Amino acid sequence of two new milk-clotting proteases from the macroalga Gracilaria edulis

by Ariestya Arlene Arbita

Submission date: 25-Mar-2023 10:19AM (UTC+0700)
Submission ID: 2045975376
File name: B1C1-03._Amino_acid_sequence_of_two_new.pdf (496.34K)
Word count: 7993
Character count: 49759

International Journal of Biological Macromolecules 211 (2022) 499–505

Contents lists available at ScienceDirect

International Journal of Biological Macromolecules

journal homepage: www.elsevier.com/locate/ijbiomac

Amino acid sequence of two new milk-clotting proteases from the macroalga *Gracilaria edulis*

Ariestya Arlene Arbita [a,c], Nicholas A. Paul [b], Julian Cox [a], Jian Zhao [a,*]

[a] School of Chemical Engineering, The University of New South Wales, Sydney, NSW 2052, Australia
[b] School of Science and Engineering, University of the Sunshine Coast, Maroochydore, Queensland 4558, Australia
[c] School of Chemical Engineering, Faculty of Industrial and Technology, Parahyangan Catholic University, Ciumbuleuit 94, Bandung 40141, Indonesia

ARTICLE INFO

Keywords:
Gracilaria edulis
Algae
Proteases
Proteomics
Protein sequencing

ABSTRACT

This study is aimed at identifying and characterising the proteases we previously extracted from the red seaweed *Gracilaria edulis* with the potential as milk-clotting enzymes. The protease extract was first analysed by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) and zymography. Two protease bands with a molecular weight of 44 and 108 kDa were identified, and analysed using in-gel digestion and liquid chromatography-tandem mass spectrometry/mass spectrometry (LC-MS/MS). Eight peptides from the LC-MS/MS analysis matched those in existing protein databases but they were not related to any protease of the genera *Gracilaria* and *Hydropuntia*. Further analysis revealed that more than 80% of the peptide sequence of the algal proteases matched with those from members of the bacteria kingdom, including *Gallaecimonas* and *Alteromonas*. Among these, twelve matching homolog proteases were identified as metalloprotease and serine protease. The results indicated that the algal proteases have a close relationship with both algae and bacteria, and suggest that the proteases might have resulted from past bacterial colonisation of the algae and subsequent horizontal gene transfer between bacteria and algae.

1. Introduction

Milk-clotting enzymes are essential in cheese-making. The earliest and most common coagulant used in cheese-making is calf rennet, derived from the abomasum of a young cow; however, it is now unable to meet the growing demand for cheese production worldwide [1]. The primary component of calf rennet is chymosin, with a lesser amount of pepsin, but their actual proportions vary depending on the calf's age [2]. Several alternatives have been explored as substitutes for calf rennet, including animal pepsins, recombinant chymosin, microbial, fungal and plant coagulants. Although these substitutes have been used in cheese making and produced cheese with comparable properties with those of calf rennet cheese, they still have various drawbacks, including issues with genetically engineered organisms (GMO), ethical and religious concerns, and excessive proteolytic activity which leads to bitterness in the cheese [3,4].

The marine ecosystem harbours a huge and diverse range of different organisms; many of which contain proteases that have the potential as milk coagulants, including the vast number of macroalgal species [5,6]. In a previous study [7], we isolated two proteases from the red seaweed *Gracilaria edulis*. The proteases exhibited good caseinolytic and milk-clotting activities and demonstrated a potential as a rennet substitute for cheese making. The proteases were found to belong to the serine and metalloprotease groups with a molecular weight of 44 and 108 kDa.

G. edulis is a marine red macroalga which is widely utilised for food and for producing agar across the tropical Indo-Pacific [8]. The seaweed is rich in nutrients, amino acids phytohormones and the polysaccharides are extracted as agar [9]. Previous studies on *G. edulis* are focused on nutritional composition and physicochemical properties [10,11]. Some of them explored the functional properties of the alga, such as alpha-amylase and alpha-glucosidase inhibition [12], and bioactive peptides [13,14]. Meanwhile, a few studies have also investigated the enzymatic activities, such as the bromoperoxidatic activity, of the species [15]. Recent studies on *G. edulis* have shifted to genomics. Liu et al. [16] reported the complete plastid genome of *G. edulis* and showed its close relationship with the other *Gracilaria* species. The NCBI sequence database of *G. edulis* and its synonym (*Hydropuntia edulis*) was 751 and 40, respectively (March 2022). These proteins are derived from the mitochondrion, plastid and chloroplast. However, among these numbers, none of them was protease, and the study about enzymes,

* Corresponding author.
E-mail address: jian.zhao@unsw.edu.au (J. Zhao).

https://doi.org/10.1016/j.ijbiomac.2022.05.038
Received 16 September 2021; Received in revised form 8 April 2022; Accepted 5 May 2022
Available online 11 May 2022

especially protease, from *G. edulis* species was very limited.

One of the main challenges of extracting proteolytic enzymes from biological matrixes is the maintenance of their activity, which can be lost during the extraction process [17]. For plant tissues such as seaweed, cell wall and the vacuoles are the two main obstacles in extracting enzymes. These need to be disrupted, usually by homogenisation. As protein solubility increases at low salt concentrations (<0.5 M), known as "salting in", a buffer of low salt concentration is commonly used as a solvent for homogenisation to allow the enzyme to dissolve in an aqueous phase [18]. For most enzymatic extracts, purification is necessary to remove non-enzymatic components to increase the unit activity. The most commonly used purification method is the "salting out" process with high salt concentrations which decrease the solubility of protein and cause the protein to precipitate based on their solubility [18]. This is usually followed by a desalting process which can be achieved by several methods, including dialysis, desalting column, ultrafiltration, three-phase partitioning (TPP) and phenol-based method [18–21]. Further purification of proteases usually involves the use of various chromatographic techniques such as gel filtration chromatography, affinity chromatography, ion-exchange chromatography and fast protein liquid chromatography (FPLC) [18].

With regard to the study on *G. edulis* proteases, the molecular properties of the enzymes, such as amino acid sequence, remain to be characterised. Such information is crucial to a deeper understanding of the biochemical properties of the enzymes and their biological relationship with other proteases [22]. Therefore, in the present study, we purified the proteases from *G. edulis* and analysed the amino acid sequence of the enzymes by in-gel digestion and liquid chromatography-tandem mass spectrometry/mass spectrometry.

2. Materials and methods

2.1. Materials and chemicals

Gracilaria edulis (synonym *Hydropuntia edulis*) was purchased fresh from Suva market, Fiji. The alga was dehydrated in an Ezidri-Ultra FD1000 food dehydrator at 60 °C for 3 days. The dried alga had a moisture content 8.8 ± 0.5%, and was kept in a sealed vacuum bag with a silica gel pack inside and stored at ambient temperature until it was processed. Bovine whole casein and individual caseins (α-, β-, and κ-casein) were purchased from Sigma-Aldrich (Sydney, Australia). All other chemicals were of analytical grade and were purchased from Sigma-Aldrich (Sydney, Australia) unless otherwise stated.

2.2. Extraction and protease purification from G. edulis

The *G. edulis* proteases were extracted and purified according to the method described in literature [23] and the detailed procedures are given in our previous study [7]. Briefly, dried alga samples (6 g) was homogenised in 20 mM phosphate buffer (pH 7.0) in a ratio of 1:20 (w/v) with an Ultra-Turrax T 25 homogeniser (IKA, Staufen, Germany) for 1 h at 24 °C, followed by centrifugation at 5000 ×*g* for 30 min at 4 °C. The supernatant was filtered with Whatman #1 filter paper and collected as the liquid crude extract (CE). The CE was mixed with ammonium sulphate at 50% (w/v) saturation and stirred for 30 min. The mixture was centrifuged at 10,000 ×*g* for 10 min at 4 °C and the pellet was collected and re-dissolved in one-tenth of the initial CE volume of 20 mM phosphate buffer (pH 7.0), which was then dialysed (3.5 kDa cut-off membrane) against 1.8 l of the same phosphate buffer at 4 °C for 24 h with constant stirring. The resultant solution was taken as purified extract (PE).

2.3. Identification of active protein bands of purified extracts with SDS-PAGE and zymography

The purified extract (PE) was first subjected to analysis by SDS-PAGE and zymography to separate and identify the active protein bands according to the procedure of the previous study [7]. Briefly, the PE (50 μl) was mixed with denaturing and reducing sample buffer (50 μl) followed by boiling at 95 °C for 5 min. The sample buffer was a solution of 2× Laemmli concentrate:β-mercaptoethanol (95: 5, v/v). The prepared samples (50 μl) and the pre-stained protein standards of 2–250 kDa (Precision Plus Protein Dual Xtra Prestained Protein Standards, Bio-Rad) were loaded into each well of 4–20% (w/v) sodium dodecyl sulphate polyacrylamide gel and electrophoresis was run at 400 mA, 200 V until the bromophenol blue dye reached the bottom of the gel. After electrophoresis, the gel was soaked in a fixing solution of acetic acid:methanol:water (7:10:83, v/v) followed by staining in 0.0275% (w/v) of Coomassie blue stain R-250 dissolved in a mixture of acetic acid:methanol:water (1:4:95, v/v). The gel was de-stained by repeatedly washing in the solution of acetic acid:methanol:water (1:4:5, v/v).

The caseinolytic activity of the PE were analysed on a 10% polyacrylamide gel (TGXTM FastCast™ Acrylamide Kit, 10%, Bio-Rad) copolymerised with 0.1% casein using Native-PAGE zymography. The extract was mixed with sample buffer at a 1:1 ratio (v/v) without denaturing and reducing conditions. Each prepared sample (30 μl) was transferred to each well of the gel and electrophoresis was conducted at 400 mA and 200 V in the running buffer (2.5 mM Tris, 19.2 mM glycine, 0.01% SDS, pH 8.3, Bio-Rad) until the bromophenol blue dye reached the bottom of the gel. After electrophoresis, the gel was washed for 15 min in the enzyme renaturing buffer (2.5% (v/v) Triton X-100, 50 mM Tris, 200 mM NaCl, 0.5 mM $ZnCl_2$, 5 mM $CaCl_2$, 0,02% (w/v) NaN_3) at room temperature with gentle agitation and the washing step was repeated four times using fresh buffer each time. The gel was transferred to 30 ml of developing buffer (50 mM Tris, 200 mM NaCl, 0.5 mM $ZnCl_2$, 5 μM $CaCl_2$, 0,02% (w/v) NaN_3) and was incubated for 2 h. The gel was then stained with 0.125% of Coomassie brilliant blue R-250 dissolved in a solution of acetic acid:methanol:water (2:5:3, v/v) and de-stained with a solution of formic acid:methanol:water (1:30:70, v/v). Protein bands with caseinolytic activity appeared as clear bands on a dark blue background of the gel. After that, the protein bands that exhibited caseinolytic activity were excised from the SDS-PAGE gel and subjected to in-gel digestion for sequence analysis by LC-MS/MS.

2.4. Protease sequencing

2.4.1. In-gel digestion

Gels that contained protein bands of interest were firstly washed with deionised water for 5 min. The bands of interest were excised using a sterile scalpel blade with excision being made as close as possible to the boundary of the band. Coomassie stain was removed by incubating the gel bands in 200 μl NH_4HCO_3 (25 mM) in acetonitrile for approximately 20 min. After removal of the solvent, the bands were incubated in 40 μl iodoacetamide (25 mM) in NH_4HCO_3 (50 mM) for 30 min at 37 °C. The bands were then washed with acetonitrile (50 μl) twice for 10 min each, followed by the addition of 40 μl trypsin (~100 ng) in NH_4HCO_3 (20 mM), and the solution incubated at 37 °C for 14 h. The bands were then washed with deionised water (50 μl), followed by incubation in 1% v/v formic acid and acetonitrile (100 μl) for 15 min. The supernatant containing the digested peptides was dried and rehydrated in deionised water (10 μl) containing 0.05% v/v heptafluorobutyric acid and 0.1% v/v formic acid.

2.4.2. LC-MS/MS analysis of amino acid sequence

The digested peptides were separated by nano-LC using an Ultimate 3000 HPLC and autosampler system (Dionex, Amsterdam, Netherlands). Samples (2.5 μl) were concentrated and desalted onto a micro C18 precolumn (300 μm × 5 mm, Dionex) with H_2O:CH_3CN (98:2, 0.05% TFA) at 15 μl/min. After a 4 min wash, the pre-column was switched (Valco 10 port valve, Dionex) into line with a fritless nano column (75 μ × ~10 cm) containing C18 media (1.9 μ, 120 Å, Dr Maisch, Ammerbuch-Entringen Germany). Peptides were eluted using a linear gradient of

H_2O:CH_3CN (98:2, 0.1% formic acid) to H_2O:CH_3CN (64:36, 0.1% formic acid) at 200 nl/min over 30 min. High voltage (2000 V) was applied to low volume tee (Upchurch Scientific, Washington, USA) and the column tip positioned ~0.5 cm from the heated capillary (T = 275 °C) of an Orbitrap Velos (Thermo Electron, Bremen, Germany) mass spectrometer. Positive ions were generated by electrospray and the Orbitrap operated in data dependent acquisition (DDA) mode. A survey scan *m/z* 350–1750 was acquired in the Orbitrap (resolution = 30,000 at m/z 400, with an accumulation target value of 1,000,000 ions) with lock-mass enabled. Up to the 10 most abundant ions (>4000 counts) with charge states > +2 were sequentially isolated and fragmented within the linear ion trap using collisionally induced dissociation with an activation q = 0.25 and activation time of 30 ms at a target value of 30,000 ions. The *m/z* ratios selected for MS/MS were dynamically excluded for 30 s.

All MS/MS spectra were searched against NCBI database (December 2018) using MASCOT (version 2.3, Matrix Science) with the following search criteria: enzyme specificity was trypsin; precursor and product ion tolerances were at 4 ppm and ± 0.4 Da, respectively; variable modification of methionine oxidation; and one missed cleavage was allowed. The ion score significance threshold was set to 0.5 and each protein was provided with a probability based Mowse (Molecular Weight Search) score [24]. The mass spectrometric analysis was carried out at the Bioanalytical Mass Spectrometry Facility, University of New South Wales, Australia.

3. Results and discussion

The results of SDS-PAGE and zymography of *G. edulis* protease are shown in Fig. S1. The SDS-PAGE showed several protein bands while the zymogram confirmed two bands at 44 kDa and 108 kDa, which exhibited dominant caseinolytic activity. These two bands were excised and subjected to amino acid sequencing by LC-MS/MS analysis. The LC-MS peptide chromatograms of the two bands were given in Figs. S2 and S3. Each peptide peak was separated and subjected to MS and the mass data collected from the LC/MS/MS analysis from both excised gels were searched on the NCBI database using Eukaryotes and Other Eukaryotes taxonomy (Tables 1 and 2). The identification of the amino acid sequence was based on similarities between the protein from the excised gels and known proteases from the NCBI protein database.

Eight peptides were found in the database with the Mascot Protein Score restricted to the related genera *Gracilaria* and *Hydropuntia* ranging from 36 to 148. The peptides included the photosystem I P700 chlorophyll *a* apoprotein A1 and A2 which are the proteins from the chloroplast *Gracilaria tenuistipitata* var. *liui* and *Gracilaria salicornia*, respectively [25,26]. The GDP-mannose-3′, 5′-epimerase which is an enzyme involved in the biosynthesis of agar and the cell wall polysaccharides [27] while the galactose-1-phosphate uridylytransferase is an enzyme that breaks complex sugars into galactose [28]. However, none of the peptides was identified as part of a protease.

The accession number of proteins in the NCBI database (March 2022) that relates to the species *Gracilaria edulis* and *Hydropuntia edulis* was 751 and 40, respectively. However, none of them is listed as protease. This suggested that the protein sequence of hydrolytic enzymes for those species might not have been identified previously and, hence, not stored in the database. Therefore, we tried to compare the LC-MS/MS results with other taxonomies available in the databases. The LC-MS/MS results were searched against different taxonomy databases, including archaea, bacteria, fungi, viridiplantae, unclassified, and other taxonomies (includes plasmids and artificial sequences). Tables S1 and S2 show the peptide fragments that matched the protease type from the excised gel of 44 and 108 kDa, respectively. The highest Mascot Protein Score was with the bacteria kingdom. More than 80% of the peptide sequence of the PE matched with the genera *Gallaecimonas* and *Alteromonas*. These bacteria are gram-negative and belong to proteobacteria phylum isolated from intertidal sediment and found in seawater [29,30].

We have previously shown that the *G. edulis* proteases belong to serine and metalloprotease types [7]; therefore, to narrow down the

Table 1
Proteins identified of PE from 44 kDa gel in NCBI Eukaryotes and Other Eukaryotes taxonomy database.

No.	Homology with protein from	Accession number	Mass	Mascot score	Peptide sequences	Sequence coverage
1	Photosystem I P700 chlorophyll a apoprotein A1 [*Gracilaria tenuistipitata var. liui*][a]	YP_063614.2	83,550	102	1 MTISSQEQET KKVQVTVDKN PVATSFEKWA KPGHFSRTLA KGPKTTTWIW NLHADAHDFD 61 SHTSSLEEVS RKIFSAHFGQ LAIIFLWLSG MYFHGAKFSN YIAWLSNPTA IKPSAQIVWP 121 IVQQEILNGD VGGGFQGVQI TSGFFQIWRA SGITTEFELY ATAIGGLFMA CLMLFAGWFH 181 YHKAAPKLEW FQNVESMMNH HLAGLLGLGC LGWAGHQIHI SIPINKLLDS GVSPQELPLP 241 HEFLVN**RELV** **SQLYPSFSKG** IIPFPTLNWN EYSDFLTFKG GLNPITGGLW LTDTAHHHLA 301 LAVLFLVAGH MYRTDWGIGH SMKEILEAHK GPFTGEGHKG IYEILTTSWH AQLAINLAMM 361 GSLSIIVAHH MYAMPPYPYI ATDYPTQLSL FTHHMWIGGF CIVGAGAHAS IPMVRDYNPA 421 QNYNNVLDRV IRHRDAIISH LNWVCIFLGF HSFGLYIHND TMRALGRSQD MFSDTAIQLQ 481 PIFAQWIQNI HSLAPSNTSP NALATASYAF GGIVIAVNNK IAMMPINLGT ADFMVHHIHA 541 FTIHVTVLIL VKGFLFSRNS RLIPDKSSLG FRFPCDGPGR GGTCQVSGWD HVFLGLFWMY 601 NSLSIAIFHF SWKMQSDVWG SVTPAGTVSH ITGGNFAQSA ITINGWLRDF LWAQASQVIQ 661 SYGSALSAYG LIFLGAHFVW AFSLMFLFSG RGYWQELIES IVWAHNKVKV APSIQPRALS 721 ITQGRAVGVA HYLLGGIGTT WAFFLARIIS VG	2%
2	Photosystem I P700 chlorophyll a apoprotein A1 [*Gracilaria tenuistipitata var. liui*][b]	YP_063614.2	83,550	118	1 MTISSQEQET KKVQVTVDKN PVATSFEKWA KPGHFSRTLA KGPKTTTWIW NLHADAHDFD 61 SHTSSLEEVS RKIFSAHFGQ LAIIFLWLSG MYFHGAKFSN YIAWLSNPTA IKPSAQIVWP 121 IVQQEILNGD VGGGFQGVQI TSGFFQIWRA SGITTEFELY ATAIGGLFMA CLMLFAGWFH 181 YHKAAPKLEW FQNVESMMNH HLAGLLGLGC LGWAGHQIHI SIPINKLLDS GVSPQELPLP 241 HEFLVN**RELV** **SQLYPSFSKG** IIPFPTLNWN EYSDFLTFKG GLNPITGGLW LTDTAHHHLA 301 LAVLFLVAGH MYRTDWGIGH SMKEILEAHK GPFTGEGHKG IYEILTTSWH AQLAINLAMM 361 GSLSIIVAHH MYAMPPYPYI ATDYPTQLSL FTHHMWIGGF CIVGAGAHAS IPMVRDYNPA 421 QNYNNVLDRV IRHRDAIISH LNWVCIFLGF HSFGLYIHND TMRALGRSQD MFSDTAIQLQ 481 PIFAQWIQNI HSLAPSNTSP NALATASYAF GGIVIAVNNK IAMMPINLGT ADFMVHHIHA 541 FTIHVTVLIL VKGFLFSRNS RLIPDKSSLG FRFPCDGPGR GGTCQVSGWD HVFLGLFWMY 601 NSLSIAIFHF SWKMQSDVWG SVTPAGTVSH ITGGNFAQSA ITINGWLRDF LWAQASQVIQ 661 SYGSALSAYG LIFLGAHFVW AFSLMFLFSG RGYWQELIES IVWAHNKVKV APSIQPRALS 721 ITQGRAVGVA HYLLGGIGTT WAFFLARIIS VG	2%
3	Photosystem I P700 chlorophyll A apoprotein A2 (chloroplast) [*Gracilaria salicornia*][b]	YP_009019621.1	82,568	91	1 MATKFPKFSQ ALSQDPTTRR IWYGIATAHD FESHDGMTEE NLYQKIFSSH FGHIAIIFLW 61 TSGNLFHVAW QGNFEQWVTQ PMKIKPIAHA IWDPHFGQPA IKAFTKGGVS YPVDITFSGV 121 YHWWYTIGMR TNNDLYNASL FLLVLSAIML FAGWLHLQPK FKPGLSWFKN NESRLNHHLS 181 GLFGLSSLAW TGHLVHVAIP ESRGQHIGWD NFTYTLPHPS GLQPFFTGNW NVYAANPDTS 241 NHIFGTSQGA GTAILTFLGG FHPQSQSLWL TDIAHHHLAI AIIFIIAGHM YKTNWGIGHN 301 IKDIIDAHRP PSGKLGKGHI GLYETITNSL HIQLGLALAS LGVITSLVAQ HMYAMPPYAF 361 MAKDFTTQAA LYTHHQYIAG FLMVGAFAHG AIFFIR**DYDP** **EENKNNVLAR** **M**LDHKEAIIS 421 HLSWVCLFLG FHTLGLYIHN DTMIAFGTPE KQILIEPVFA QWIQASSGKA LYGFDTLLSS 481 SSNIATKASG NIWLPGWLEA INSNTNSLFL TIGPGDFLVH HAIALGLHTT TLILVKGALD 541 ARGSKLMPDK KDFGYSFPCD GPGRGGTCDI SAWDAFYLSV FWMLNTIGWV TFYWHWKHMT 601 IWQGNVNQFN ESSTYLMGWL RDYLWLNSSP LINGYNPYGM NNLSVWAWMF LFGHLIWATG 661 FMFLISWRGY WQELIETLAW AHF**RTPLANL** **IRWKDKPVAL** **SIVQARL**WGL VHFSVGYILT 721 YAAPVIASTS GKFG	5%
4	GDP-mannose-3′, 5′-epimerase [*Gracilaria changii*][b]	AGL46678.1	38,875	36	1 MTDTKR**RVLV** **TGAGGFIGGH** **LAKR**LKDEGN HVVGADWKEN EYFKVEEFCS EFMLLDLRDK 61 ENCMKA**S**KDC EDIYNLAADM GGMGFIQSNH SVILYNNAMI SFNMLEAART SGCKRFYYAS 121 SACIYPEGAQ LDPNNPGLKE SDAWPAQPQD AYGLEKLASE ELAIHYGKDF GMKTRIGRFH 181 NIYGPQGTWK GGREKAPAAF CRKSITSTTE FEMWGDGNQT RSFCYVDDAV EGVIRLMNSD 241 FTEPLNIGSD EMISMNDMGK MVLGFAAKDI PIKHIPGPEG VRGRNSDNTL CQKVLGWVPG 301 IKLEDGLRRT FDWIKSQIEE EKAKGIDNDY STSTVVGTHA PTDSKASKRK	5%

[a]Derived from NCBI Eukaryotes database.
[b]Derived from NCBI Other Eukaryotes database.
Shared amino acids are shown as bold and underline characters.

Table 2
Proteins identified of PE from 108 kDa gel in NCBI Eukaryotes and Other Eukaryotes taxonomy database.

No.	Homology with protein from	Accession number	Mass	Mascot score	Peptide sequences	Sequence coverage
1	Photosystem I P700 chlorophyll a apoprotein A1 [*Gracilaria tenuistipitata var. liui*][a]	YP_063614.2	83,550	103	1 MTISSQEQET KKVQVTVDKN PVATSFEWWA KPGHFSRTLA KGPKTTTWIW NLHADAHDFD 61 SHTSSLEEVS RKIFSAHFQQ LAIIFLWLSG MYFHGAKFSN YIAWLSNPTA IKPSAQIVWP 121 IVGQEILNGD VGGGFQGVQI TSGFFQIWRA SGITTEFELY ATAIGGLFMA CLMLFAGWFH 181 YHKAAPKLEW FQNVESMMNH HLAGLLGLGC LGWAGHQIHI SIPINKLLDS GVSPQELPLP 241 HEFLVNRELV SQLYPSFSKG IIPFFTLNWN EYSDFLTFKG GLNPITGGLW LTDTAHHHLA 301 LAVLFLVAGH MYRTNWGIGH SMKEILEAHK GPFTGEGHKG IYEILTTSWH AQLAINLAMM 361 GSLSIIVAHH MYAMPFPYFY ATDYPTQLSL FTHHMWIGGF CIVGAGAHAS IFNV**RDYNPA** 421 **QNYNNVLDRV** IRHRDAIISH LNWVCIFLGF HSFGLYIHND TNRALGRSQD MFSDTAIQLQ 481 PIFAQWIQNT HSLAPSNTSP NALATASYAF GGDVIAVNNK IAMMPINLGT ADFMVHHIHA 541 FTIHVTVLIL VKGFLFSRNS RLIPDKSSLG FRFPCDGPGR GGTCQVSGWD HVFLGLFWMY 601 NSLSIAIFHF SWKMQSDVWG SVTPAGTVSH ITGGNFAQSA ITINGWLRDF LWAQASQVIQ 661 SYGSALSAYG LIFLGAHFVW AFSIMFLFSG RGYWQELIES IVWAHNKV**KV APSIQPRA**LS 721 ITQGRAVGVA HYLLGGIGTT WAFFLARIIS VG	3%
2	Photosystem I P700 chlorophyll A apoprotein A2 (chloroplast) [*Gracilaria salicornia*][b]	YP_009019621.1	82,568	148	1 MATKTFKPSQ **ALSQDYTTRR** IWYGLATAHD FESHDGMTEE NLVQKIFSSH FGHIAIIFLW 61 TSGNLFHVAW QGNFEQAVTQ PMKIKPIAHA IWDPHFGQPA IKAFTKGGVS YPVDITFSGV 121 YHWWYTIGMR TNNDLYNASL FLLVLSAIML FAGWLHLQPK FKPGLSWFKN NESRLNHHLS 181 GLFGLSSLAW TGHLVHVAIP ESRGQHIGWD NFTYTLPHPS GLQPFFTGNW NVYAANPDTS 241 NHIFGTSQGA GTAILTFLGG FHPQSQSIWL TDIAHHHLAI AIIFIIAGHM YKTNWGIGHN 301 IKDIIDAHKP PSGKLGKGHI GLYETITNSL HIQLGLALAS LGVITSLVAQ HMYAVPPYAF 361 MAKDPTTQAA LYTHHQYIAG FLMVGAFAHG AIFFIRDYDP **EENKNNVLAR** MLDHKEAIIS 421 HLSWVCLFLG FHTLGLYIHN DTMIAFGTPE KQILIEPVFA QWIQASSGKA LYGFDTLLSS 481 SSNIATKASG NIWLPGWLEA INSNTNSLFL TIGPGDFLVH HAIALGLHTT TLILVKGALD 541 ARGSKLMPIK KDFGYSFPCD GPGRGGTCDI SAWDAFYLSV FWMLNTIGWV TFYWHWKHMT 601 IWQGNVNQFN ESSTYIMGWL RDYLWLNSSP LINGYNPYGM NNLSVWAWMF LFGHLIWATG 661 FMFLISWRGY WQELIETLAW AHE**MTPLNKL IRWKDKPVAL SIVQARL**VGL VHFSVGYILT 721 YAAFVIASTS GKFG	7%
3	Photosystem I P700 chlorophyll a apoprotein A1 [*Gracilaria tenuistipitata var. liui*][b]	YP_063614.2	83,550	134	1 MTISSQEQET KKVQVTVDKN PVATSFEWWA KPGHFSRTLA KGPKTTTWIW NLHADAHDFD 61 SHTSSLEEVS RKIFSAHFQQ LAIIFLWLSG MYFHGAKFSN YIAWLSNPTA IKPSAQIVWP 121 IVGQEILNGD VGGGFQGVQI TSGFFQIWRA SGITTEFELY ATAIGGLFMA CLMLFAGWFH 181 YHKAAPKLEW FQNVESMMNH HLAGLLGLGC LGWAGHQIHI SIPINKLLDS GVSPQELPLP 241 HEFLVNRELV SQLYPSFSKG IIPFFTLNWN EYSDFLTFKG GLNPITGGLW LTDTAHHHLA 301 LAVLFLVAGH MYRTNWGIGH SMKEILEAHK GPFTGEGHKG IYEILTTSWH AQLAINLAMM 361 GSLSIIVAHH MYAMPFPYFY ATDYPTQLSL FTHHMWIGGF CIVGAGAHAS IFNV**RDYNPA** 421 **QNYNNVLDRV** IRHRDAIISH LNWVCIFLGF HSFGLYIHND TNRALGRSQD MFSDTAIQLQ 481 PIFAQWIQNT HSLAPSNTSP NALATASYAF GGDVIAVNNK IAMMPINLGT ADFMVHHIHA 541 FTIHVTVLIL VKGFLFSRNS RLIPDKSSLG F**RFPCDGPGR** G**G**TCQVSGWD HVFLGLFWMY 601 NSLSIAIFHF SWKMQSDVWG SVTPAGTVSH ITGGNFAQSA ITINGWLRDF LWAQASQVIQ 661 SYGSALSAYG LIFLGAHFVW AFSIMFLFSG RGYWQELIES IVWAHNKV**KV APSIQPRALS** 721 **ITQGRA**VGVA HYLLGGIGTT WAFFLARIIS VG	6%
4	Galactose-1-phosphate uridylytransferase [*Gracilaria gracilis*][b]	AAB88705.1	42,417	42	1 MSASFDYTEH PHRRYNPLSA RWILCSPHRA KRPWQQSVED LPFDERPEYD PKDYLGPGNF 61 RVNGSVQNPK YDTTFVFDND FQALLDNTPH GEVGSVEDND LLVAKAVRGK CRVCFSPKL 121 NLTVAEMTVE EIKHVVDAWL EEYDTLSKLD YIGHVQIFEN KGQMMGCSNP HPHQQIWASE 181 FVPEEPRIVL ENLKAYHEKK GTHMLEDYVK IEMEAKERIV CENDTFLAVV PFWATWPFEV 301 HDYAHFMMHF YPPLLRSATV RKFMVGFEML GESQ**BDLTAE QAAARL**RACS EVHYNHARKK 361 VEGNGDASK	3%

[a]Derived from NCBI Eukaryotes database.
[b]Derived from NCBI Other Eukaryotes database.
Shared amino acids are shown as bold and underline characters.

search, these two types of protease were manually included in the database search (Tables 3 and 4). The metalloprotease type of protease was detected only in the 44 kDa band while the serine-type protease was identified in both the 44 and 108 kDa bands. There are nine serine and metalloproteases in the database with matching sequences to the 44 kDa band (Table 3), while there are three serine proteases with matching sequences to the 108-protease band (Table 4). Also, the sequence coverage of the 44 kDa band was generally higher than that of the 108 band, with the highest coverage was for zinc metalloprotease (*Gallaecimonas xiamenensis* 3-C-1) at 23%, followed by low quality protein (the homolog of the protein may not be derived directly from the genome sequence [31]) alkaline serine protease (*Vibrio* sp. JCM 18904) (17%).

Interestingly, the Mascot score for both protein bands was above 50, which confirmed the identity of the protein [32]. The range of the Mascot score of the 44 kDa band was from 55 for the metalloprotease yebA (*Clostridium* sp. CAG: 1193) to 1008 for the zinc metalloprotease (*Gallaecimonas xiamenensis* 3-C-1). The protease identified from the 108 kDa band exhibited lower Mascot scores, ranging from 54 for alkaline serine protease (*Kangiella koreensis*) to 281 for subtilase family serine protease (*Alteromonas macleodii* str. 'Balearic Sea AD45').

The identification parameters of these proteins were comparable to those in previous studies. A metalloprotease purified from *Paenibacillus* spp. BD3526 showed a Mascot score of 65 with a sequence coverage of 29% [33]. Some plant proteases from *Cynara cardunculus* L. and *Withania coagulans* were also claimed to be identified with the Mascot score range of 47–82 and sequence coverage range of 5–39% [34,35].

With regard to molecular mass, the homology proteins showed that the mass ranged from 28,736 to 139,974 Da and 64,283 to 139,974 Da for the excised gels of 44 kDa and 108 kDa, respectively. These, however, differed from the masses of the homolog protein in Tables 3 and 4. Such difference indicated that the protease has not been previously reported or the homology protein may consist of several protease subunits [36]. Similar finding has been reported by Hang et al. [37] who found that the mass of a novel metalloproteinase enzyme extracted from *Paenibacillus* spp. BD23526 was different from the matching hypothetical protein. They argued that the mass difference suggested a novel enzyme or an existing zymogen.

The fact that the protein sequence of the isolated seaweed protease sequences had higher matching scores with bacteria than with other eukaryotes could be due to two causes. The first and obvious cause is contamination of the algal sample by bacteria. This, however, is considered very unlikely because for this to occur, the amount of protein from the contaminated bacteria must be overwhelmingly greater than that from the alga. This would mean that there must have been massive and visually noticeable bacterial growth on the algal sample, which was not observed. The second, and more likely cause of the results might be attributed to several evolutionary and environmental factors. The first is bacterial colonisation of plants and their subsequent symbiosis through coevolution. A recent study by Souza et al. [38] suggests that because of the plant colonisation by bacteria, some of the plant metabolite productions are stimulated by the bacteria. Moreover, symbiosis with bacteria also enables the plant to adapt and exchange signals, including those in the formation of rRNA and other major cell organelles [39]. Pace [40] reports that the respiratory, photosynthetic and metabolism systems of eukaryotic cells are possibly derived from the bacterial symbionts, *e.g.*, the protein sequence of plant mitochondria and chloroplasts may be obtained from proteobacteria and cyanobacteria.

The second factor is the horizontal gene transfer (HGT) theory, explained by Nikolaidis, Doran and Cosgrove [41], which suggests that the transmission of genetic materials between bacteria and plant is a common occurrence in nature. They evaluated the transfer of genetic material from eukaryotes to prokaryotes and found that non-enzymatic proteins of the plant were transferred from plant to bacteria. Similarly, Givens et al. [42] discovered similarities between the proteins of the stress-adaptive enzyme in a plant to that in bacteria. Finally, the HGT affects evolution traits between eukaryotes and bacteria [41]. Furthermore, Givens et al. [42] claimed that there is an evolutionary phylogenetic relationship between plants and bacteria which is built through ancient symbiosis and nucleus gene transfer between them. For instance, the plant catalytic enzyme involved in the synthesis of α-carotene was

Table 3
Protease identified of PE from 44 kDa gel in NCBI Bacteria taxonomy database.

No.	Homology with protein from	Accession number	Mass	Mascot score	Peptide sequences	Sequence coverage
1	Zinc metalloprotease [*Gallaecimonas xiamenensis* 3-C-1]	EKE76739.1	76,696	1008	1 MNKLALAVLS TAAFAPLAMA ANIVDIKPGS PRAEALKPDQ DLVKMLGLPQ TQELRKKVIK 61 RGGIVKTRIK QRFMGVEVFG LSMAADMDTM GIYHNLTGQY VEGLEGDLAD VKPAFNEKAA 121 QAKAVAYHGL ATPERGKTSL FIWLDEEQKA HLVYLVDQFV GGDEPRRPMT IIDAHSGEVL 181 KSWDQLAFQD ATGPGGNQ**KT GQYFYGTDFG YLDVDSNCRM DNTNVTTVNM NNSTSGGSVF** 241 **QFTCPENTYK S**VNGAYAPLN DAHYFGGVIF NMYSDWLQTA PLTFKLTMRV HYGNSYENAF 301 WDGSAMTFGD GGSTFYPLVS LDVSAHEVSH GFTEQNSGLI YSGQSGGINE AFSDMAGEAA 361 EFYM**KGSNDW MVGAEIFKGS GALRYMDQPS RDGSSIDNAS DYYSGMDVHY SSGVYNRAFY** 421 **LIANSSGWDT RKA**FELFALA NQVYWNASSD YDDAACGVVQ AATDLGFSVA DVQTAFTTVG 481 VDSSCGSTPP PEGGELTNGV PETGISGASG SEQFWTLDVP AGASNLVFNT SGGSGDADLY 541 V**RFGSAPTSS SYDCRPYKSG NSESCSFSAP QSGTYHVMLR G**YSAYSGLTL TGSYDADSTT 601 PGSYENGTDT NIPDNNTTGI QSVIAVDRSG DSGTVSVQVD IVHSYRGDLQ IDLLAPNGQS 661 FRLKSIGSDS ADDVHETYSV NASGIDSSGT WTLQVSDRYS QDTGYLDYWS LSFN	23%
2	LOW QUALITY PROTEIN: alkaline serine protease [*Vibrio sp.* JCM 18904]	GAJ70088.1	71,249	586	1 MHYTTSKKQG LFSLSAVAIA LTTSMSAFAA PPMGFDESKL PTKYIVKFKE DAVSRSSMGG 61 NSFWGGHESR RSLFLEQVKA RKVEKLGNRA IYSVELDGDD LEPLRNRSDV EYVEVDPPRY 121 LLSETTPWGY EAVNAQLLAD FNAGNRTVCI IDSGYDISHN DLSGNRVAGT NDSGTGSWSD 181 PGNNNAHGTH VAGTIAAIAN TEGVKGVMPN QNVNLHIVKV FNESGWGYSS GLV**KAIQTCA** 241 **DNGANVVNMS LGGSQSSRT**E QNALQNIYDQ GVLLIAAAGN DGNTAHSYPA SYDSVMSVAA 301 VDNQNDHAAF SQSTNQVEIS GPGVAILSTV TVGEG**KLSDI TLNGVSQFDR G**IVPHNRLIN 361 NGSSYVPDPI AGSVTATLDS CDVSGGNFNC GDMSGKICLT E**RIGNQSSGN YPEVDAVQAC** 421 **YNAGARAAIV YSNSELPGLQ NPFLVDTNNA YRM**VSVTVDR AFGQELLGGY VGQEITVATT 481 SGEDYEYYNG TSMATPHVTG VAGLVWSYHP TCTAAQVRNA V**KTATDIDVA GRD**N**RTGHGL** 541 **VNAEAAKL**FL DAGCNGPDGG SSSDSSFSNT SPVAIPDNKS SGAISAVEVD RSGDSGTVSI 601 DVDISHTYIG DLRVTLTSPT GGEVVLHNNT GGSANDIKNT FQADFSGFES QGTWELKAVD 661 SARRDTGTIN SWTLTFQ	17%
3	Subtilase family serine protease [*Alteromonas macleodii* str. 'Balearic Sea AD45']	AFT95466.1	139,974	326	1 MTFEANKKIV GLQGTIHTRI IMKKTILSTV IATALAGSAF GALAISNQLS GYGVEQQNAI 61 EASFDKKNIL ANAIQANSDS VRVIVQLTDV PMAQFSAVNP SVSSMAAHKG KKVNFESSAA 121 KEYKSFLESQ QQSVIQSIKS FDKSFKADMS YTAAFNGFAG IVSKSALDQL SSLSTVKAVY 181 PDVIHHAQMD ASLDLIGAVE TWEQLGGKEN AGAGVKVAII DSGIRPQNPL FSGENFEAPP 241 ADTLPTDDYC SEVPDFCNNK LIVARAADIV EGFAVVEEEY ESPLGFNGHG THVAGTAVGN 301 YGVMAERDGA EAEISGVAPA AYIMVYKGLY ATPANPASSS GASSMLLSML EAALTDGADI 361 VNNSWGGGAG GNPNGSIYED VLEAMHDAGV VTVFAAGNDG PNEGTIGCPG CSEDVITVAN 421 TTTGRLFANE VTIEGDTTIG SIPALYSVGN PAIVFDSPIT APVVYAGEVD AANVEGCDAF 481 AAGAFDGAIA LISRGTCGFV T**KIENAEAAG ATAVLVHNVD GRG**EAPIIMG GLSEAQTIPS 541 LMLPATSGQE LASLAVATDE ALNVTIGSDI V**RVVSDSLAD IMNESSSRGP NGDPSFLKPN** 601 **IAAPGTRIFS GESPDAPGHE GESFSFKN**GT SMASPHVAGA AALLKQMHPD WTAQQIKSAL 661 VTSSIRDVLK EDATTQADNF DMGAGRLDLP RATTVELTYS DLSLVDGNCY LNCEMSITVT 721 NTSDEDITVD ATAMFNDPAI SATVTPQMAT LPAGASAEIM VAVDVTTAST GSWSFGGINW 781 ADTDDTTTDY FIPVAVYPIS TDRPELFSSD VSTQVAEEGQ LVRVSASATN QNVTGMIGVT 841 GTVDHKFEID PSTIAAVKNG NQEPVSYDAD KGQIYWEGAL NTSSYSLDAD TTIAALTGGY 901 LPMAALGVQP LTCSGSCDDT SITISDLPEM TYLGKAYTSM QISSNGYISL GTSSGNVTTP 961 TPAVMPNSNE PNNVLAPFWT DLDLLGTDAG DEGTGNLYAA SLNGPNGSLL VIEWENAQLW 1021 GIPDTSFNFQ IWYDFAFDTI HFVYGNMDSP QYATVAGFEN PTGTAGFTLG ALTSAGAQGT 1081 LPVSGDEFIL ASSPGDEIEF SYNGTVVPSS EYMDDMLTVN EDASVTANIL ANEMDSTIIN 1141 TFTMESLSGD FRTFTPITID KAPLDPSTVE VTTEPANGTV TVNEDGTVTY APNTNFFGED 1201 SFTYTVRVEG TIDENEEPVE GDIVGEGTVT VAVAGVQDAP VISISAPTSV DEGESYTVTA 1261 SATDPDGDDV TITVNGMATS SITDTAPSHE QARQVTVEVT ATDGIDTTTE TVTIRVNDTS 1321 GGGSMGWIAL LLAPAVYLRR RMKRS	6%
4	Homoserine O-acetyltransferase [*Gallaecimonas xiamenensis*]	WP_008486463.1	42,190	255	1 MRYWILALGL CWALGAQALP LVEKQ**RFALD SFTTQSGTTL KE**VAVGWEAY GTLNADKSNA 61 ILITHFFSGS SHAAG**KYSDQ DPLPGYWDAI IGPGKA**IDTD KYYVLSVDTL ANANAFDPHV 121 ITTGPASINP ATGKPYGLDF PVVSIRDFVN VQKALLDKLG INKLHAVVGA SMGSFQALDW 181 AVAYPDKVER MVSVIGAGQM DPWTVYGLER WSDPIKADPA WHNGHYYEQG QPKTGLT**KAV** 241 **AYIIYDATYP DGFNSRY**TPP TDAAPKSDIR AGYKSVDELM GHAAIRAHFQ DANAILYLVR 301 ASQMFLAGYN GKLEDNLAKV SAKTLFMPAS HDRLLVPKMA RSTYETLKKL GKDSQYQEID 361 GVWGHLDGLV NIQSQGQVLK DFLDK	14%
5	C-terminal processing peptidase-1, Serine peptidase, MEROPS family S41A [*Frischella perrara*]	AJA45688.1	80,420	126	1 MKKYNYYLSI KRCLSSLIIF TACLTSVSQA KEYSIDQLPK LESDPIHQVV SRRVTNYFTQ 61 SHFRKFDLDG VFSSKIFDRY FKLLDSNKTI FIKSDIDTFR ERQTGLGKEL RDGDLRTAFD 121 IYNLSLKKRF ERYQFALAQL KEPMDFSTSE SINFKRDDLA WAVSEQELDD YWRKRVKYDE 181 LSLALSGKKE LEIREILTKR YNQILRTLVQ VKPEDAFQVF MNAFAREIDP HTSYLAPRTK 241 RDFDSEMSLS FEGIGATLSQ EDDYTRIVSF VTGGPAEKSK QLAIGDRIIG VGQKNNPIED 301 VIGWRLDDIV DKIRGPKGTI VKLEILPAGN NSKTKIIEIK RDKIHFEDRE AKLTIKQTAQ 361 GKVALIDIPS FYMGLTDRVV KLLTEANKNN VSGIVIDLRN NGGGSLAEVI SLTGLFIEKG 421 PVVQVKDNLQ SVVVYDDRDE SVQYVGPIVV MVN**RYSASAS EIFSAALQDY GRA**VIVGEDT 481 YGKGTVQTSR NIAYPIDATI HPNWPALGGV QYTIQKFYRI NGGSTQLRGV MPDIEMSPLR 541 YIDDTGERYL DNALPWDSVA VADYHVLFDI KSILPELKQH HLERIKNDPE FNYIEADIKK 601 YNDNKDQQYV VSLNKVEREK KQKETDNEEL LRMNERLKLA GKQAISKLED LPKDFKVKDA 661 YLDEAEAILF DLAKLYPDIK VSQLPSNAIM LDLNSPIMND KESK	3%
6	Putative metallopeptidase [*Alteromonas mediterranea* 615]	AGP77421.1	106,174	116	1 MSGKLALRTL CMALGVIVVT ACQHAPVDHA QSQAEAATNS SNTVSIPYEK YTLENGLTVI 61 LHEDHSDPLV HVDVTYHVGS AREDVGKSGF AHFFEHMMFQ GSKHVADEQH FKVITESGGN 121 LNGTTNTDRT NYFETVPANQ LEKVLWLESD RMGYLLEAVD QTKFENQRET VKNERAQRVD 181 NQPYGLRYEL NGEALYPEGH PYSWMTIGYV EDLDRVNVND LKAFFKRWYG PNNAVLTIGG 241 DIDVAKTKAW IKKYFGEIPA GPAVEEPEPQ PVTLTETRYM TLEDKVHLPL LQITYPTVYG 301 RHEDEAPLDV LADILGGGKT SLFYKNLVKE GMAVQAVVSH PC**RELACEFQ LLALANPAKI** 361 TSLSTLQEVL NQTLKEFETR GVTADDLART KGQIEARTVF GLQSVSGKVS ALAANETFYQ 421 TPDLIAEDIE RYNAVTADDV MRVYNKYIKD ANSVVLSVVP KGQVQLAAAE QTFERPVRNI 481 HVETVDVAGE EAFTSAPSSF DRSVMPKAGD APVVEVPDYW EAELANGIKI LGVTSTETPT 541 VTLTLGMDGG MLLDSEGKAG TAYLTALIMN ETTKHYSNEA LASELAKLGS SIRFSTAGRY 601 SQVYVSTLTK HLDETLALLK EKLFNPAFTE EDFERMKE**RV VQGLQQQAKT** PSSLARRARD 661 LILFGEDNRV SLPDEGTLET VQSITLDDVK TFYKNYYSPD KASIVAVGNL SKKDMVETLD 721 FIGQWQGNAY EFADYSDFPQ YNQNQIFLID SPEAVQSVVY IVDRSLPFDA TGDHFKSRLM 781 NFPLGGAFNS RINLNLREDK GFTYGANSGF VGGKTLGWFE VSTDLTAANT GEGIKEILGE 841 IERYRSEGVE KAEIDFMRNA FTLSDALEFE TPTSKARFLR QLLSYGLEKG YREAQLDIIN 901 NIDKESIDAL AKQYLNLDKM QIIVVGDKAK ILPQLNALSM PIIELSVEGN RREALN	3%
7	CDP-diacylglycerol/serine O-phosphatidyltransferase [*Nitrosococcus halophilus* Nc4]	ADE13838.1	28,736	76	1 MDSEQKPLLA AGAGNSRRRR GIYLLPNLLT TAGLFAGFYA IVAAMNGRFE AAAVAIFVAM 61 LMDGVDGRIA RLTNTQSAFG AEYDSLADMV SFGLGPALVI YEWTLHSLGK LGWLAAFLYA 121 ACAALRLA**RF NTQVGIADKR Y**FQGLASPSA AACLAGLVWF AVDAGIAGAA LLLPAFILTI 181 LTAVFMVSNI RYYSFKGLDS REKVPFVAIL LVVLIFVLIS TDPPKVLFIA FLLYALSGPV 241 VTLVQIRQHR IQRRAPLGST KVDDDEK	5%
8	Alkaline serine protease [*Kangiella koreensis*]	WP_015780787.1	64,283	57	1 MKLNQIVAGI LLATGVTGAY AGEFKTASVP SKAIKGQYIV VLKDDAVAQN MGQFSSNASE 61 QAIQMMTENL SRKYQAQVQR TYTKALKGGT FTLSEKAAQK LAQDPNVLLV EEDQIVSLNA 121 TQNNATWGID RVDQRNLPLS GTYTYNTTAS NVNAYIIDTG ILNSHSEFGG RSVSGIDTVD 181 NDNDATDCNG HGTHVAGTVG GSTYGIAKSV NLIGVRVLNC SGSGTLSGVI DGIDWVASNH 241 TKPAVANMSL GGGASSSIDT AVANLVSSGV TVVVAAGNDN SSACNYSPAR EPSAITVGST 301 TSSDSRSSFS NYGSCLDIYA PGSSITSAWS NGGTNTISGT SMASPHVAGV AALYLADNPS 361 ATPSQVDTAI TDSATPGVVS DAKSGSPNLL LYSLFDGTNP DPDPDPEPGN ELENGVSVTF 421 SGAQYSETDF TFEVPSAASN VSFDMSGGSG DADLYV**KFGS APTTSSYDCR PYRN**GNNESC 481 DFSAQTGTYY VMVRGYTSYS NVSLVATHDG DGSEPPAEGG SETYPNLSGA SGSWTHYYVD 541 IPAGMSSLDV QMSGGSGDAD LYVRRGAQPT TSSYDCRPYR WGNDETCSFS NPAQDRWYIS 601 IRGYSSYSGV TLQIDWQ	3%
9	Metalloprotease yebA [*Clostridium sp.* CAG: 1193]	CCY45094.1	55,024	55	1 MSTKEKLVAI SITILLSVFS IIPIYEFKVS AKTTPRDLYR VYLNGKNIGV IESKEKLEKY 61 INEDQKELKD EFGVDTVYLP KGLYISKYST YNATLTDEKD IYETIKQTEH FTIKGYTITI 121 KKEDENKEDI KINVLNKTDF ENAVTSVVKA FVPNEELDVF LSGEDITIKD TGKKIEDLYI 181 KEDIKIKESY ISSDEQILLD ERSVTKYLLF GEDIDEKKYI VKAGDTIESI AEDNKLAVEE 241 LLVVNQDLKS KNNLLSIGEE ISVALISPVI TVVEEEHLVE VKTIPYATQI EYDNSLGWGL 301 TKIKQEGKDG SQIVTQ**KLKY ENGEIIRA**LI SDTKVIDEVI NQVKVIGTMS SIQIDPSDIP 361 DAGDWYWPTL QPSIITSPYG WRWGALHDGT DISGTGNGSP IFAANNGVVY KVFYDGIGGY 421 QIIIAHENNY YTWYAHLSAQ YVTAGQKVTR GTKIGAMGCT GSACTGPHLH FGAYKGIPGR 481 GGISFNAMTL YR	2%

Shared amino acids are shown as bold and underline characters.

found in homology with those of prokaryotes and archaea [43]. This finding identified a direct line of evolution from archaea and some groups of bacteria *via* cyanobacteria and green algae to plants.

Ramanan et al. [44] reported that the mutualistic symbiosis between algae and bacteria has synergistically influenced their metabolism, which has played a key role in the primary endosymbiosis stage of algae revolution. In this stage, the algal ancestor used a cyanobacterium organelle as a photosynthesis agent. The host cell of the cyanobacterium organelle has been argued to have originated from either bacteria or archaea. According to Schönknecht et al. [45], in the extreme environment, the HGT from bacteria and archaea to eukaryotes algae was likely to happen due to the environmental adaptation and helped the algae to survive under high/low temperature and other stressful conditions. They reported that the protein-coding genes of *Galdieria sulphuraria*, a microbial eukaryotes red alga, were in homology with bacteria for at least 5%. Pennisi [46] also reported that two species of the red algae

Table 4
Protease identified of PE from 108 kDa gel in NCBI Bacteria taxonomy database.

No.	Homology with protein from	Accession number	Mass	Mascot score	Peptide sequences	Sequence coverage
1	Substilase family serine protease [*Alteromonas macleodii* str. 'Balearic Sea AD45']	AFT95466.1	139,974	281	1 MTFENKKIV GLQGTIHTRI IMKKTILSTV IATALAGSAF GALAISNQLS GYGVEQQNAI 61 EASFSKKNIL AWAIQANSIE VRVIVQLTIV PNAQFSAVNP SVSSNAAHKG KKVNFESSAA 121 KEYKSFLESQ QQSVIQSIKS FDKSFKADMS YTAAFNGFAG IVSKSALDQL SSLSTVKAVY 181 PDVIHHAQMD ASLDLIGAVE TWEQLGGKEN AGAGVKVAII DSGIRPQNPL FSGENFEAPP 241 ADTLPTDDYC SEVPDFCNNK LIVARAADIV EGFAVVEEEY ESPLGFNGHG THVAGTAVGN 301 YGVMAERDGA EAEISGVAPA AYLMVYKGLY ATPANPASSS GASSMLLSML EAALTDGADI 361 VNNSWGGGAG QNPNGSIYED VLEAMHDAGV VIVFAAGNDG PNEGTIGCPG CSEDVITVAN 421 TTTGRLFANE VTIEGDTTIS SIPAISVGNP PAIVFDSPIT APVVAGEVD AANVEGCDAF 481 AAGAFDGAIA LISRGTCGFV TKIENAEAAG ATAVLVHNVD GRGEAPILMG GLSEAQTIPS 541 IMLPATSGQE LASLAVATDE AINVTIGSDI VMVVSDSLAD IMNESSSRGP NGDPSFLKPN 601 IAAPGTRIPS GESPSDAPGHE GESFSFKNGT SMASPHVAGA AALLKQKNPD WTAQQIKSAL 661 VTSSIRDVLK EDATTQADNF DMGAGRLDLP RATTVELTYS DLSLVDGNCV LNCEMSITVT 721 NTSDEDITVD ATAMNNDPAI SATVTPQMAT LPAGAGAAEM VAHVVTTAST GSWSFGGINW 781 ADTDDTTTDY FIPVAVYPIS TDRPELFSSD VSTQVAEEQQ LVRVSASATN QNVTQMIGVT 841 GTVDHKFEID PSTIAAVKNG NQEPVSYDAD KQQIWEGAL NTSSYSLDAD TTIAALTGGY 901 LPMAALGVQP LTCSGSCDDT SITISDLPEM TYLGKAYTSM QISSNGYISL GTSSGNVTTP 961 TPAVMPNSNE PNNVLAPFWT DLDLLGTDAG DEGTGNLYAA SINGPNGSLL VIEWENAQIW 1021 GIPDTSFNFQ IWYDFAFDTI HFVYGNMDSP QYATVAGFEN PTGTAGFTLG AIZSAGAQGT 1081 LPVSGDEFIL AESPGDEIEF SYNGTVVPSS EVNDNLTVN EDASVTANIL ANEMESTIN 1141 TFTMESLSGD FRTFTPITID KAPLDPSTVE VTTEPANGTV TVNEDGTVTY APNTNFFGED 1201 SFTYTVRVEG TIDENEEPVE GDIVGEGTVT VAVAGVQDAP VISISAPTSV DEGESYTVTA 1261 SATDPDGDDV TITVNGMATS SITDTAPSHE QARQVTVEVT ATDGIDTTTE TVTIRVNDTS 1321 GGGSMGWIAL LLAPAVYLRR RMKRS	4%
2	Substilase family serine protease [Alteromonas macleodii ATCC 27126]	AFS37436.1	130,116	195	1 MAQFSAVNPS VSSMAAHKQQ KVNFKSNAAK EYQSFLEAQQ QSVIQSIKSF DKSFKADMSY 61 TAAFNGFAGI VSKSALDKLS SLSTVKAVYP DLVHHAQMDA SLDLIGAVQT WEQFGGKESA 121 GAGVKVAIID SGIRPENPLF SGENFTAPDA ATLPTDDYCS EVPDFCNNKL IVARAADIVE 181 GFSVVEEEYE SPLGFNGHGT HVAGTAVGNY GVMAERDGAE AEISGVAPAA YLMVYKGLYA 241 TPANPASSSG ASSMLLSMLE AALTDGADIV NNSWGGGAGG NPNGSVYEDV LEAMHDAGVV 301 TVFAAGNDGP GAGTIGCPGC SEDVITVANT TTGRLFANEV TVEGDTTLGS IPALYSVGNP 361 AIVFDSPITA PVVYAGEVDA ANVEGCDAFA AGAFDGAIAL ISRGTCGFVT KIENAEAAGA 421 TAVLIHNVDG RGEAPILMGG LSEAQTIPSL MLPATPGLAL ANLAETTEES LNVTIGSDIV 481 KVVSDSLADI MNESSSRGPN GDPTFLKPNI AAPGTRIPSG ESPDAPGHED QNFSFKNGTS 541 MASPHVAGAA ALLKQMHPIW TAQQIKSALV TSSIRDILKE DASTAADNFD MGAGRLDLPR 601 ATTVELTYSD LSLVDGNCYL NCEMSITVTN TSDEEVTVDA TAMFNDSAIS ATVTPQMATL 661 AAGAGAEIMV AVDVTPASTG SWSFGGINWA DTDDTTTDYF IPVAIYPISS DEPSIFDSIV 721 STQVAAEGEL VVTEAFATNT DVTGMSITG EIDHKYAINP STISAVKNGN QEPVSYDADS 781 GTVYWEGAIN TASFSMNADT TIGDILEGFG LDRYVPMASI GVAPLECSGS CDDASIEVSL 841 PRTITYLGSE YTTMQISSNG FVSLGTSSGN ITTPPAEVLP SLSEPNNIVA PYWTDFDLAG 901 GGAGAGNIYA VSLTTGHFVV EWENAQIWNE DGTSFNFQLW FEYETGNVNF VYGPMDSPVY 961 ATVVGAENIS GEVGTTYAAL TSAGLQGTLP VEGDEFVLTA NAGDEVTISV AGTVVPSSEY 1021 MDDMLSVNED ASVTANILAN EMDSTIINKF TMDSLSGEFR TTTPTITDKA PLDPSTVVIS 1081 GEPANGTVTV NDDGTVTYAP NTDFFGEDSF TYTVRVEGTI DENEEPVEGE IVGEGTVTVA 1141 VAGVQDAPVI SISAPTSVDE GQSYTVTASA TEADGDDVTI TINGVERTSF TDTAPSNEQS 1201 NRVTVQVTAS DGIETTTESV TIRVNDKSGG SMGWIALLLA PAVYLRRRMK RS	2%
3	Alkaline serine protease [*Kangiella koreensis*]	WP_015780787.1	64,283	54	1 MKLNQIVAGI LLATGVTGAY AGEFKTASVP SKAIKQQYIV VLKDDAVAQN MGQFSSNASE 61 QAIQMMTENL SRKYQAQVQR TVTKALKGGT FTLSEKAAQK LAQDPNVLLV EEDQIVSLNA 121 TQNNATWGID RVDQRNLPLS GTYTYNTTAS NVNAYIIDTG ILNSHSEFGG RSVSGIDTVD 181 NDNDATDCNG HGTHVAGTVG GSTYGIAKSV NLIGVRVLNC SGSGTLSGVI DGIDWVASNH 241 TKPAVANMSL GGGASSSIDT AVANLVSSGV TVVVAAGNDN SSACNYSPAR EPSAITVGST 301 TSSDSRSSFS NYGSCLDIYA PGSSITSAWS NGGTNTISGT SMASPHVAGV AALYLADNPS 361 ATPSQVDTAI TDSATPGVVS DAKSGSPNLL LYSLPDGTNP DPDPDPEPGN ELENGHSVTF 421 SGAQYSETDF TFEVPSAASN VSFDMSGGSG DADLYVKFGS APTTSSYDCR PYRNGNNESC 481 DFSAQTGTYY VMVRGYTSYS NVSLVATHDG DGSEPPAEGG SETYPNLSGA SGSWTHYYVD 541 IPAGMSSLDV QMSGGSGDAD LYVRRGAQPT TSSYDCRPYR WGNDETCSFS NPAQDRWYIS 601 IRGYSSYSGV TLQIDWQ	3%

Shared amino acids are shown as bold and underline characters.

Cyanidiophyceae consist of 6% prokaryotic DNA. However, to date, there have been no reports on the phylogenetic relationship and gene transfer of hydrolytic enzymes between bacteria and algae. Therefore, the finding of this research has opened new research opportunities for investigation into algal lineage and their evolutionary relationship with bacteria.

4. Conclusion

This study evaluated the amino acid sequence of proteolytic enzymes from the red seaweed *G. edulis*.

LC-MS/MS analysis of the two protease bands excised from SDS-PAGE gels showed that the enzymes belong to the metalloprotease and serine protease families, which have not been previously reported for the algal species. However, none of the peptides were related to the proteases in the existing protein databases for the genera *Gracilaria* and *Hydropuntia*, indicating that the current protein databases for the algae are probably incomplete. Rather, more than 80% of the peptide sequences matched with those from members of the bacteria kingdom, including *Gallaecimonas* and *Alteromonas*. These results demonstrate a strong relationship between algae and bacteria, which was probably caused by evolutionary and environmental forces, such as bacterial colonisation of plants, their subsequent symbiosis and horizontal gene transfer. However, further studies are needed to confirm (or refute) the hypothesis and examine how such gene transfer may have occurred. For the practical aspect, the utilisation of the milk-clotting enzymes for cheese-making application should be investigated.

CRediT authorship contribution statement

Ariestya Arlene Arbita: Methodology, Validation, Formal analysis, Investigation, Visualisation, Project administration, Writing-Original draft preparation. **Nicholas A. Paul**: Resources, Supervision, Writing-Reviewing and editing. **Julian Cox**: Supervision, Reviewing. **Jian Zhao**: Conceptualisation, Methodology, Validation, Supervision, Writing-Reviewing.

Declaration of competing interest

The authors declare that there is no conflict of interest in this study.

Acknowledgement

AA Arbita thanks the Indonesia Endowment Fund for Education (LPDP), the Republic of Indonesia, for financing her PhD scholarship. We thank C Morris from the University of the South Pacific for the supply of the dried seaweed from Fiji and S. Tiitii from the Ministry of Agriculture and Fisheries for the supply of the dried seaweed from Samoa under the ACIAR Project FIS/2010/098. We would like to thank Dr Ling Zhong for her help in liquid chromatography and mass spectrometry analysis. Mass spectrometric analysis for this work was carried out at the Bioanalytical Mass Spectrometry Facility, UNSW, and was supported in part by infrastructure funding from the New South Wales Government as part of its co-investment in the National Collaborative Research Infrastructure Strategy.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijbiomac.2022.05.038.

References

[1] F.D. Nicosia, I. Puglisi, A. Pino, C. Caggia, C.L. Randazzo, in: Plant Milk-clotting Enzymes for Cheesemaking, 2022, pp. 1–16.

[2] A. Andrén, in: A.L. Kelly, L.B. Larsen (Eds.), Milk-Clotting Enzymes - Agents of Change: Enzymes in Milk and Dairy Products, Springer International Publishing, Cham, 2021, pp. 349–362, https://doi.org/10.1007/978-3-030-55482-8_14.

[3] M. Mozzon, R. Foligni, C. Mannozzi, F. Zamporlini, N. Raffaelli, L. Aquilanti, Clotting properties of Onopordum tauricum (willd.) aqueous extract in milk of different species, Foods 9 (2020) 1–15, https://doi.org/10.3390/foods9060692.

[4] Y. Zhang, T. Geary, B.K. Simpson, Genetically modified food enzymes: a review, Curr. Opin. Food Sci. 25 (2019) 14–18, https://doi.org/10.1016/j.cofs.2019.01.002.

[5] A. Rojas, R.F. Doolittle, The occurrence of type S1A serine proteases in sponge and jellyfish, J. Mol. Evol. 55 (2002) 790–794, https://doi.org/10.1007/s00239-002-2361-x.

[6] S. Sharmila, L.J. Rebecca, G. Susithra, V. Dhanalakshmi, Partial purification of protease from seaweed, Int. J. Appl. Biotechnol. Biochem. 2 (2012) 81–85. http://www.ripublication.com/ijabb.htm.

[7] A.A. Arbita, N.A. Paul, J. Cox, J. Zhao, Extraction, partial purification and characterization of proteases from the red seaweed Gracilaria edulis with similar cleavage sites on κ-casein as calf rennet, Food Chem. 330 (2020), 127324, https://doi.org/10.1016/j.foodchem.2020.127324.

[8] M.D.N. Meinita, N. Akromah, N. Andriyani, D.Harwanto Setijanto, T. Liu, Molecular identification of Gracilaria species (Gracilariales, rhodophyta) obtained from the South Coast of Java Island, Indonesia, Biodiversitas 22 (2021) 3046–3056, https://doi.org/10.13057/biodiv/d220759.

[9] S.S. Mahusook, F.A.A. Rajathi, H.N.S. Maharifa, R. Sharmila, Comparative study of agarophytes-Gracilaria edulis and Gelidiella acerosa as biostimulant and application of agar for water-holding in soil and plant growth promotion, Agric. Sci. Dig. 41 (2021) 21–27, https://doi.org/10.18805/ag.D-5108.

[10] R. Sakhthivel, K.P. Devi, Evaluation of physicochemical properties, proximate and nutritional composition of Gracilaria edulis collected from Palk Bay, Food Chem. 174 (2015) 68–74.

[11] J. Debbarma, B. Madhusudana Rao, L. Narasimha Murthy, S. Mathew, G. Venkateshwarlu, C.N. Ravishankar, Nutritional profiling of the edible seaweeds Gracilaria edulis, Ulva lactuca and sargassum sp, Indian J. Fish. 63 (2016) 81–87, https://doi.org/10.21077/ijf.2016.63.3.60073-11.

[12] M. Seethalakshmi, Alpha Amylase and Alpha Glucosidase Inhibition Activity of Selected Edible Seaweeds From South Coast Area of India 9, 2017.

[13] A. Arulkumar, T. Rosemary, S. Paramasivam, R.B. Rajendran, Phytochemical composition, in vitro antioxidant, antibacterial potential and GC-MS analysis of red seaweeds (Gracilaria corticata and Gracilaria edulis) from Palk Bay, India, Biocatal. Agric. Biotechnol. 15 (2018) 63–71, https://doi.org/10.1016/j.bcab.2018.05.008.

[14] T.L. Gunathilaka, K.W. Samarakoon, P. Ranasinghe, L.C.D. Peiris, In-vitro antioxidant, hypoglycemic activity, and identification of bioactive compounds in phenol-rich extract from the marine red algae Gracilaria edulis (Gmelin) silva, Molecules 24 (2019) 1–17, https://doi.org/10.3390/molecules24203708.

[15] T. Suthiphongchai, P. Boonsiri, B. Panijpan, Vanadium-dependent bromoperoxidases from Gracilaria algae, J. Appl. Phycol. 20 (2008) 271–278, https://doi.org/10.1007/s10811-007-9243-y.

[16] T. Liu, X. Tang, X. Jia, X. Wu, M. Huang, J. Zeng, The complete plastid genome and phylogenetic analysis of Gracilaria edulis, Mitochondrial DNA Part B 4 (2019) 2598–2599, https://doi.org/10.1080/23802359.2019.1642161.

[17] R. Eisenthal, M.J. Danson, Enzyme Assays. http://opus.bath.ac.uk/id/eprint/4231, 2002.

[18] K.C. Duong-Ly, S.B. Gabelli, Salting out of Proteins Using Ammonium Sulfate Precipitation, 1st ed., Elsevier Inc., 2014 https://doi.org/10.1016/B978-0-12-420119-4.00007-0.

[19] A.T. Phillips, M.W. Signs, Desalting, concentration, and buffer exchange by dialysis and ultrafiltration, Curr. Protoc. Protein Sci. 38 (2004) 1–15, https://doi.org/10.1002/0471140864.ps0404s38.

[20] W. Wang, Q.J. Liu, H. Cui, Rapid desalting and protein recovery with phenol after ammonium sulfate fractionation, Electrophoresis 28 (2007) 2358–2360, https://doi.org/10.1002/elps.200600743.

[21] K. Hafid, J. John, T.M. Sayah, R. Domínguez, S. Becila, M. Lamri, A.L. Dib, J. M. Lorenzo, M. Gagaoua, One-step recovery of latex papain from Carica papaya using three phase partitioning and its use as milk-clotting and meat-tenderizing agent, Int. J. Biol. Macromol. 146 (2020) 798–810, https://doi.org/10.1016/j.ijbiomac.2019.10.048.

[22] U. Andjelković, D. Josic, Mass spectrometry based proteomics as foodomics tool in research and assurance of food quality and safety, Trends Food Sci. Technol. 77 (2018) 100–119, https://doi.org/10.1016/j.tifs.2018.04.008.

[23] Q. Sun, B. Zhang, Q.-J. Yan, Z.-Q. Jiang, Comparative analysis on the distribution of protease activities among fruits and vegetable resources, Food Chem. 213 (2016) 708–713, https://doi.org/10.1016/j.foodchem.2016.07.029.

[24] D.J.C. Pappin, P. Hojrup, A.J. Bleasby, Rapid identification of proteins by peptide-mass fingerprinting, Curr. Biol. 3 (1993) 327–332, https://doi.org/10.1016/0960-9822(93)90195-T.

[25] J.C. Hagopian, M. Reis, J.P. Kitajima, D. Bhattacharya, M.C. de Oliveira, J. C. Hagopian, Comparative analysis of the complete plastid genome sequence of the red alga Gracilaria tenuistipitata var. liui provides insights into the evolution of rhodoplasts and their relationship to other plastids, J. Mol. Evol. (2004) 464–477, https://doi.org/10.1007/s00239-004-2638-3.

[26] M.A. Campbell, G. Presting, M.S. Bennett, A.R. Sherwood, Highly conserved organellar genomes in the Gracilariales as inferred using new data from the Hawaiian invasive alga Gracilaria salicornia (Rhodophyta), Phycologia 53 (2014) 109–116, https://doi.org/10.2216/13-222.1.

[27] R.-S. Siow, S. Teoh, S.-S. Teo, S.-M. Phang, C.-L. Ho, Molecular cloning and characterization of GDP-mannose-3′,5′- epimerase from Gracilaria changii, J. Appl. Phicology 25 (2013) 1309–1318, https://doi.org/10.1007/s10811-013-9987-5.

[28] A.O. Lluisma, M.A. Ragan, Characterization of a galactose-1-phosphate uridylyltransferase gene from the marine red alga Graeilaria gracilis, Curr. Genet. 34 (1998) 112–119, https://doi.org/10.1007/s002940050374.

[29] A.C. Parte, LPSN - list of prokaryotic names with standing in nomenclature, Nucleic Acids Res. 42 (2014) 613–616, https://doi.org/10.1093/nar/gkt1111.

[30] J.P. Euzéby, List of bacterial names with standing in nomenclature: a folder available on the internet, Int. J. Syst. Bacteriol. 47 (1997) 590–592, https://doi.org/10.1099/00207713-47-2-590.

[31] K. Pruitt, T. Murphy, G. Brown, M. Murphy, RefSeq Help, National Center for Biotechnology Information (NCBI), 2011.

[32] J.P. Kayser, J.L. Vallet, R.I. Cerny, Defining parameters for homology-tolerant database searching, J. Biomol. Tech. 15 (2004) 285–295.

[33] R. Majumder, S.P. Banik, S. Khowala, Purification and characterisation of κ-casein specific milk-clotting metalloprotease from Termitomyces clypeatus MTCC 5091, Food Chem. 173 (2015) 441–448, https://doi.org/10.1016/j.foodchem.2014.10.027.

[34] M. Salehi, M.R. Aghamaali, R.H. Sajedi, S.M. Asghari, E. Jorjani, Purification and characterization of a milk-clotting aspartic protease from Withania coagulans fruit, Int. J. Biol. Macromol. 98 (2017) 847–854, https://doi.org/10.1016/j.ijbiomac.2017.02.034.

[35] A. Ben Amira, J. Bauwens, E. De Pauw, S. Besbes, H. Attia, F. Francis, C. Blecker, Identification of proteins from wild cardoon flowers (Cynara cardunculus L.) by a proteomic approach, J. Chem. Biol. (2017) 25–33, https://doi.org/10.1007/s12154-016-0161-9.

[36] J.A. Rivett, High molecular mass intracellular proteases, Biochem. J. 263 (1989) 625–633, https://doi.org/10.1042/bj2630625.

[37] F. Hang, Q. Wang, Q. Hong, P. Liu, Z. Wu, Z. Liu, H. Zhang, W. Chen, Purification and characterization of a novel milk-clotting metalloproteinase from Paenibacillus spp. BD3526, Int. J. Biol. Macromol. 85 (2016) 547–554, https://doi.org/10.1016/j.ijbiomac.2016.01.028.

[38] R.S.C. de Souza, J.S.L. Armanhi, N.de B. Damasceno, J. Imperial, P. Arruda, Genome sequences of a plant beneficial synthetic bacterial community reveal genetic features for successful plant colonization, Front. Microbiol. 10 (2019), https://doi.org/10.3389/fmicb.2019.01779.

[39] G.M. Preston, B. Haubold, P.B. Rainey, Bacterial genomics and adaptation to life on plants: implications for the evolution of pathogenicity and symbiosis, Curr. Opin. Microbiol. 1 (1998) 589–597.

[40] N.R. Pace, A molecular view of microbial diversity and the biosphere, Science (80-.) 276 (1997) 734–740, https://doi.org/10.1126/science.276.5313.734.

[41] N. Nikolaidis, N. Doran, D.J. Cosgrove, in: Plant Expansins in Bacteria and Fungi : Evolution by Horizontal Gene Transfer and Independent Domain Fusion 31, 2013, pp. 376–386, https://doi.org/10.1093/molbev/mst206.

[42] R.M. Givens, M.H. Lin, D.J. Taylor, U. Mechold, J.O. Berry, V.J. Hernandez, Inducible expression, enzymatic activity, and origin of higher plant homologues of bacrterial RelA/SpoT stress proteins in Nicotiana tabacum, J. Biol. Chem. 279 (2004) 7495–7504, https://doi.org/10.1074/jbc.M311573200.

[43] G. Sandmann, Molecular evolution of carotenoid biosynthesis from bacteria to plants, Physiol. Plant. 116 (2002) 431–440, https://doi.org/10.1034/j.1399-3054.2002.1160401.x.

[44] R. Ramanan, B. Kim, D. Cho, H. Oh, H. Kim, Algae – bacteria interactions: evolution, ecology and emerging applications, Biotechnol. Adv. 34 (2016) 14–29, https://doi.org/10.1016/j.biotechadv.2015.12.003.

[45] G. Schönknecht, W.H. Chen, C.M. Ternes, G.G. Barbier, R.P. Shrestha, M. Stanke, A. Bräutigam, B.J. Baker, J.F. Banfield, R.M. Garavito, K. Carr, C. Wilkerson, S. A. Rensing, D. Gagneul, N.E. Dickenson, C. Oesterhelt, M.J. Lercher, A.P.M. Weber, Gene transfer from bacteria and archaea facilitated evolution of an extremophilic eukaryote, Science (80-.) 339 (2013) 1207–1210, https://doi.org/10.1126/science.1231707.

[46] E. Pennisi, Algae suggest eukaryotes get many gifts of bacteria DNA, Science (80-.) 363 (2019) 439–440, https://doi.org/10.1126/science.363.6426.439-b.

Amino acid sequence of two new milk-clotting proteases from the macroalga Gracilaria edulis

ORIGINALITY REPORT

18%	5%	17%	%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Ariestya Arlene Arbita, Nicholas A. Paul, Julian Cox, Jian Zhao. "Extraction, partial purification and characterization of proteases from the red seaweed Gracilaria edulis with similar cleavage sites on κ-casein as calf rennet", Food Chemistry, 2020 Publication	14%
2	unsworks.unsw.edu.au Internet Source	4%

Exclude quotes	Off	Exclude matches	< 3%
Exclude bibliography	On		