurofisiología del Instituto Multidisciplinar de Biología Celular (IMBICE), La Plata, B1906APO Buenos Aires, Argentina

Introduction

Proteolytic enzymes, also referred as peptidases, are essential for the survival of prokaryotic and eukaryotic organisms. They are encoded by approximately 2% of each genome. Several distinct mechanisms exist for the regulation of peptidase activity. One of such mechanisms is by peptides and proteins that directly or indirectly inhibit them (Rawlings et al. 2004). Protein peptidase inhibitors can be grouped into 74 families on the basis of primary sequence similarities as well as their 3D structure. Furthermore, 51 of those families are grouped into 38 clans, according to the MEROPS Database (Rawlings et al. 2012).

Relevant to this work, Bowman–Birk inhibitors (BBIs) belong to the MEROPS inhibitor family I12, clan IF. These inhibitors take their name from the trypsin and chymotrypsin inhibitor isolated from soybeans by Bowman (1946) and further characterized by Birk et al. (1963). Interestingly, BBIs can contain one, two, four, five or even more inhibitor units per molecule (Rawlings et al. 2004). They mainly inhibit serine peptidases of the S1 family, but also inhibit S3 peptidases (Laskowski and Kato 1980; Rawlings et al. 2004). Members of this family have a duplicated structure and generally possess two distinct inhibitory sites. These inhibitors are primarily found in plants and in particular in the seeds of legumes, as well as in cereal grains. Most legume BBIs inhibit trypsin at the first (N-terminal) reactive site and chymotrypsin at the second (C-terminal) reactive site. The amino acid residues in the reactive site of the inhibitors are usually designed as P4, P3, P2, P1, P′1, P′2, P′3, P′4, P′5, P′6. The P1 residue is involved in the so-called "primary" contact region and confers inhibitory specificity: Arg or Lys for trypsin; Leu, Phe, or Tyr for chymotrypsin; Ala for elastase (Laskowski and Kato 1980). In cereals, they exist in two forms, one of which is a duplication of the basic structure (Tashiro et al. 1987). BBIs interact with their counterpart enzymes via a solvent-exposed surface loop that adopts a canonical proteinase inhibitory conformation. Upon binding, the resulting noncovalent complex renders the proteinase inactive. This inhibitory mechanism is commonly seen in the majority of serine proteinase inhibitory proteins. A particular feature of the BBI protein, however, is that the interacting loop is a particularly well-defined, disulfide-linked, short beta-sheet region (McBride and Leatherbarrow 2001; McBride et al. 2002; Brauer et al. 2003). Some features make BBIs interesting as possible oral medication since they can withstand high temperatures and resist a wide range of pH fluctuations as well as proteolytic enzymes present in the gastrointestinal tract. Moreover, they are non-allergenic and bioavailable (Park et al. 2007; Losso 2008; Marin-Manzano et al. 2009).

From a drug discovery point of view, several human serine proteases (implicated in inflammatory responses) have been shown to be sensitive to BBI inhibition. Among them, elastase, alpha-chymotrypsin, chymase and cathepsin G can be found in that group (Larionova et al. 1993; Ware et al. 1997). Based on the above, finding additional and perhaps unknown serine proteases sensitive to BBI inhibition can be expected in the near future (Dai et al. 2012). BBIs have been shown to be effective suppressors of carcinogenesis and to diminish inflammatory responses (Qi et al. 2005; Souza et al. 2014; Utrilla et al. 2015). Recently, it has been reported that BBIs express powerful suppressive effects on tumor progression of two prevalent cancers: gastric adenocarcinoma and colorectal adenocarcinoma (Fereidunian et al. 2014). This inhibition has also been reported in autoimmune encephalomyelitis with the positive additional effect to attenuate neuronal loss, making it an attractive candidate for oral therapy in multiple sclerosis (Touil et al. 2008; Dai et al. 2012).

In the present study, MpBBI emerges as the first BBI isolated from the family Moraceae that has been systematically purified, cloned and structural/functionally characterized using both experimental and in silico techniques.

Materials and methods

Plant material

Fruits of *Maclura pomifera* (Raf.) Schneid. (Moraceae) were collected in La Plata, Argentina. Individual fruits were carefully cleaned with distilled water, chopped and seeds manually harvested and stored at −20 °C. As customary, a voucher specimen was deposited by Dr. Ana María Arambarri in the Herbarium of Área de Botánica, Facultad de Ciencias Agrarias y Forestales, Universidad Nacional de La Plata, Argentina, under the accession code LPAG 5445.

Crude extract preparation and preliminary purification

A crude extract was obtained from seeds of *M. pomifera* following a previously described method (Lazza et al. 2010). Briefly, seeds (approximately 50 g) were crushed with 250 mL of 50 mM Tris–HCl buffer (pH 7.2) containing 0.15 M NaCl, the homogenate was centrifuged and the resulting supernatant was filtered to obtain a crude extract (CE). Then, one volume of crude extract was treated with four volumes of chilled acetone at −20 °C and the precipitate was redissolved with 50 mM Tris–HCl buffer (pH 7.2); this preparation, named "redissolved acetone precipitate" (RAP), was freeze-dried and stored.

Chromatographic purification

Size-exclusion chromatography of RAP (first purification step) was carried out on a K15/30 column filled with Sephadex G50 Fine (GE Healthcare-Bioscience AB, Uppsala, Sweden). RAP proteins resolved into three main fractions; aliquots that exhibited inhibitory activity higher than 50% were pooled. Active aliquots were submitted to anion-exchange chromatography using an Äkta Purifier and a 5 mL HiTrap Q HP column (GE Healthcare-Bioscience AB), equilibrated with 50 mM Tris–HCI buffer at pH 7.2. After a washing step with the aforementioned buffer, proteins were eluted with a 0–0.3 M NaCl linear gradient at a flow rate of 1 mL min^{-1}. The main fraction, which exhibited inhibitory activity, was then pooled and submitted to reverse-phase (RP)-HPLC. RP-HPLC was carried out on a Phenomenex C18 analytical column (i.d. = 4.6 mm, length = 150 mm with a 5 µm particle size) using a GE HPLC system (Äkta Purifier 10) equilibrated with 0.1% (v/v) trifluoroacetic acid (TFA) in deionized water. A 0.5 mg aliquot of lyophilized protein was injected into the column, and its separation was achieved using an acetonitrile gradient (20–30%, 60 min) in 0.1% (v/v) TFA at a flow rate of 1 mL/min. The eluent was monitored at 215 nm. Once again, the main fraction showing inhibitory activity was pooled and re-submitted to RP-HPLC to further purify it at the same conditions previously described.

Determination of trypsin inhibitory activity

Antitrypsin activity for bovine trypsin was assayed using the chromogenic synthetic substrate *N*α-benzoyl-L-arginine 4-nitroanilide hydrochloride (BAPNA) in 0.05 M Tris–HCl buffer (pH 8.0), containing 0.04 M $CaCl_2$ at 37 °C according to Erlanger et al. (1961) with minor modifications. Briefly, 100 µL of trypsin (13 µM) solution and 100 µL of sample or buffer (considered as a blank) were incubated for 1 min at 37 °C, then 750 µL of buffer was added and the assay was initiated by the addition of 50 µL of 20 mM BAPNA in DMSO. The reaction rate was determined by measuring the absorbance at 410 nm for 3 min. For accuracy, the reaction was measured in the linear portion in the 40–60% inhibition range.

Effect of temperature and pH on the stability of the inhibitor

To test the MpBBI thermal stability, several inhibitor solutions (pH 7.0) were heated for 30 min in water bath at various temperatures (37–100 °C) and then cooled to 0 °C. Residual trypsin inhibitory activity was measured using BAPNA as substrate at 37 °C.

To measure the pH stability of the inhibitory activity, samples of the inhibitor previously lyophilized were dissolved using Good's buffers (Good et al. 1966) at different pH values (2.0, 3.0, 4.0, 5.0, 6.0, 7.0, 8.0, 9.0, 10.0, 11.0, 12.0, and 13.0). These mixtures were kept for 30 min at 37 °C and quickly afterward adjusted back to pH 8.0. Residual trypsin inhibitory activity was evaluated as described above. Analysis of variance (ANOVA) was performed using GraphPadInStat version 3.10, for Windows, GraphPad Software (La Jolla, CA, USA; http://www.graphpad.com.

Inhibitory activity and inhibition constant (K_i) determination

The substrate Abz-KLRSSKQ-EDDnp was used; fluorescence resonance energy transfer (FRET) peptide hydrolysis was carried out at λ of 320 nm and λ of 420 nm (excitation and emission wavelengths, respectively, for 2-amino-benzoic acid, Abz) in a Shimadzu RF-1501 spectrofluorometer (Shimadzu Corporation, Kyoto, Japan) previously calibrated with standard solutions of hydrolyzed Abz (Korkmaz et al. 2008). The enzymes evaluated were trypsin and subtilisin A. Determinations were performed at 37 °C in 50 mM Tris–HCl buffer containing 5 mM $CaCl_2$, pH 8.0. Product formation was monitored by the increase of fluorescence. Peptidase inactivation kinetics was obtained according to the following equation: $[P] = \frac{v_z}{\lambda}(1 - e^{-\lambda t}) + d$, where $[P]$ is the product concentration formed by the hydrolysis of the FRET peptide, v_z is the velocity of substrate hydrolysis for zero time, λ is the first-order rate constant and d is the basal fluorescence before addition of the enzyme. Reactions were continuously monitored by fluorescence measurements of hydrolysis of the substrates. The values of K_i and k_4 were derived from the relation of the first-order rate constant (λ) with the inhibitor concentration $[I]$, where the association rate constant is described for $k4$, I represents the inhibitor concentration, S the substrate concentration, and Km is the *Michaelis–Menten* constant: $\lambda = \frac{k_4 \times [I]}{K_i \times \left(1 + \frac{[S]}{Km}\right) + [I]}$. The inhibitory reaction was done in different concentrations of inhibitors, and all the plots of λ vs $[I]$ were non-linear. The K_i and $k4$ parameters were obtained by non-linear regression using the Grafit software (Erithacus Software, Horley, Surrey, UK) according to the abovementioned equation described by Baici et al. (2009). Errors were less than 5% for each of the obtained kinetic parameters.

Molecular weight determination by matrix-assisted laser desorption/ionization time of flight (MALDI-TOF) MS

MALDI-TOF MS was used to determine both molecular weight and purity of MpBBI. These kinds of analyses were

performed in a Bruker ultrafleXtrem TOF/TOF spectrometer. Samples were prepared by mixing equal volumes of a saturated solution of the matrix (sinapinic acid) in 0.1% TFA in water/acetonitrile (2:1), and a 1–10 μM protein solution. From this mixture, 1 μL was spotted on the sample slide and allowed to evaporate to dryness. Protein Calibration Standard I (Bruker Daltonics) was used as standard for mass calibration.

Peptide mass fingerprinting (PMF) assay

MpBBI purified by RP-HPLC was denaturated with 6 M guanidinium chloride and 50 mM ammonium bicarbonate. Then, all Cys residues were reduced and alkylated with 10 mM DTT and 25 mM iodoacetamide, respectively. Finally, it was digested with 25 ng of trypsin (Gold-trypsin for sequencing grade, Promega) for 3 h at 37 °C.

MALDI-TOF MS was used for protein identification by PMF. Analyses were performed using an ultrafleXtrem MALDI-TOF mass spectrometer (Bruker Daltonics, Bremen, Germany). The sample was spotted on a MTP 384 GroundSteel and mixed with freshly prepared matrix solution [10 mg/mL of α-cyano-4-hydroxy-cinnamic acid (HCCA) in aqueous solution containing 30% acetonitrile and 0.1% TFA]. An external calibration was performed using standard peptide calibrants. Peptide masses were acquired within the 800–4000 m/z ca range. Comprehensive peak assignments were accomplished using the BioTools software package (Bruker Daltonics). The routinely used MASCOT search tool was applied for the identification of tryptic maps (http://www.matrixscience.com).

Cloning and sequence analysis of MpBBI cDNA

Total RNA was extracted from unripe *M. pomifera* seeds and the 3′RACE (rapid amplification of cDNA ends)-PCR method was applied to obtain MpBBI cDNA. About 2 g of unripe seeds was crushed in a mortar with liquid nitrogen. Fractions of 0.1 g were mixed with 350 μL of Kit buffer (NucleoSpin® RNA Plant, Macherey–Nagel, Düren, Germany), containing 3.5 μL of β-mercaptoethanol. Total RNA was extracted as recommended by the manufacturers, using spin columns that trap nucleic acids. Single-stranded cDNA was prepared by 3′-RACE using the First-Strand cDNA Synthesis Kit for RT-PCR, AMV (Roche Diagnostics-Roche Applied Science, Indianapolis, IN, USA) and a specific polydT primer (RoR1polydT = 5′-CCGGAATTCACTGCAGGGTACCCAATACGACTCACTATAGGGCTTTTTTTTTTTTTTTT-3′). The N-terminal sequence (Lazza et al. 2010) was used to design a specific oligonucleotide primer (5′-GCTAGGGAACCTAAGTTTTCTACTCATTGCGA-3′) to obtain the complete cDNA. PCR products were analyzed by agarose gel electrophoresis and the cDNA was cloned in *E. coli* (XL1-Blue) using the pGEM-T Promega vector. Clones containing the plasmidic cDNA with the abovementioned insert were sequenced. The consensus sequence of the serine peptidase inhibitor was obtained by the analysis of the information from sequencing (DNA Baser).

MpBBI homology modeling

The MpBBI FASTA sequence was obtained from UNIPROT (http://www.uniprot.org) with the following code: I7JDE8_MACPO. A search in the protein data bank (PDB, http://www.rcsb.org) was carried out utilizing "Bowman–Birk" as a search term. Three-dimensional structures were downloaded and loaded into the molecular operating environment (MOE) software for further exploration. BBI-like molecules were isolated from molecular complexes (if present) and their FASTA sequences were aligned using a multiple sequence alignment utilizing MOE default values (blosum62 matrix, tree-based buildup, gap start 7, iteration limit 100, gap extend 1, failure limit 1 and with their actual secondary structure). The BBI-like molecule with the maximum sequence identity to MpBBI (2ILN) was curated and employed as template for homology modeling in default mode. These structures were checked and corrected for chemically consistent atom and bond type. Hydrogen atoms were added using the all-atom force field AMBER99 and partial charges assigned to all models (Ponder and Case 2003). Relaxation was carried out to convergence criteria of 0.05 kcal/mol and a dielectric constant (ε) of 3. Water molecules (if present) were deleted. Monomers of bound trypsin molecules were included. All non-hydrogen atoms were held fixed during the energy minimization step. The optimized proteins were then saved in an MOE molecular database. Homology modeling was performed with the "homology model" feature of MOE using the MpBBI FASTA sequence aforementioned and with 2ILN as template in default mode. The final model was subjected to the Protonate 3D application to optimize ionization states and proton placement in the closing model prior to the final refinement step.

Results

Purification

The crude extract obtained from mature seeds of *M. pomifera* was purified by acetone precipitation, gel filtration and ion-exchange chromatography. This purification scheme allowed the isolation of a fraction with trypsin inhibitory activity, which presented a single band in SDS-PAGE (Lazza et al. 2010). However, more sensitive MALDI-TOF MS analysis indicated the presence of

multiple isoforms (data not shown). Therefore, the inhibitory fraction from exchange chromatography was submitted to RP-HPLC; this way five major molecular forms with inhibitory activity were detected and separated (Fig. 1a). The fraction 4 was selected and purified by rechromatography (Fig. 1b), its homogeneity was confirmed by MALDI-TOF MS and its molecular mass was of 6653.55 Da (Suppl. Fig. S1). This protease inhibitor was named MpBBI (*Maclura pomifera* Bowman–Birk inhibitor) since as mentioned above this inhibitor belongs to the type BBI family of inhibitors. The inhibitor was subjected to PMF analysis using MASCOT (URL http://www.matrixscience.com) as a search tool to identify tryptic maps. No matches with other plant inhibitor were found suggesting that this peptidic sequence corresponds to a new plant protease inhibitor.

Inhibitory properties and stability

The MpBBI was able to strongly inhibit serine protease trypsin, but no inhibition was detected for the non-specific serine protease subtilisin A. The dissociation constant of protease inhibitor complex (K_i) value was 0.66 nM, while the relationship $k4/K_i$ (the inactivation constant) was 33,636 mM^{-1} s^{-1} (Fig. 2). MpBBI's thermal and pH stabilities wwere tested under extreme temperature and pH ranges. It was found to be heat stable when subjected to 100 °C for 10 min at neutral pH (Fig. 3a). MpBBI was also stable in a wide array of pH solutions; after being left at 37 °C for 30 min, it retained its full trypsin inhibitory activity in the pH range of 2–13 (Fig. 3b).

Cloning and sequence analysis of MpBBI

The consensus sequence was obtained from four selected clones using the CLUSTAL W multiple sequence alignment program. Nucleotide sequences of the four clones were identical. This sequence was deposited in the European Nucleotide Archive (http://www.ebi.ac.uk/ena) under the CCI55634.2 code. The aforementioned protease inhibitor is constituted by 405 pb, of which the 195 first ones codify for a polypeptide chain. The mature inhibitor has 65 amino acids and possesses 10 cysteine residues. Basic physicochemical properties were calculated: molecular mass, 7379.3; pI, 6.10 and extinction molar coefficient, 9105 M^{-1} cm^{-1} (assuming all pairs of Cys residues form disulfide bonds); data were obtained by analysis with ProtParam tool from the ExPASy Proteomics server (http://www.expasy.org).

The PMF for the chromatographically purified trypsin inhibitor was compared to the theoretical fingerprint of the inhibitor cloned and sequenced from *M. pomifera* seeds. More than 75% of intensity coverage of the spectrum was explained by the theoretical sequence of the chromatographically purified inhibitor, with a 100% of the sequence

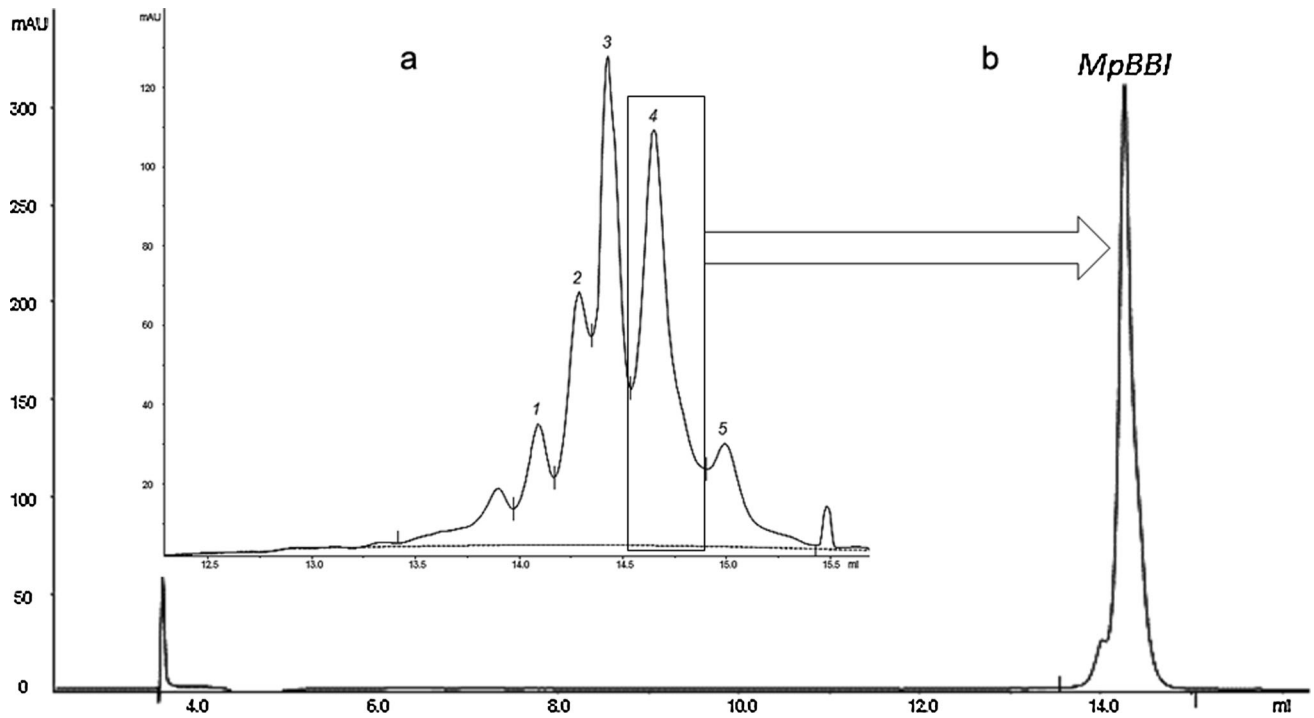
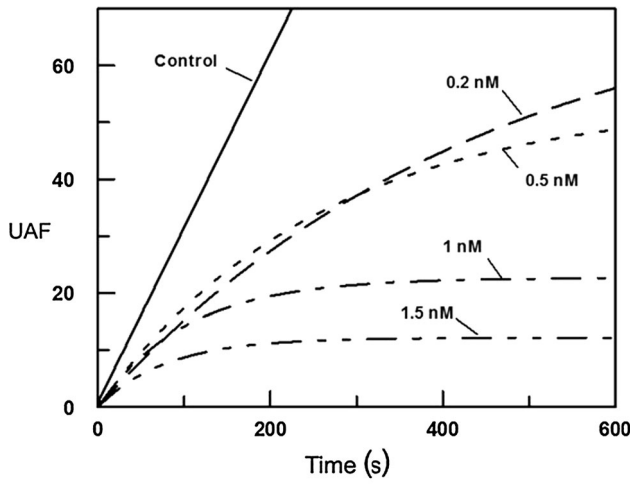


Fig. 1 Reverse-phase-HPLC of the inhibitory fraction and rechromatographic analysis of fraction 4. The inhibitory fraction from exchange chromatography was submitted to reverse-phase chromatography using HPLC (**a**). Fraction 4 rechromatography (**b**) allows obtaining the characterized MpBBI



K_i (nM)	k4 (s^{-1})	k4 / K_i (mM^{-1} s^{-1})
0.66	0.0222	33636

Fig. 2 Inhibitory activity of MpBBI against trypsin. The inhibitor found was able to inhibit trypsin. Inhibition kinetic parameters are shown at the *bottom*; errors were less than 5% for each of the parameters

coverage (Fig. 4; Table 1). However, it should be noted that the peptide sequence (I7JDE8_MACPO) derived from nucleotide sequence, contained six extra amino acids at C terminus (VMPHQQ) versus the purified inhibitor sequence.

Homology modeling

The Bowman–Birk inhibitor from *Medicago scutellata* seeds (PDB: 2ILN; 34% identity and 40% similarity with MpBBI) was used as a template for 3D elucidation via homology modeling. The best scored (GBVI scoring) homology model of the MpBBI superimposed onto the 2ILN X-ray showed a similar overall 3D shape with two loops and five intramolecular disulfide bridges (Figs. 5, 6). The Ramachandran plot of the best scored model suggested that model quality was optimal (Fig. 6) with a vast majority of amino acids falling within allowed regions. MpBBI showed a non-symmetrical "double-headed" arrangement that is usually termed loop 1 and 2 in the BBIs. The most conserved region (CTRMNPPQC), which corresponds to loop 2 (Figs. 5, 6, 7), relates to the canonical loop "formula", $CTP_1SXPPQC$ seen in other BBIs (P_1

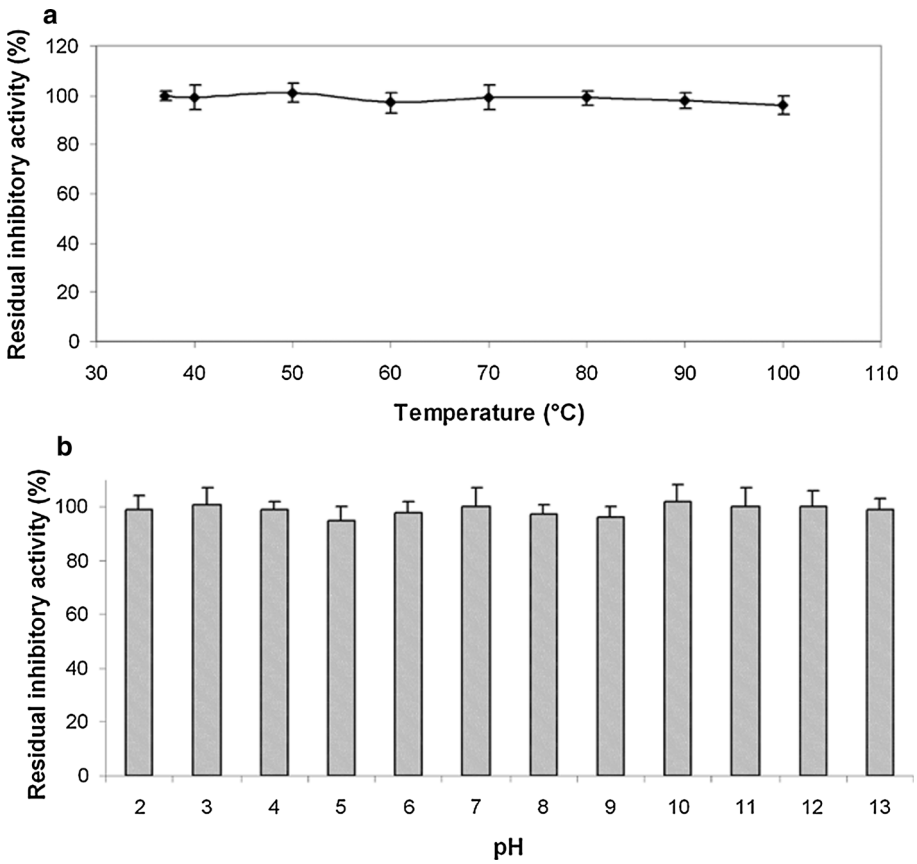


Fig. 3 Residual inhibitory activity of MpBBI under extreme temperature and pH values. The inhibitor found was able to inhibit trypsin from around 30–100 °C (**a**) and in pH values ranging from 2 to 14 (**b**)

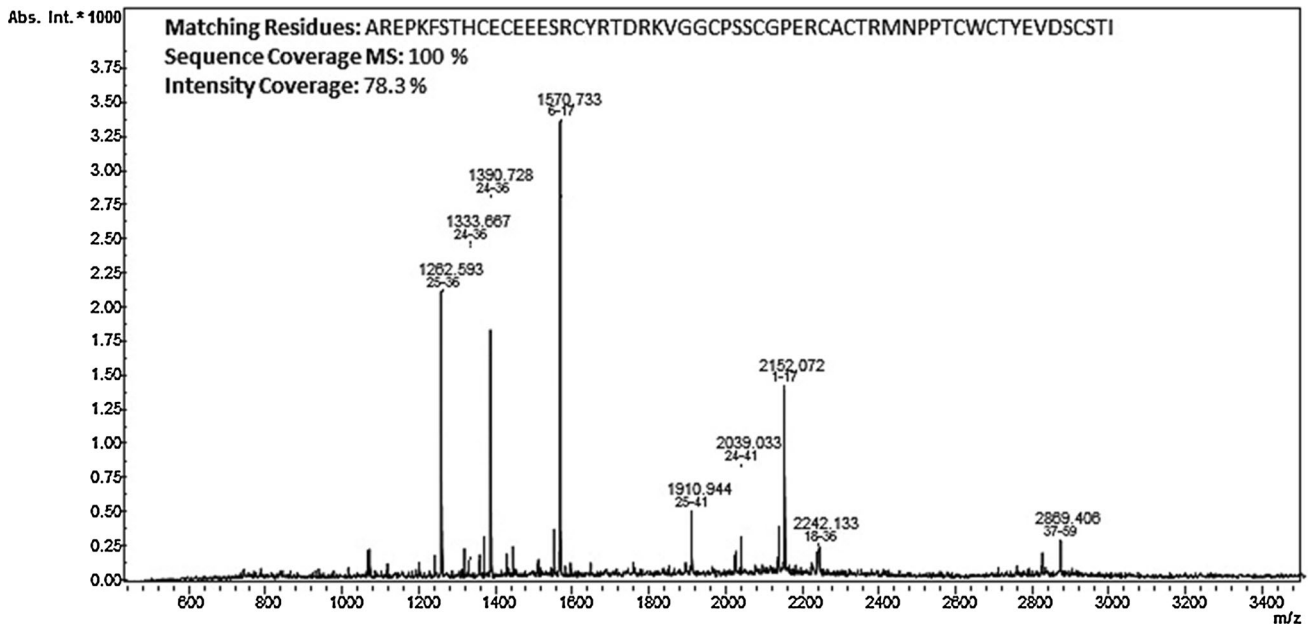


Fig. 4 Sequence validation of MpBBI by PMF. The MpBBI purified from *M. pomifera* seeds was digested with trypsin (previously reduction and alkylation of Cys with 10 mM DTT and 50 mM IAA). The digestion mixture was analyzed by MALDI-TOF MS and compared with the theoretical fingerprint of the inhibitor cloned and sequenced

Table 1 Identification of trysin-digested peptides of inhibitor purified with theoretical PMF of MpBBI

Experimental PMF (m/z)	Theoretical PMF (m/z)	Fragment position	Peptide fragment sequence
2152.072	2151.918	1–17	AREPKFSTHCECEEESR
1570.733	1570.590	6–17	FSTHCECEEESR
2242.133	2241.991	18–36	CYRTDRKVGGCPSSCGPER
1390.728	1390.620	24–36	KVGGCPSSCGPER
1262.593	1262.525	25–36	VGGCPSSCGPER
2039.033	2038.867	24–41	KVGGCPSSCGPERCACTR
2869.406	2869.118	37–59	CACTRMNPPTCWCTYEVDSCSTI
MpBBI full-length sequence			AREPKFSTHCECEEESRCYRTDRKVGGCPSSCGPERCACTRMNPPTCWCTYEVDSCSTI

indicates the residue that determines the specificity of inhibition; X indicates any of the 20 residues). It has been shown that if Arg or Lys is present in the P_1 position, antitrypsin activity can be expected (Qi et al. 2005). MpBBI-Arg41 is deeply buried in the catalytic trypsin-binding site similar to the BBI from 2ILN. This residue seems to be close enough to trypsin Asp189 side chain and Gly219 backbone to establish ionic and hydrogen bond interactions, respectively. Thus, the Arg41 residue in the P1 of MpBBI position allows to explain the subnanomolar trypsin inhibition activity. A structural particularity about this loop is that in MpBBI, Ser is replaced by a Met, diminishing the hydrogen bond potential but increasing the hydrophobic potential of that region.

Strangely, the MpBBI loop 1 (CEEESRC) showed no sequence identity or similarity compared to other BBIs. The calculation of electrostatic potential maps of crystallized BBIs vs MpBBI showed a sharp contrast between them and MpBBI in those regions (Fig. 7). MpBBI revealed a large electronegative region (given by the three aspartates) entirely different from the electropositive regions of BBIs (2ILN, 3MYW, 2G83, and 2R33). The BBIs electropositive regions (Arg10-2ILN, Lys10-2G83, Lys10-2R33, and Lys9-3MYW) are flanked by polar (Thr) and neutral (Ile, 2G83, 2R33 and 3MYW) amino acids, except for 3MYW that has an extra electropositive region.

Discussion

Protease inhibitors from legumes and non-legumes are effective in preventing or suppressing carcinogenic processes in a wide variety of in vitro and in vivo model

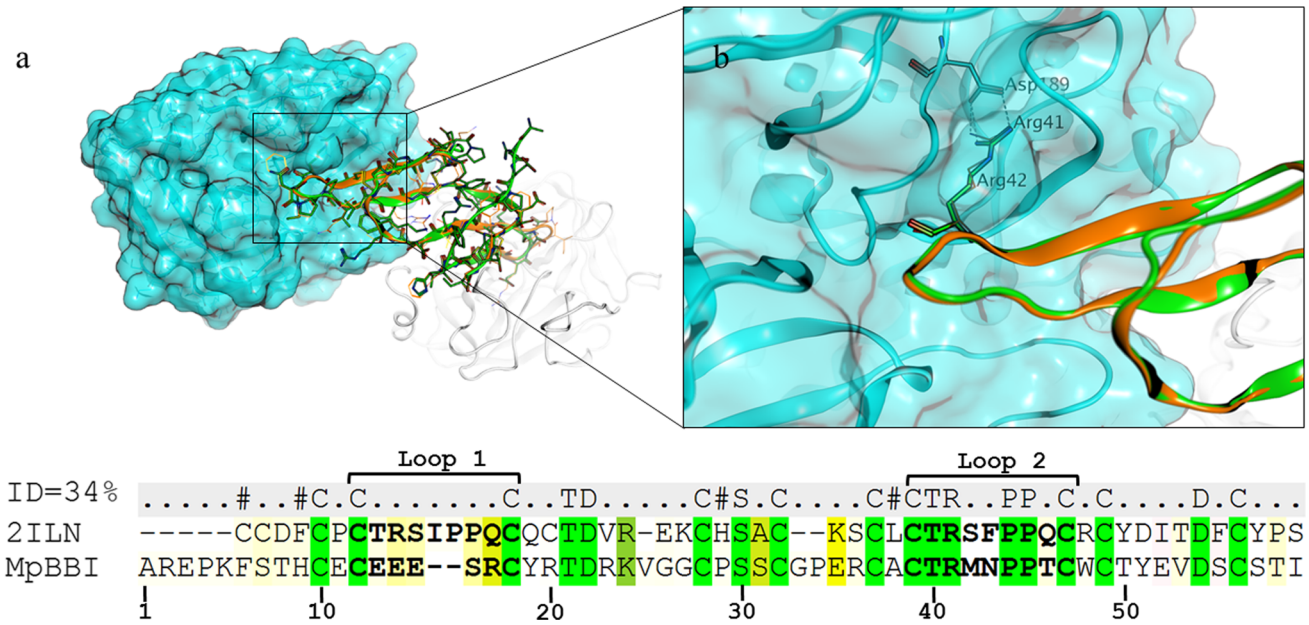


Fig. 5 Homology model of MpBBI and trypsin Connolly surface based on 2ILN. The 2ILN-based MpBBI (*green* atoms and ribbons) model and X-ray (*orange* atoms and ribbons) can be seen in (**a**, **b**). The sequence alignment used in model building is shown at the *bottom*; residues are *colored* based on their sequence identity

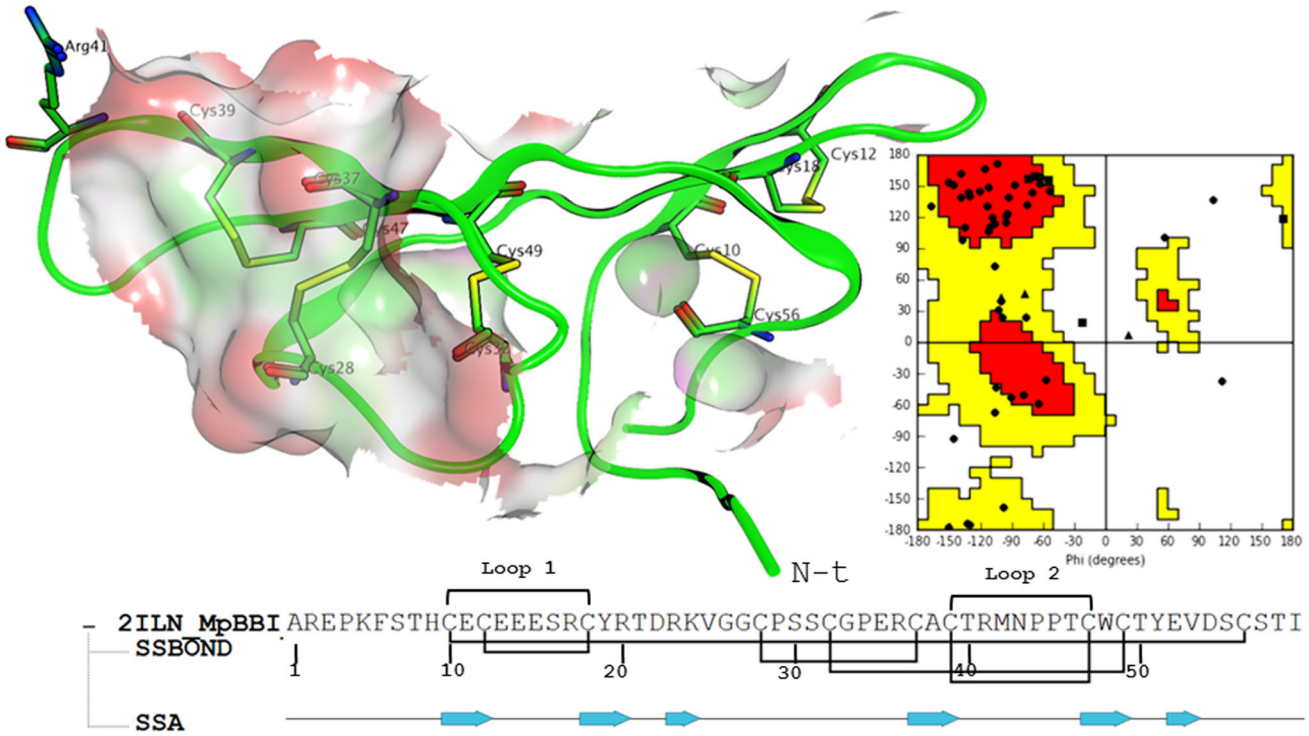


Fig. 6 Three-dimensional structure of 2ILN-based MpBBI, Ramachandran plot and solvent-exposed surface of its disulfide bonds. The molecular surface of the five disulfide bridges seen in 2ILN-MpBBI was *colored* based on its solvent exposure (*red*, solvent exposed, *green* hydrophobic and magenta hydrophilic features). All disulfide bonds showed complete solvent exposure and optimal Ramachandran plot allocation of model amino acids

systems (Kennedy 1998). Particularly, BBIs are the most characterized inhibitors for their role as carcinogenesis suppressors (Souza et al. 2014). BBIs are found in mono- and dicotyledons, especially in leguminous seeds, but there have been no reports so far of BBIs isolated from species belonging to Moraceae family (dicot plants).

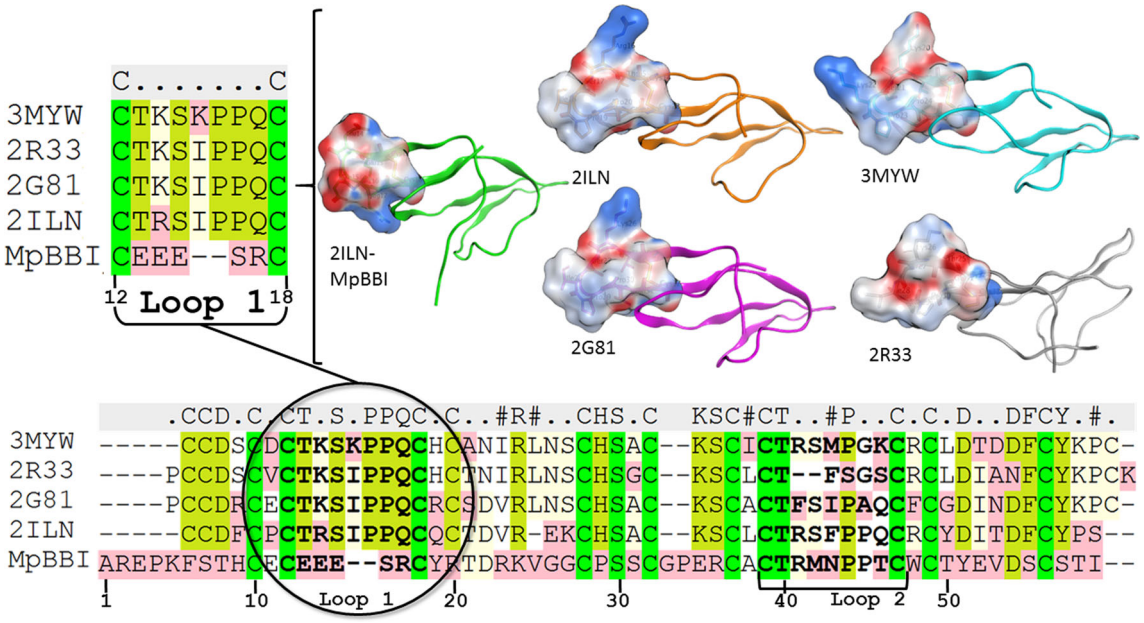


Fig. 7 Loop 1 electrostatic potential surface of MpBBI vs. high sequence identity dicots BBIs. Multiple sequence alignment helps to highlight the MpBBI (2ILN-derived) unusually large electronegative region (CEEESRC), termed Loop 1 when compared to other BBIs that possess electropositive regions

From seeds of *M. pomifera* (Moraceae) using a combined chromatographic approach was isolated and purified to homogeneity a new protease inhibitor by MALDI-TOF MS. The inhibitor showed remarkable thermal and pH stabilities consistent with previous findings that most of BBIs are stable to pH and heat, while Kunitz type inhibitors are more sensitive to heat (Deshimaru et al. 2002; Osman et al. 2002). The inhibitor named MpBBI displayed potent inhibitory activity towards trypsin proteases (with K_i value in the 10^{-10} M range); similar inhibitory capacity was also determined for Bowman–Birk proteinase inhibitor from *Clitoria fairchildiana* seeds (Dantzger et al. 2015). BBI-type inhibitors isolated from different leguminous plants possessed K_i values for trypsin between 0.1 and 52.0 nM (Prasad et al. 2010). The molecular mass of MpBBI was also similar to the other BBIs (Birk 1985; Swathi et al. 2014; Dantzger et al. 2015). These facts suggest the presence of a novel molecular entity from *M. pomifera* seeds that may be related to BBIs.

Further, the corresponding cDNA was cloned and sequenced. A default BLAST search allocated this inhibitor within the BBI category, due to its common residues with other BBIs. Based on all of the above, this inhibitor has been properly termed as an MpBBI.

PMF of fraction chromatographically purified explained more than 75% of the intensity coverage for the inhibitor obtained from cloning and sequencing from the deposited nucleotide sequence with UNIPROT code I7JDE8_MACPO (Fig. 4). This suggests that the fraction was originated from the aforementioned gene. Interestingly, the nucleotide sequence has some extra six amino acids at the C terminus (VMPHQQ) versus the purified inhibitor sequence. The presence of multiple genes and the possibility of hydrolysis support the idea that a large number of isoinhibitors can be present and coexist within *M. pomifera* seeds (Kalume et al. 1995; Wang et al. 2014).

In classical BBIs, the presence of electropositive amino acids in loop 1 allows recognition and binding to the electronegative active sites of trypsin and chymotrypsin. Strikingly, the MpBBI model showed a remarkably large electronegative sequence in loop 1 consisting of three aspartic acid residues (Fig. 7). This observation may suggest characteristic or distinctive biological activities that are yet to be defined. To our knowledge, in BBIs this type of loop (CEEESRC) has never been previously described in the literature.

Joshi et al. (2013) have made a summary of the number of disulphide bonds that can be found in the most common families of plant serine proteases inhibitors; dicotyledons produce BBIs with seven disulphide bonds and monocotyledons with five disulphide bonds. However, as it is shown in Fig. 6, the MpBBI model (2ILN-based) showed five solvent-exposed disulphide bonds (C10–C56, C12–C18, C28–C37, C32–C49, C39–C47) which would make MpBBI the first dicot with five disulphide bonds. Under the light of the distinctive loop 1 of MpBBI, it may not be surprising to find other differences with dicotyledons, which may well be their disulphide bonds content. A model that suggests that all MpBBI sulfhydryls are forming disulphide bonds could have additional weight from a

stability point of view. This can be related to a decrease in entropy due to disulphide bonds formation, which would favor the folded state over the unfolded state. In turn, the folded state could help to stabilize the native conformation of plant serine protease inhibitors and the disulphide bonds protect the protein from cellular damage to maintain its integrity. All of these findings suggest that the intramolecular interactions and loop amino acid content are collectively responsible for the functional activity of the BBIs. This can occur through the maintenance of the scaffold framework, conformational rigidity and shape complementarity of interacting loops against their counterpart active sites (Joshi et al. 2013).

Conclusions

In the present work, a combined approach to isolate, purify, clone, characterize and elucidate the 3D structure of a novel Bowman–Birk inhibitor molecular entity was applied. Since this is a new BBI-type protease inhibitor obtained from the seeds of a dicotyledonous plant belonging to the Moraceae family, no structural information has been described up to date. To surmount these shortcomings and to shed some light on its 3D structure, in silico approaches such as homology modeling and electrostatic surface calculations were performed. An X-ray structure with optimal sequence identity to MpBBI (2ILN) was found and used as a template for 3D elucidation via homology modeling. From the comparative model created, Arg41 (P1) emerged as a key interacting residue deeply buried in trypsin S1 (catalytic) site. The main interaction established seems to be Arg41–Asp189 ionic bond. The model revealed unique characteristics: (1) their double-headed regions are substantially different from each other; (2) the amount of disulphide bonds lies in the monocot group. Since *M. pomifera* is a dicot, it could be the first BBI from this plant group not having seven disulphide bonds. Perhaps the most remarkable modeling finding here is that one of the non-identical heads (loop 1) of MpBBI possesses a unique sequence of acidic amino acids (CEEESRC) that confers that head a distinctive electronegative nature, which may or may not be involved in inhibitory activity. In contrast, loop 2 is the most conserved region (CTRMNPPQC) in MpBBI related to other BBIs. These findings suggest that MpBBI possesses unique structural characteristics that could result in singular biological activities.

Author contribution statement LM.I.L. and S.A.T. conceived and designed research. The MALDI-TOF MS analyses were carried out by S.A.T. M.I. conducted homology modeling analysis and wrote the manuscript. C.M.L. carried out purification and characterization experiments. D.M.A. conducted kinetics experiments, M.A.J. and X.D. helped with methodological approaches. N.O.C. and F.X.A. participated in the design of the study. All authors read and approved the final manuscript.

Acknowledgements This work was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP-Project-12/50191-4R, D.A. and M.A.J.), CYTED (Red temática PROMAL 210RT0398, C.M.L., N.O.C., F.X.A., L.M.I.L. and S.A.T.) and CONICET (PIP 0297, L.M.I.L. and S.A.T.). The MALDI-TOF MS analyses were carried out in the Proteomics and Bioinformatics Facility of the Universitat Autònoma de Barcelona (SePBioEs-UAB) by S.A.T.

Compliance with ethical standards

Conflict of interest The authors declare that they have no competing interests.

References

Baici A, Schenker P, Wächter M, Rüedi P (2009) 3-Fluoro-2,4-dioxa-3-phosphadecalins as inhibitors of acetylcholinesterase. A reappraisal of kinetic mechanisms and diagnostic. Methods Chem Biodiv 6:261–282

Birk Y (1985) The Bowman–Birk inhibitor. Int J Pept Protein Res 25:113–131

Birk Y, Gertler A, Khalef S (1963) A pure trypsin inhibitor from soya beans. Biochem J 87:281–284

Bowman DE (1946) Differentiation of soy bean antitryptic factors. Proc Soc Exp Biol Med 63:547–550

Brauer AB, Nievo M, McBride JD, Leatherbarrow RJ (2003) The structural basis of a conserved *P2* threonine in canonical serine proteinase inhibitors. J Biomol Struct Dyn 20:645–656

Dai H, Ciric B, Zhang G-X, Rostami A (2012) Interleukin-10 plays a crucial role in suppression of experimental autoimmune encephalomyelitis by Bowman–Birk inhibitor. J Neuroimmunol 245:1–7

Dantzger M, Ilka MV, Scorsato V, Marangoni S, Rodrigues Macedo ML (2015) Bowman–Birk proteinase inhibitor from Clitoria fairchildiana seeds: isolation, biochemical properties and insecticidal potential. Phytochemistry 118:224–235

Deshimaru M, Hanamoto R, Kusano C, Yoshimi S, Terada S (2002) Purification and characterization of proteinase inhibitors from wild soja (*Glycine soja*) seeds. Biosci Biotechnol Biochem 66:1897–1903

Erlanger B, Kokowsky N, Cohen W (1961) The preparation and properties of two new chromogenic substrates of trypsin. Arch Biochem Biophys 95:271–278

Fereidunian A, Sadeghalvad M, Oscoie MO, Mostafaie A (2014) Soybean Bowman–Birk protease inhibitor (BBI): identification of the mechanisms of BBI suppressive effect on growth of two adenocarcinoma cell lines: AGS and HT29. Arch Med Res 45:455–461

Good NE, Winget GD, Winter W, Connolly TN, Izawa S, Singh RMM (1966) Hydrogen ion buffers for biological research. Biochemistry 5:467–477

Joshi RS, Mishra M, SureshCG GuptaVS, Giri AP (2013) Complementation of intramolecular interactions for structural-functional stability of plant serine proteinase inhibitors. Biochim Biophys Acta 1830:5087–5094

Kalume E, Sousa MV, Morhy L (1995) Purification, characterization, sequence determination, and mass spectrometric analysis of a trypsin inhibitor from seeds of the Brazilian tree *Dipteryxalata* (Leguminosae). J Protein Chem 14:685–693

Kennedy AR (1998) Chemopreventive agents: protease inhibitors. Pharmacol Ther 78:167–209

Korkmaz B, Attucci S, Juliano MA, Kalupov T, Jourdan ML, Juliano L, Gautier F (2008) Measuring elastase, proteinase 3 and cathepsin G activities at the surface of human neutrophils with fluorescence resonance energy transfer substrates. Nat Protoc 3:1–9

Larionova NI, Gladysheva IP, Tikhonova TV, Kazanskaia NF (1993) Inhibition of cathepsin G and elastase from human granulocytes by multiple forms of the Bowman–Birk type of soy inhibiton. Biokhimiia 58:1437–1444

Laskowski M, Kato I (1980) Protein inhibitors of proteinases. Annu Rev Biochem 49:593–626

Lazza CM, Trejo SA, Obregón WD, Pistaccio LG, Caffini NO, López LMI (2010) A novel trypsin and α-chymotrypsin inhibitor from *Maclura pomífera* seeds. Lett Drug Design Discov 7:244–249

Losso JN (2008) The biochemical and functional food properties of the Bowman–Birk inhibitor. Crit Rev Food Sci Nutr 48:94–118

Marin-Manzano MC, Ruiz R, Jimenez E, Rubio LA, Clemente A (2009) Anti-carcinogenic soybean Bowman–Birk inhibitors survive fecal fermentation in their active form and do not affect the microbiota composition in vitro. Br J Nutr 101:967–971

McBride JD, Leatherbarrow RJ (2001) Synthetic peptide mimics of the Bowman–Birk inhibitor protein. Curr Med Chem 8:909–917

McBride JD, Watson EM, Brauer AB, Jaulent AM, Leatherbarrow RJ (2002) Peptide mimics of the Bowman–Birk inhibitor reactive site loop. Biopolymers 66:79–92

Osman MA, Reid PM, Weber CW (2002) Thermal inactivation of tepary bean (*Phaseolus acutifolius*), soybean and lima bean protease inhibitors: effect of acidic and basic pH. Food Chem 78:419–423

Park JH, Jeong HJ, de Lumen BO (2007) In vitro digestibility of the cancer-preventive soy peptides lunasin and BBI. J Agric Food Chem 55:10703–10706

Ponder JW, Case DA (2003) Force fields for protein simulations. Adv Protein Chem 66:27–85

Prasad ER, Dutta-Gupta A, Padmasree K (2010) Purification and characterization of a Bowman–Birk proteinase inhibitor from the seeds of black gram (*Vigna mungo*). Phytochemistry 71:363–372

Qi RF, Song ZW, Chi CW (2005) Structural features and molecular evolution of Bowman–Birk protease inhibitors and their potential application. Acta Biochim Biophys Sinica 37:283–292

Rawlings ND, Tolle DP, Barrett AJ (2004) Evolutionary families of peptidase inhibitors. Biochem J 378:705–716

Rawlings ND, Barrett AJ, Bateman A (2012) MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. Nucleic Acids Res 40:D343–D350

Souza LDC, Camargo R, Demasi M, Santana JM, Sá CM, de Freitas SM (2014) Effects of an anticarcinogenic Bowman–Birk protease inhibitor on purified 20S proteasome and MCF-7 breast cancer cells. PLoS One 9(1):e86600

Swathi M, Lokya V, Swaroop V, Mallikarjuna N, Kannan M, Gupta AD, Padmasree K (2014) Structural and functional characterization of proteinase inhibitors from seeds of *Cajanus cajan* (cv. ICP 7118). Plant Physiol Biochem 83:77–87

Tashiro M, Hashino K, Shiozaki M, Ibuki F, Maki Z (1987) The complete amino acid sequence of rice bran trypsin inhibitor. J Biochem 102:297–306

Touil T, Ciric B, Ventura E, Shindler KS, Gran B, Rostami A (2008) Bowman–Birk inhibitor suppresses autoimmune inflammation and neuronal loss in a mouse model of multiple sclerosis. J Neurol Sci 271:191–202

Utrilla MP, Peinado MJ, Ruiz R, Rodriguez-Nogales A, Algieri F, Rodriguez-Cabezas ME, Clemente A, Galvez J, Rubio LA (2015) Pea (*Pisum sativum* L.) seed albumin extracts show anti-inflammatory effect in the DSS model of mouse colitis. Mol Nutr Food Res 59:807–819

Wang J, Li X, Xia X, Li H, Liu J, Li QX, Li J, Xu T (2014) Extraction, purification, and characterization of a trypsin inhibitor from cowpea seeds (*Vigna unguiculata*). Prep Biochem Biotechnol 44:1–15

Ware JH, Wan XS, Rubin H, Schechter NM, Kennedy AR (1997) Soybean Bowman–Birk protease inhibitor is a highly effective inhibitor of human mast cell chymase. Arch Biochem Biophys 344:133–138