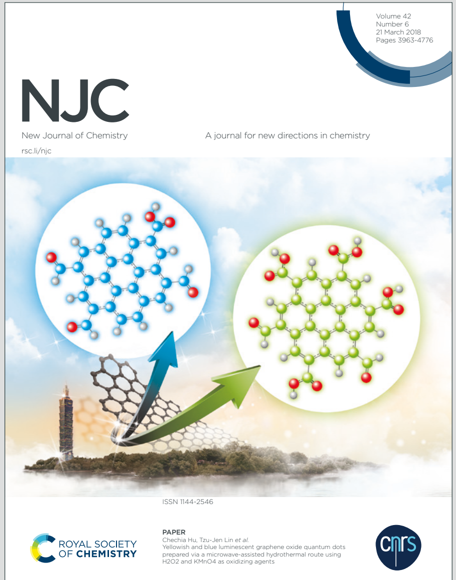
View Article Online
View Journal

New Journal of Chemistry

A journal for new directions in chemistry

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: J. Miloševi, L. Vrhovac, F. urkovi, B. Jankovi, S. N. Malkov, J. Lah and N. Polovi, *New J. Chem.*, 2020, DOI: 10.1039/D0NJ02938F.



This is an Accepted Manuscript, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this Accepted Manuscript with the edited and formatted Advance Article as soon as it is available.

You can find more information about Accepted Manuscripts in the Information for Authors.

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard Terms & Conditions and the Ethical guidelines still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this Accepted Manuscript or any consequences arising from the use of any information it contains.

rsc.li/njc

View Article Online
DOI: 10.1039/D0NJ02938F

ARTICLE

Isolation, identification, and stability of Ficin 1c isoform from fig latex

aReceived 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

Jelica Milošević [a], Lidija Vrhovac [a,b], Filip Đurković [a], Brankica Janković [a,c], Saša Malkov [d], Jurij Lah [e] and Natalija Đ. Polović *[a]

Latex of common fig (*Ficus carica*) is a rich protein source with a high level of proteolytic activity contributing to its defensive role. The divergent group of cysteine proteases known as ficin (EC 3.4.22.3) represents the majority of latex protein content and shows activity towards fig parasites. Both classical and novel biochemical techniques suggest the intricate pattern of ficin expression and activity profiles. Even though structurally related, different ficin isoforms show some differences in pI values enabling their separation using ion-exchangers. A single alkaline isoform was purified and identified based on the available transcriptomic data as Ficin 1c. This isoform shows both general proteolytic and gelatinolytic activity suggesting a biological role in the degradation of a broad range of natural substrates. The insight into the Ficin 1c structure also provided some functional clues. The secondary structure content and the overall fold are similar to related proteases of the same and other plant sources resulting in similar unfolding routes. Stability assessment of Ficin 1c in comparison to ficin isoform mixture showed that isoform diversity might lead to increased protease stability.

Introduction

Latex is a milky liquid exudate produced by lactiferous cells in about 10% of flowering plant species. A common characteristic of latex produced by different plants is the presence of many secondary metabolites. Lactiferous cells of common fig (*Ficus carica* var. Brown Turkey) produce latex rich in cis-1,4-polyisoprene, proteins, alkaloids, terpenoids, tannins, sterols [1]. All these compounds make latex the primary protector from pathogens after the injury of younger parts of the plant [2]. Biochemical protection is provided by various enzymes, including proteases, oxidases, glucosidases, phosphatases [1, 3], chitinases [4], and collagenases [5]. Dominant among fig latex proteins are cysteine proteases collectively named ficin (EC 3.4.22.3). These enzymes belong to the C1 family (clan CA) of cysteine proteases and are functionally related to papain from papaya latex, and bromelain from pineapple [6]. Plant proteases of this family have physiological roles in defense against herbivore insects and phytopathogens[7].

Protein-level investigations of fig latex show multiple forms of ficin. Most of the articles published report the existence of about five isoforms based on the number of different peaks in the ion-exchange chromatography [8-10]. Attempts to separate these isoforms date from about 50 years ago [11]. Complications with ficin separation and identification come as a result of various enzyme expression profiles in fig latex during the ripening season [12], autolysis [10], and genomic/sequential differences among varieties and cultivars [13]. The total number of ficins expressed in common fig seems to be much higher than suggested by the experiments, as indicated by recent transcriptomic data [14]. Publications regarding the proteome and transcriptome analysis of fig latex have emerged in the last few years as well. Multi-omics approach has been applied to prove the differences between commercial-ripe and tree-ripe fig [15], as well as between different organs of the young plant [14].

Considering their natural roles, plant cysteine proteases, and fig latex as a rich source of these enzymes have a long tradition of medical use [16]. Besides medical applications, proteolytic enzymes are an indispensable part of food and beverages, pharmaceutical and detergent industry, comprising about 60% of global enzyme sales [9]. Ficin, as an easily accessible protease mixture present in high concentration in fig latex, has a great potential for commercial uses. Characterization of both mixture and individual chromatographically separated isoforms, in terms of protein stability, gives new insights into their functional stability. We have already published a study on the increased stability of a mixture of ficin isoforms, as present in fig latex, compared to the singular form of papain in broader conditions [17]. This study has opened a possibility that, beside altered substrate specificity among them, the increased functional stability of multiple ficin isoforms in the mixture could be the benefit of their divergent evolution in fig latex. Here we report the first identification of purified ficin isoform based on novel transcriptome data. We further wanted to

a. Department of Biochemistry, University of Belgrade - Faculty of Chemistry, Serbia
b. Present address: Department of Chemistry and Molecular Biology, University of Gothenburg, Sweden
c. Present address: Department of Chemistry, University of Zurich, Switzerland
d. University of Belgrade - Faculty of Mathematics, Serbia
e. University of Ljubljana, Faculty of Chemistry and Chemical Technology, Slovenia
Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/x0xx00000x

Published on 02 September 2020. Downloaded by University of New England on 9/5/2020 9:34:15 PM.

New Journal of Chemistry Accepted Manuscript

Published on 02 September 2020. Downloaded by University of New England on 9/5/2020 9:34:15 PM.

investigate the differences in the general protein stability of one isolated and identified ficin isoform and the mixture of all ficin isoforms of the same latex sample. Results of ficin's activity towards natural substrates and its structural characterization are presented as well.

Experimental

Reagents

SP-Sepharose, Coomassie brilliant blue (CBB) R-250 and G-250, Bovine serum albumin (BSA), and HPLC grade acetonitrile were purchased from Sigma-Aldrich (Steinheim, Germany). Unstained protein molecular weight markers were purchased from Thermo Scientific (Rockford, IL, USA). The rest of the chemicals were analytical grade commercial products used without further purification.

Collection of fig latex and water latex fraction (WLF) preparation

Latex was collected from one fig tree (*Ficus carica* var. Brown Turkey) from the private orchard in Bar, Montenegro, in mid-June. Freshly collected latex was immediately stored at -20°C. Debris-free water latex fraction (WLF) was prepared as previously described [5]. Briefly, latex was thawed at 4°C, centrifuged to remove insoluble gum (45 min at 18000xg at 4°C), and finally, the supernatant obtained was three-times extracted with two volumes of petroleum ether. Bradford method was used for protein concentration determination with bovine serum albumin used as a standard [18].

Ion-exchange (IEX) chromatography for isoform separation

SP-Sepharose (3 mL) was packed in a 0.8 x 6 cm column and equilibrated with 30 mM citrate buffer of pH 5.0. WLF (1mL) was ten times diluted in the same buffer to prepare the starting sample and then loaded onto the column. Unbound proteins were washed with 5 column volumes (CV) - 15 mL. Bound proteins were eluted with 0-400 mM sodium chloride gradient in 30 mM citrate buffer of pH 5.0. The length of the gradient was 30 CV, and fractions volume was 0.3 mL. Protein concentration was estimated by the Bradford method to construct the chromatogram. Bound fractions under 5 major peaks were pooled.

Reverse-phase chromatography for isoform analysis

Reverse-phase chromatography was performed on a Discovery® BIO Wide Pore C5-5 10 cm x 4.6 mm column (Supelco, Bellefonte, PA, USA) using Äkta Purifier 10 System (GE Healthcare, Uppsala, Sweden). Bound peaks I-V from the IEX were prepared for the analysis by the addition of 10 mM ammonium acetate, and 1 mL of the sample was applied onto the column. Proteins were eluted in 1 mL/min flow with a 0-70% acetonitrile gradient of 10 CV length (both buffers A and B contained 25 mM ammonium acetate). Fractions of 0.1 mL were collected. Acetonitrile and ammonium acetate were removed from the fractions using SpeedVac Concentrator immediately after the elution. Proteins were resuspended in 0.1 mL of 30 mM citrate buffer of pH 5.0 and prepared for electrophoretic analysis.

View Article Online
DOI: 10.1039/D0NJ02938F

SDS PAGE

SDS PAGE was performed on Hoefer Dual Gel Mighty Small SE 245 vertical electrophoresis system (Hoefer, Holliston, MA, USA). It was applied to analyze proteins eluted in fractions of both, ion-exchange and reverse-phase chromatography. Samples were prepared under reducing conditions (in a sample buffer containing 60 mM Tris pH 6.8, 25% glycerol, 1% SDS, 14.4 mM 2-mercaptoethanol, and 0.1% bromophenol blue), and incubated 10 minutes at 90°C before applying to discontinuous electrophoresis system with 4% stacking and 12% resolving gel [19]. Proteins were stained with Coomassie Brilliant Blue (Serva, Heidelberg, Germany).

Trypsin Mass Fingerprinting for isoform identification

Reverse-phase analysis of IEX peaks I-V showed that, unlike peaks I-IV, peak V contained dominantly one ficin isoform. Fractions eluted from the reverse-phase analysis of this peak were applied to SDS PAGE. Single bands detected were cut and incubated with 0.001% trypsin (Pierce™ Trypsin Protease, MS-Grade) in 100 mM ammonium bicarbonate buffer for 24 h at 37°C. Peptide fragments were additionally extracted from the gel by incubation with 50 µL of 50% acetonitrile 0.1% trifluoroacetic acid for 30 minutes at 37°C. Acetonitrile and trifluoroacetic acid were removed using SpeedVac Concentrator.

Peptides obtained were injected onto a reversed-phase C18 column (RR HT, 1.8 µm, 4.6 × 50 mm) coupled with a Zorbax Eclipse XDB-C18 installed on 1200 series HPLC system (Agilent Technologies). The separation was performed in a segmented acetonitrile gradient (5–95% for 10 min with 95% for 5 min). All solutions contained 0.2% formic acid. The mass spectrometer (6210 Time-of-Flight (TOF) LC-MS system (G1969A; Agilent Technologies, Santa Clara, CA, USA)) was run in positive electron spray ionization (ESI) mode with a capillary voltage of 4000 V, fragmentor voltage of 200 V and mass range of 100–3200 m/z. The peptide masses were searched against the ficin sequences from transcriptome data in the NCBI database using Protein Prospector program.

General proteolytic activity

General proteolytic activity was determined towards casein as a substrate. WLF and purified ficin isoform were diluted to 0.4 mg/mL concentration in 100 mM Tris-HCl buffer of pH 8.0 containing 20 mM cysteine. Each sample in the volume of 100 µL was incubated with 500 µL of 0.65% casein dissolved in the same buffer for 30 min at 37°C. Reactions were stopped by the addition of 100 µL of 5% trichloracetic acid, followed by 15 min centrifugation step at 12000 x g to remove remaining casein. The absorbance of ten-fold diluted supernatant containing soluble peptides was measured at 280 nm. One arbitrary unit of caseinolytic activity (U) was defined as the amount of enzyme that increases the absorbance at 280 nm for 1 per minute of reaction [5]. Specific activity was defined as the

New Journal of Chemistry Accepted Manuscript

Please do not adjust margins

View Article Online
DOI: 10.1039/D0NJ02938E

Published on 02 September 2020. Downloaded by University of New England on 9/5/2020 9:34:15 PM.

number of arbitrary units per milligram of proteins in a sample. All measurements were performed in triplicates.

Gelatinolytic activity

The activity of WLF and purified ficin isoform towards gelatin was measured using the ninhydrin method [20]. Ten microliters of protein samples were incubated with 100 µL of 20 mg/mL gelatin solution for 5 h at 37°C in 100 mM Tris-HCl buffer of pH 8. After the addition of 110 µL of 20% (w/v) PEG 6000, samples were incubated for 1 h at 4°C and finally centrifuged for 30 min at 12000 x g at 4°C. Ninhydrin solution (2% ninhydrin, 50% propylene-glycol, 0.16% $SnCl_2$ in 50 mM citrate buffer of pH 5.0) in the volume of 200 µL was added to 20 µL of two-times diluted supernatant and incubated for 15 min at 100°C. After cooling to room temperature, 400 µL of 50% 1-propanol was added, the solution was two-times diluted with 100 mM Tris-HCl buffer of pH 8.0, and the absorbance was measured at 540 nm. One unit of gelatinolytic activity (GDU) was defined as the amount of enzyme that releases gelatin peptides equivalent to 1 µmol of glycine in color produced with ninhydrin in 5 h. Specific activity was defined as the number of units per milligram of proteins in a sample. All measurements were performed in triplicates.

Determination of secondary structures

Secondary structures were determined for purified ficin isoform by Fourier Transform Infrared Spectroscopy (FT-IR) with an attenuated total reflectance (ATR) at 1 cm^{-1} resolution (Nicolet 6700 FTIR, software OMNIC, Version 7.0, Thermo Scientific, USA). Ten microliters of each sample in 1 mg/mL protein concentration were applied onto diamond crystal (Smart Orbit, Thermo Scientific, USA), and the solvent was evaporated to thin-film using the nitrogen stream. Samples spectra were collected in 64 scans. Secondary structure content was analyzed in the Amide I region, essentially as described for papain [21, 22].

3D structure modeling

Ficin 1c 3D structure model was built using the Phyre 2.0 Server. The sequence of C-terminal 214 amino acids corresponding to mature protein was uploaded to the server. Twenty templates (including papain, caricain, ficin A, and ficin D) from Protein data bank were selected by the program to model Ficin 1c 3D structure based on heuristics to maximize confidence, percentage identity and alignment coverage [23].

Fluorescence measurements

Perkin Elmer LS 50 spectrofluorimeter (Perkin Elmer, Waltham, MA) was used for all fluorescence measurements and spectra recordings.

Intrinsic protein fluorescence was measured to monitor protein denaturation upon urea treatment**.** Emission spectra of ficin mixture and purified isoform in concentration 0.03 mg/mL were scanned in the presence of increasing urea concentrations (0, 0.4, 0.8, 1.2, 1.6, 2.0 M). All measurements were done in 50 mM sodium phosphate buffer at temperatures 10, 20, 30, 40, and 50°C. Samples were incubated in the presence of urea overnight at room temperature. Emission spectra were scanned from 300 to 450 nm with 295 nm excitation with excitation slit 8.5 nm, emission slit 3.3 nm, and Scan speed 60 nm/min in 1 cm cuvette. Spectral data were used to construct the urea-induced denaturation curves by plotting the maximal emission (at between 355 and 357 nm in all samples) as a function of urea concentration.

ANS fluorescence. The exposure of hydrophobic surfaces was monitored by 8-Anilinonaphthalene-1-sulfonic acid (ANS) fluorescence spectra collection. Proteins in concentration 0.03 mg/mL were incubated without urea or in the presence of 1 M urea for 1 hour at room temperature. ANS was added to the final concentration 12 µM and incubated 30 minutes in the dark at room temperature. The experiment was conducted in two different buffers – 50 mM sodium phosphate of pH 7.0 and 50 mM glycine buffer of pH 3.0. Emission spectra were collected with 380 nm excitation, with 5 nm excitation slit, 15 nm emission slit, Scan speed 60 nm/min. Emission was scanned in the range from 400 to 600 nm.

Results and discussion

Purification of ficin isoforms

The most recent transcriptomic data suggest the existence of 11 ficin isoforms (ficins 1a-c; 2a-c; 3-5, and 6a and 6b) in common fig latex of *F. carica* var. Masui Dauphine, as presented in Supplementary Table S1. On the other hand, there are four 3D structures of ficin isoforms deposited in Protein DataBank, isolated from fig latex of *F. carica* var. Chetoui, grown in north Morocco. Simple inspection suggests that their molecular masses are very similar, all between 23 and 24 kDa, which is in concordance with the experimental data [8, 24, 25]. On the other hand, they show a range of isoelectric points promising for isoform separation.

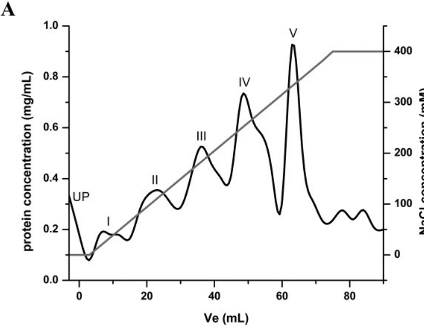
In order to separate ficin isoforms present in the latex, we performed ion-exchange chromatography. Cation exchangers have been successfully applied for the separation of ficin isoforms previously [9, 10]. Usually, the most acidic isoforms were characterized regarding the substrate specificity profile. However, due to the lack of transcriptomics derived data, none of them was identified to enable the comparison with the growing number of records from omics methodologies. Chromatography was performed at pH 5.0 to suppress autoproteolysis since cysteine and serine proteases have lowered activity at acidic pH values due to the protonation of catalytic amino acid residues in their active sites [5, 9]. As shown in Figure 1a, latex fractionation resulted in 5 major bound peaks, similar to previously published results on fractionation of ficin isoforms [8, 9]. However, the abundance of proteins under those peaks was different. In this work, the most abundant protein fractions were under peaks IV and V, while other reports suggested that more abundant were bound fractions that elute with lower ionic strength. Here we

New Journal of Chemistry Accepted Manuscript

Please do not adjust margins

View Article Online
DOI: 10.1039/D0NJ02938F

analyzed latex collected in the period when the fruits are not yet ripe, so the expression of different isoforms could differ as suggested previously [12]. Electrophoretic analysis (Figure 1b) showed that bands of all bound peaks have relative mobilities close to the 25 kDa marker, as expected for ficin isoforms (Supplementary Table S1). As visible from Figure 1a, one isoform, with an elution volume of 65 mL, under the peak, V, appeared to be best separated with the most symmetrical peak. The yield of purification of that isoform was 21.5% of total proteins from the starting sample (Table 1).



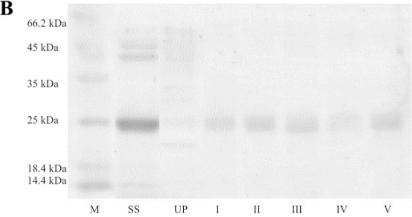


Figure 1. A) Chromatographic separation of water latex fraction proteins on SP-Sephadex column. The column was equilibrated with 30 mM sodium citrate buffer at pH 5.0. Bound proteins were eluted using a linear gradient of 0-0.4 M NaCl in the same buffer. Protein concentration in eluted peaks was estimated by the Bradford method. B) SDS PAGE analysis of eluted peaks in 12% resolving gel. M – molecular weight standards, SS – starting sample, UP – unbound proteins, I – V bound protein peaks.

Activities of purified fig latex protein fractions toward natural substrates

Plant cysteine proteases have a high level of general proteolytic activity. In the case of ficin, the general proteolytic activity could be attributed to protection against insects and fungi, as previously shown for papain [7]. The substrate used for the measurement of the general proteolytic activity was commonly used casein [5, 9, 12]. Another substrate, gelatin, being denatured collagen, has regions of specific amino acid sequence Pro-Hyp-Gly. The gelatinolytic activity could hint proteolytic potency in the hydrolysis of the connective tissue matrix [26]. We had previously characterized the enzyme with collagenolytic activity in fig latex with pI value 5 [5]. In the chromatographic analysis presented here, unbound protein fraction from the cation exchanger is likely to contain this acidic protein, as can be concluded from a high degree of gelatinolytic activity shown in Table 1. On the other hand, bound fractions show various degrees of both general proteolytic and gelatinolytic activity. These results are supported by the recent paper showing that, besides collagenases, cysteine proteases from the latex of different *F. carica* cultivars also contribute to the total gelatinolytic activity [27]. However, the ratio of gelatinolytic to general proteolytic activity is lower in bound fractions rich in mostly alkaline ficin isoforms (Table 1, Supplementary Table S1).

Table 1. Specific activities of purified fig latex proteins toward casein (general proteolysis substrate) and gelatin. UP – unbound proteins; I-V – bound protein fractions pulled under peaks I-V.

Sample	Yield, %	General proteolytic activity, U/mg	Gelatinolytic activity, GDU/mg
UP	9.5	11.9 ± 0.6	190 ± 4
I	6.7	25 ± 1	132 ± 3
II	13.3	27.3 ± 0.8	83 ± 5
III	18.8	21.5 ± 0.5	101 ± 4
IV	23.5	24.2 ± 0.6	39 ± 2
V	21.5	24.0 ± 0.4	84 ± 3

Identification of Ficin 1c

In order to test the purity of the proteolytic enzymes fractionated by cation-exchange chromatography, a reverse-phase chromatography was performed. Bound peaks I-IV showed multiple peaks in reverse-phase chromatography (data not shown), suggesting the presence of more than one isoform or their proteolytic fragments. This is in agreement with the new transcriptomics data that revealed the existence of more isoforms than the chromatography profile of ficins suggested. Here we focused on the most alkaline isoform, which has not been characterized before (peak V). The reverse-phase chromatogram and electrophoretic profiles of IEX bound peak V are shown in Figure 2, suggesting the dominant presence of one isoform with two smaller impurities of slightly higher and lower elution volumes. The peaks were subjected to trypsin mass fingerprinting to identify which ficin isoform they represent.

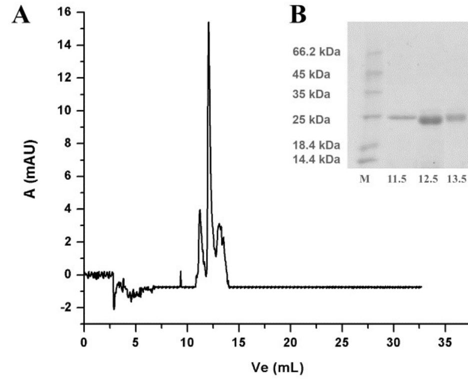


Figure 2. A) Reverse-phase chromatography of ficin isoform under the peak V from IEX chromatography. The analysis was performed on Discovery® BIO Wide Pore C5-5 10 cm x 4.6 mm column using a 10 CV long gradient of acetonitrile 0-70% in 25 mM ammonium acetate. The elution was monitored by measuring the absorbance at 280 nm. B) SDS PAGE analysis of fractions under the peaks in 12% resolving gel. M – molecular weight standards, 11.5, 12.5, 13.5 – elution volumes of samples (mL) under the peaks.

Published on 02 September 2020. Downloaded by University of New England on 9/5/2020 9:34:15 PM.

New Journal of Chemistry Accepted Manuscript

This journal is © The Royal Society of Chemistry 20xx

View Article Online
DOI: 10.1039/D0NJ02938F

Table 2. Peptide mass fingerprint of the isolated ficin isoform. Sequences given in the bold correspond to the Ficin 1c specific peptides.

m/z Submitted	MH^+ Matched	Delta Da	Start	End	Missed Cleavages	Sequence
679.1800	679.3232	-0.143	67	72	0	(R)GGWMTK(A)
750.0800	750.3451	-0.265	96	102	0	(K)**GECNVTK**(A)
750.0800	750.3781	-0.298	113	118	0	(R)**YESVPR**(N)
785.3100	785.4152	-0.105	198	205	0	(R)NVAEPAGK(C)
878.8000	878.4400	0.360	95	102	1	(R)**KGECNVTK**(A)
885.2000	885.4716	-0.272	73	79	0	(K)AYDYIIK(N)
1015.3100	1015.5207	-0.211	1	8	0	(-)LPETVDWR(I)
1061.4300	1061.5586	-0.129	103	112	0	(K)ASQTVATIDR(Y)
1105.6600	1105.4810	0.179	182	190	0	(R)**NSWGNWGER**(G)
1178.3000	1178.5333	-0.233	206	216	0	(K)**CGIAMHSTYPV**(-)
1568.7200	1568.8391	-0.119	127	142	0	(K)**AVANQPVSVTIEAGGR**(A)
1628.7500	1628.7663	-0.0163	80	94	0	(K)NGGITSQSNYPYTAR(K)
1696.5900	1696.9341	-0.344	126	142	1	(K)**KAVANQPVSVTIEAGGR**(A)
1757.1600	1756.8613	0.299	80	95	1	(K)NGGITSQSNYPYTARK(G)

Protein Prospector analysis showed the highest overlap between purified ficin isoform and Ficin 1c from transcriptomic analysis (48.6% coverage or 105 out of 216 amino acid residues) with 7 specific out of 14 matched peptides. Sequences of other ficin isoforms: Ficin 1a and Ficin 3 were non-specifically recognized with significantly lower coverage of about 14% (Supplementary Table S2) as expected due to the sequence identity and close evolutionary relationships between them and Ficin 1c (Supplementary Figures S1 and S2) and could be a consequence of the smaller peak present in the sample. In respect to trypsin mass fingerprinting C-terminally truncated Ficin 1c is present as an impurity, as well. However, the folded Ficin 1c protein content is estimated to be about 85-90% under the peak V. Proteomic identification of Ficin 1c matched with its ion-exchange behavior as well. Being a highly alkaline protein (with the theoretical pI value of 8.64, Supplementary Table 1), Ficin 1c is eluted from the cationic exchanger under the peak V using concentrations of sodium-chloride higher than 0.3 M (Figure 1). Sequences of ficin isoforms isolated from another source (Ficins B, C, and D) were matched with 12-20% coverage (Supplementary Table S2).

Secondary structure of Ficin 1c

To further characterize the identified isoform in the structural sense, secondary structure content was determined in purified Ficin 1c by Fourier-transform infrared spectroscopy (FTIR). The deconvoluted spectrum is presented in Supplementary Figure S3. Ficin 1c has 25.6% of β-sheets and 23.4% of α-helices, which does not differ significantly from the secondary structure content of ficins from *F. carica* var. Chetoui with the known tertiary structure. The entire secondary structure content of Ficin 1c is also similar to related cysteine proteases from other sources - papain and caricain, as presented in Table 3.

Table 3. Secondary structure contents of ficin and related cysteine proteases.

Protein	Secondary structure content, %				
	β-strand	α-helix	Turns	Unordered	Error/ reference
Ficin 1c	25.6	23.4	18.1	32.9	0.8
Ficin A	22-26	19-24	18-19	35-40	Reference[8]
Ficin B-D	22-24	19-23	18	36-40	Reference[8]
Papain	20	29	18	38	Reference[8]
Caricain	20	26	15	42	Reference[8]

3D structure model

3D structure model of mature Ficin 1c was built using Phyre 2.0 Server. All the templates covered 99% of the Ficin 1c sequence and had a high local similarity. Additional information on the template library is presented in Supplementary Table S3. Ficin 1c structure is modeled with dimensions (Å): X:52.538 Y:60.697 Z:42.676, and is presented in Figure 3, together with papain and ficins A-D.

Published on 02 September 2020. Downloaded by University of New England on 9/5/2020 9:34:15 PM.

New Journal of Chemistry Accepted Manuscript

This journal is © The Royal Society of Chemistry 20xx

View Article Online
DOI: 10.1039/D0NJ02938F

Published on 02 September 2020. Downloaded by University of New England on 9/5/2020 9:34:15 PM.

New Journal of Chemistry Accepted Manuscript

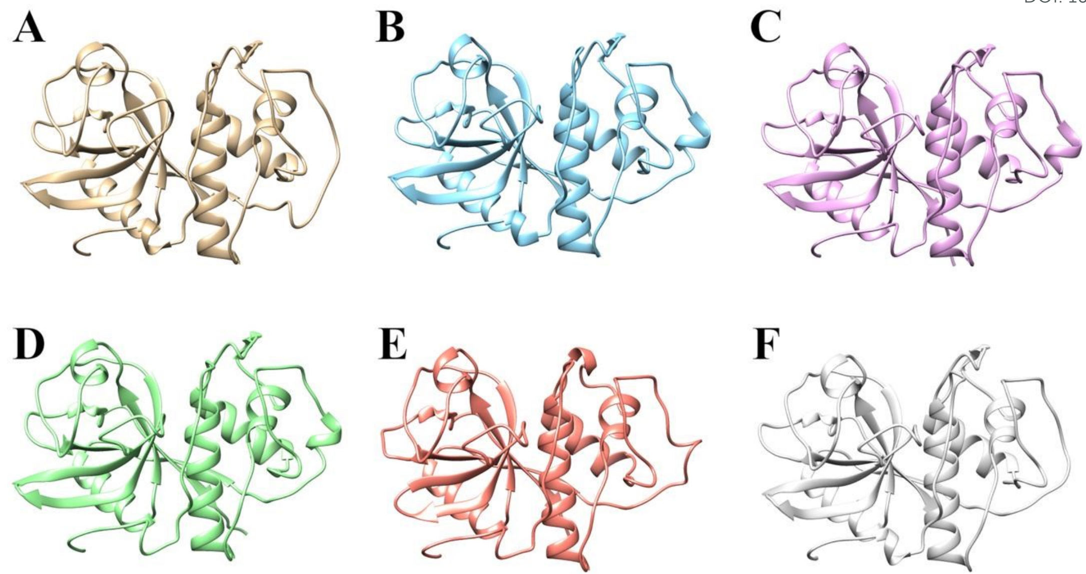


Figure 3. 3D structure comparison of A) Ficin 1c (model); B) Ficin D (PDB 4YYW); C) Ficin C (PDB 4YYU); D) Ficin B (PDB 4YYS); E) Papain (PDB 1KHQ) and F) Ficin A (PDB 4YYQ).

As 99% of residues were modeled at 100% confidence, we assume that the core of the protein had been modeled reliably. Ficin 1c showed the overall fold typical for papain-like cysteine proteases, having one α-helix-rich and one β-sheet-rich domain. The main but still minor difference between the modeled structure of Ficin 1c and Ficins B, C, and D was the lack of a short helix in the middle of the unordered region within the α-helix-rich domain. The same helix is missing in the structures of Ficin A and papain, although, despite model confidence, there is the possibility that the orientation of the surface loops is not that well-defined.

Ficin 1c stability assessment

Ficin 1c denaturation was monitored by the measurement of its intrinsic fluorescence in the presence of chaotropic agent – urea, in the temperature range 10-50°C. Urea was used in concentrations lower than 2 M leading to a stable intermediate state (Figure 4). Further increase of urea concentration up to 8 M results in a complete denaturation of the protein, making it a two-step process. We focused on the first transition to a stable intermediate. The addition of urea to 0.4 M concentration led to the transition towards this intermediate state observed by the increase in emission (Figure 4). It is likely to be the consequence of structural perturbations in the α-helix-rich domain known to be less stable and prone to reorganization in structurally related proteins [22].

On the other hand, a more stable β-sheet-rich domain was not affected in this range of urea concentrations (0.4 to 2 M), resulting in a stable intermediate structure. Similar behavior was reported for papain upon urea treatment [28, 29], and in low concentrations of another chaotropic agent - guanidine hydrochloride [30, 31]. The same pattern of Ficin 1c and papain behavior in the presence of urea can be explained by similarities of their overall fold, domain distribution, the intactness of hydrophobic core of the proteins (Figure 3), conservation of cysteine residues (forming disulfide bridges) (Supplementary Figure S1), secondary structures content (Table 3) and distribution (Figure 3).

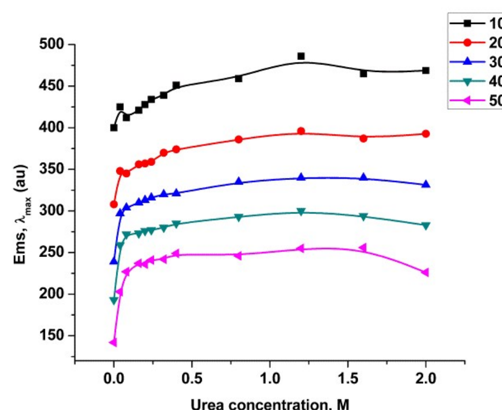


Figure 4. Urea denaturation of Ficin 1c in the temperature range 10-50°C.

Ficin 1c stability was further characterized in neutral and acidic conditions of pH 3. Emission spectra of Ficin 1c were recorded at both pH 7 and 3 without urea and in the presence of 1 M urea (Figure 5). The results on destabilization of Ficin 1c are concordant with the article published by Devaraj *et al.* indicating that another, yet unidentified, ficin isoform is more susceptible to urea unfolding at low pH of about 3 [32]. The concentration of urea of 1 M was in the range of concentrations leading to the first transition to the stable intermediate (Figure 4). Its formation by pH shift in Ficin 1c is proven by both methods - intrinsic fluorescence and ANS fluorescence measurement (Figure 5). Regardless pH, the

This journal is © The Royal Society of Chemistry 20xx

View Article Online
DOI: 10.1039/D0NJ02938F

presence of 1 M urea increases the exposure of hydrophobic patches resulting in increased ANS fluorescence. Still, in the case of pH 3, the molten globule intermediate state is favored both in Ficin 1c and previously investigated ficin isoforms [24], resulting in higher hydrophobic exposure even without urea added. The previously described ficin isoform [9] was eluted from cationic exchanger at low salt concentrations, and according to its amino acid composition is acidic protein, contrary to Ficin 1c. Nevertheless, both ficin isoforms showed similar behavior concerning urea denaturation and the existence of a folding intermediate at acidic pH value. Altogether, Ficin 1c showed the same stability concerning urea and low pH destabilization as another purified ficin isoform and papain.

Ficin 1c increases considerably at the lowest concentration of urea applied (0.4 M), and all the spectra with varying urea concentrations stuck close to each other, in the case of ficin isoform mixture, perturbations leading to the increase in emission are more gradual. The variety of isoforms in a mixture with different transition points can explain this difference. As some of the isoforms are probably more stable than others, the transition to intermediate state is more gradual, suggesting higher stability in lower concentrations of urea in the case of isoform mixture compared to Ficin 1c alone.

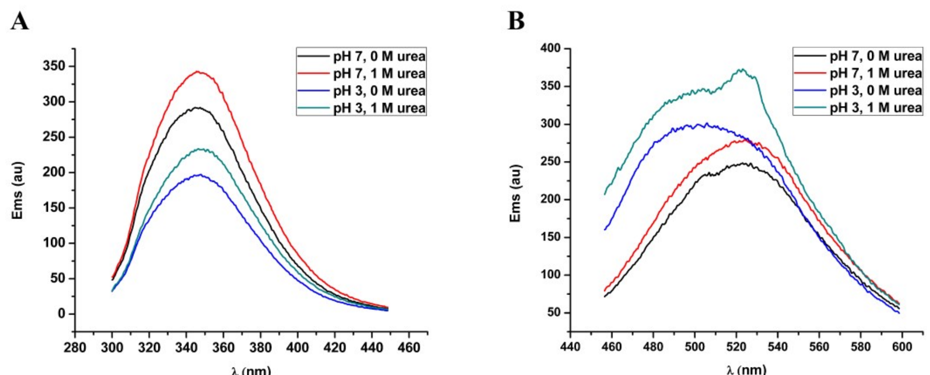


Figure 5. A) Emission spectra of Ficin 1c at pH 7 and pH 3 with (1 M) and without urea. B) ANS fluorescence spectra of Ficin 1c at pH 3 and 7 with (1 M) and without urea

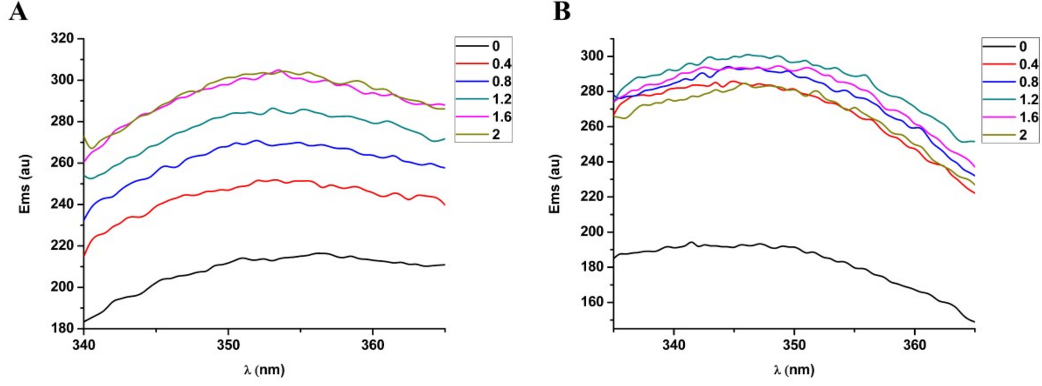


Figure 6. Emission fluorescence spectra of A) ficin mixture and B) Ficin 1c recorded at 40°C (excitation wavelength 295 nm) in the presence of 0 - 2 M urea (concentrations provided in the legend).

Comparison of Ficin 1c and ficin isoform mixture unfolding

The differences in the unfolding of purified Ficin 1c and a mixture of ficin isoforms in the same latex of not fully ripe fig are noticeable by comparing their emission spectra in the presence of increasing urea concentration (Figure 6). The same range of urea concentrations (0 - 2 M) was applied, resulting in a stable intermediate rather than a fully denatured state. The distribution of spectra of Ficin 1c and ficin isoform mixture differ significantly (Figure 6). While emission in the case of

A similar pattern of prolonged cold stability in the presence of ethanol of ficin isoforms mixture compared to single polypeptide papain was noticed in our previous work [20]. Due to the absence of aggregation-specific β-sheets, the mechanism that stabilized a mixture of ficin isoforms is probably a consequence of the lowered ability of polypeptides with different sequences to form nuclei for aggregation and consequent cold denaturation. Despite dealing with denaturation caused by a chaotropic agent, the same mechanism can play a role here, since low concentrations of

Published on 02 September 2020. Downloaded by University of New England on 9/5/2020 9:34:15 PM.
New Journal of Chemistry Accepted Manuscript

View Article Online
DOI: 10.1039/D0NJ02938F
Published on 02 September 2020. Downloaded by University of New England on 9/5/2020 9:34:15 PM.
New Journal of Chemistry Accepted Manuscript

urea (<1 M) does not prevent polypeptide aggregation in papain-like proteases [28]. The gradual unfolding of ficin isoform mixture compared to easily destabilized Ficin 1c (Figure 6) supports the hypothesis that the diversity of sequences provides stabilization. As all isoforms contribute to latex properties, the expression of multiple isoforms in the unripe fig might be beneficial for their broad substrate specificity and higher stability, providing better protection in the ripening period when the plant is susceptible to parasites and herbivore attacks.

Conclusions

We isolated alkaline isoform of ficin from *F. carica* and identified this isoform as Ficin 1c based on the available transcriptome data. Purified isoform is specific towards casein representing its general proteolytic activity and towards gelatin, suggesting its broad role in natural substrate proteolysis. Secondary structure content and the overall fold of purified Ficin 1c is in concordance with the data regarding other characterized ficin isoforms, as well as the entire family of papain-like plant cysteine proteases. Its structural features result in stable unfolding intermediate formation in low urea and low pH conditions. This single isoform appeared to be less resistant to denaturation than the mixture of ficin isoforms, suggesting that, besides broader substrate specificity, prolonged stability of ficin isoforms mixture could be another beneficial consequence of divergent evolution of latex proteases.

Conflicts of interest

There are no conflicts to declare.

Acknowledgments

This work was supported by the Ministry of Education, Science and Technological Development of Republic of Serbia, Contract number: 451-03-68/2020-14/200168

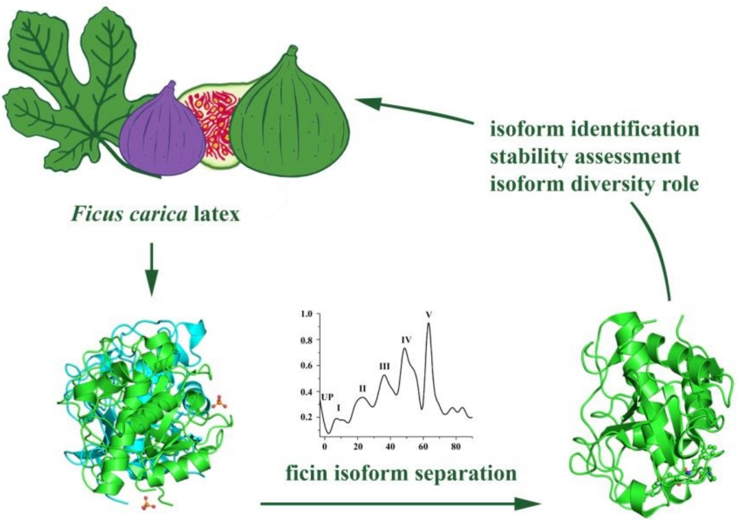
References

1 C. Villard, R. Larbat, R. Munakata and A. Hehn, *Planta*, 2019, **249**, 617-633.
2 2. A. P. Oliveira, L. R. Silva, P. B. Andrade, P. Valentao, B. M. Silva, R. F. Goncalves, J. A. Pereira and P. Guedes de Pinho, *J Agric Food Chem*, 2010, **58**, 10855-10863.
3 K. Konno, *Phytochemistry*, 2011, **72**, 1510-1530.
4 C. L. Chiang, Y. M. Chang, C. T. Chang and H. Y. Sung, *J Agric Food Chem*, 2005, **53**, 7579-7585.
5 B. Raskovic, O. Bozovic, R. Prodanovic, V. Niketic and N. Polovic, *J Biosci Bioeng*, 2014, **118**, 622-627.
6 N. Gonzalez-Rabade, J. A. Badillo-Corona, J. S. Aranda-Barradas and C. Oliver-Salvador Mdel, *Biotechnol Adv*, 2011, **29**, 983-996.
7 K. Konno, C. Hirayama, M. Nakamura, K. Tateishi, Y. Tamura, M. Hattori and K. Kohno, *Plant J*, 2004, **37**, 370-378.
8 M. Azarkan, A. Matagne, R. Wattiez, L. Bolle, J. Vandenameele and D. Baeyens-Volant, *Phytochemistry*, 2011, **72**, 1718-1731.
9 K. B. Devaraj, P. R. Kumar and V. Prakash, *J Agric Food Chem*, 2008, **56**, 11417-11423.
10 H. Zare, A. A. Moosavi-Movahedi, M. Salami, M. Mirzaei, A. A. Saboury and N. Sheibani, *Phytochemistry*, 2013, **87**, 16-22.
11 V. C. Sgarbieri, S. M. Gupte, D. E. Kramer and J. R. Whitaker, *J Biol Chem*, 1964, **239**, 2170-2177.
12 B. Raskovic, J. Lazic and N. Polovic, *J Sci Food Agric*, 2016, **96**, 576-582.
13 L. Solorzano Zambrano, G. Usai, A. Vangelisti, F. Mascagni, T. Giordani, R. Bernardi, A. Cavallini, R. Gucci, G. Caruso, C. D'Onofrio, M. F. Quartacci, P. Picciarelli, B. Conti, A. Lucchi and L. Natali, *Genom Data*, 2017, **13**, 64-66.
14 S. Kitajima, W. Aoki, D. Shibata, D. Nakajima, N. Sakurai, K. Yazaki, R. Munakata, T. Taira, M. Kobayashi, S. Aburaya, E. H. Savadogo, S. Hibino and H. Yano, *Planta*, 2018, **247**, 1423-1438.
15 Y. Cui, Z. Wang, S. Chen, A. Vainstein and H. Ma, *BMC Plant Biol*, 2019, **19**, 146.
16 E. P. Lansky, H. M. Paavilainen, A. D. Pawlus and R. A. Newman, *J Ethnopharmacol*, 2008, **119**, 195-213.
17 J. Milosevic, B. Jankovic, R. Prodanovic and N. Polovic, *Amino Acids*, 2019, **51**, 829-838.
18 M. M. Bradford, *Anal Biochem*, 1976, **72**, 248-254.
19 U. K. Laemmli, *Nature*, 1970, **227**, 680-685.
20 Y. Zhang, Y. Fu, S. Zhou, L. Kang and C. Li, *Anal Biochem*, 2013, **437**, 46-48.
21 B. Raskovic, M. Popovic, S. Ostojic, B. Andelkovic, V. Tesevic and N. Polovic, *Spectrochim Acta A*, 2015, **150**, 238-246.
22 B. Raskovic, N. Babic, J. Korac and N. Polovic, *J Serb Chem Soc*, 2015, **80**, 613-625.
23 L. A. Kelley, S. Mezulis, C. M. Yates, M. N. Wass and M. J. Sternberg, *Nat Protoc*, 2015, **10**, 845-858.
24 K. B. Devaraj, P. R. Kumar and V. Prakash, *Int J Biol Macromol*, 2009, **45**, 248-254.
25 S. Haesaerts, J. A. Rodriguez Buitrago, R. Loris, D. Baeyens-Volant and M. Azarkan, *Acta Crystallogr F Struct Biol Commun*, 2015, **71**, 459-465.
26 B. G. Raskovic and N. D. Polovic, *J Herb Med*, 2016, **6**, 73-78.
27 K. Nishimura, K. Higashiya, N. Ueshima, T. Abe and K. Yasukawa, *J Food Sci*, 2020, **85**, 535-544.
28 F. Edwin, Y. V. Sharma and M. V. Jagannadham, *Biochem Biophys Res Commun*, 2002, **290**, 1441-1446.
29 H. A. Sathish, P. R. Kumar and V. Prakash, *Indian J Biochem Biophys*, 2002, **39**, 155-162.
30 V. Mohan, N. Das, A. Das, V. Mishra and P. Sen, *J Phys Chem B*, 2019, **123**, 2280-2290.
31 B. Sengupta, A. Chaudhury, N. Das and P. Sen, *Protein Pept Lett*, 2017, **24**, 1073-1081.
32 K. P. Devaraj KB, Prakash V, *Process Biochem*, 2011, **46**, 458–464.

This journal is © The Royal Society of Chemistry 20xx

View Article Online
DOI: 10.1039/D0NJ02938F

Published on 03 September 2020. Downloaded by University of New England on 9/5/2020 9:34:15 PM.



Purified alkaline ficin isoform, identified as Ficin 1c regarding fig transcriptome, shows decreased stability compared to the ficin isoform mixture.

New Journal of Chemistry Accepted Manuscript