Cloning and heterologous expression of subtilisin SAPN, a serine alkaline protease from Melghiribacillus thermohalophilus Nari2AT in Escherichia coli and Pichia pastoris

Sondes Mechri, Nadia Zaraî Jaouadi, Khelifa Bouacem, Fawzi Allala, Aicha Bouraoui, Céline Ferard, Hatem Rekik, Alexandre Noiriel, Abdelkarim Abousalham, Amel Bouanane-Darenfed, et al.

Cloning and heterologous expression of subtilisin SAPN, a serine alkaline protease from *Melghiribacillus thermohalophilus* Nari2A^T in *Escherichia coli* and *Pichia pastoris*

Sondes Mechri [a], Nadia Zaraî Jaouadi [a], Khelifa Bouacem [b], Fawzi Allala [b], Aicha Bouraoui [c], Céline Ferard [c], Hatem Rekik [a], Alexandre Noiriel [d], Abdelkarim Abousalham [d], Amel Bouanane-Darenfed [b], Hocine Hacène [b], Florence Lederer [c], Laura Baciou [c, **], Bassem Jaouadi [a, *]

[a] *Laboratoire de Biotechnologie Microbienne, Enzymatique et de Biomolécules (LBMEB), Centre de Biotechnologie de Sfax (CBS), Université de Sfax, Route Sidi Mansour Km 6, BP 1177, Sfax 3018, Tunisia*

[b] *Laboratoire de Biologie Cellulaire et Moléculaire (LBCM), Equipe de Microbiologie, Faculté des Sciences Biologiques (FSB), Université des Sciences et de la Technologie Houari Boumediene (USTHB), BP 32, El Alia, Bab Ezzouar, Alger 16111, Algeria*

[c] *Institut de Chimie Physique (ICP), UMR 8000, CNRS, Université Paris Saclay, Bat. 350, 15, Avenue Jean Perrin, 91405 Orsay Cedex, France*

[d] *Univ Lyon, Université Lyon 1, Institut de Chimie et de Biochimie Moléculaires et Supramoléculaires (ICBMS), UMR 5246 CNRS, Génie Enzymatique, Membranes Biomimétiques et Assemblages Supramoléculaires (GEMBAS), Bât Raulin, 43 Bd du 11 Novembre 1918, F-69622 Villeurbanne Cedex, France*

---

*Correspondence to: B. Jaouadi, CBS, Université de Sfax, BP 1177, Sfax 3018, Tunisia

**Corresponding authors.

Email addresses: bassem.jaouadi@cbs.rnrt.tn (B. Jaouadi); laura.baciou@universite-paris-saclay.fr (L. Baciou).

ABSTRACT

The *sapN* gene, encoding the extracellular subtilisin SAPN, a serine alkaline protease from *Melghiribacillus thermohalophilus* Nari2A^T, was isolated, sequenced, and heterologously expressed in *Escherichia coli* BL21(DE3)pLysS using pUT57 and pTrc99A vectors and in *E. coli* BL21-AI™ using the Gateway™-pDEST™17 vector. Conversely, three cassettes encoding pre-pro-subtilisin (rSAPN/SP-Pro-M), pro-subtilisin (rSAPN/Pro-M), and the mature-subtilisin (rSAPN/M) were hyperexpressed in *Pichia pastoris* SMD1168 and X33 using the pPICZαC vector. rSAPNs were purified, characterized, and compared to wild-type SAPN. The deduced amino acid sequences exhibited high similarity with subtilisins from *Bacillus* strains. The highest sequence identity (96%) was observed with the *B. licheniformis* MP1 protease, with a 10-residue difference. Compared to SAPN and untagged rSAPNs, $(His)_6$-tagged enzymes showed the highest activity and stability at alkaline pH and high temperature, the highest hydrolysis degree on crab and shrimp by-products, and the best catalytic efficiency. It was found that His_6-rSAPN/SP-Pro-Ms expressed in *P. pastoris* strains was more active than those produced in *E. coli*. To initiate structure-function relationships, a 3D-model of the Pro-SAPN was built based on the available structures of common subtilisins. These data constitute a pivotal first step toward the creation of new efficient rSAPNs with enhanced catalytic properties and high potential for biotechnological and industrial uses.

***Keywords*:** *Melghiribacillus thermohalophilus*; Subtilisin; Expression; *E. coli*; *P. pastoris*; Comparative modeling.

1. Introduction

Widely studied in protein engineering, subtilisin (initially obtained from *Bacillus subtilis* [1, 2]), an alkaline serine protease, is the main industrial enzyme used as a model for understanding the importance of biocatalyst improvements [3]. It belongs to the α/β-hydrolase fold protein superfamily [4, 5]. The main representatives of bacterial subtilisins are subtilisin E of *B. subtilis* [6], subtilisin Novo or BPN′ of *B. amyloliquefaciens* [7], and subtilisin Carlsberg of *B. licheniformis* [8]. This biocatalyst has the specificity to hydrolyze peptide bonds, esters, and amides [9]. Subtilisins possess a typical nucleophilic serine residue in the canonical catalytic triad "aspartate/histidine/serine" (D32/H63-64/S220-221, numbering from the subtilisin sequences of *B. licheniformis* and *B. amyloliquefaciens*, respectively) [5]. The serine (S22/S221) initiates the hydrolysis through a nucleophilic attack of the polypeptide backbone carbonyl group. The histidine residue (H63/H64) extracts one proton from the serine hydroxyl group, while the aspartate residue (D32) maintains the protonated state of the imidazole group and orients it to the optimal position [10]. Thus far, there are 274 different protease families reported in the MEROPS peptidase database (http://merops.sanger.ac.uk/), mainly organized by amino acid sequence relationships. The subtilisin-like alkaline serine proteases constitute two distinct families that are subtilases S8 and sedolisins S53. In the S8 family, the catalytic residues are ordered as given above. However, in the S53 family, H63/H64 is substituted by a glutamate residue (E63/E64) [11].

Subtilisins are initially expressed as an inactive precursor, called pre-pro-subtilisin (42 kDa), composed of a signal peptide (SP) fused with a propeptide (Pro) at the N-terminus of the protein [12]. The SP serves the purpose of translocation to the extracellular matrix. It is excised during secretion [13]. The protein is then released as a prodomain (34 kDa) and matures in three steps: (i) pro-subtilisin folding, (ii) propeptide autoprocessing, and (iii)

propeptide removal [14]. The mature forms of subtilisins E and BPN' consist of 275 amino acids (aa) against 274 aa for subtilisin Carlsberg [5].

The leading factor prompting attention to subtilisins is the industrial utility of this type of enzyme [9, 15, 16]. Both the scientific community and industrial specialists have welcomed attempts to engineer a subtilisin because it would help to meet all industrial requirements. To satisfy these requirements in addition to provide alternative evolutionary properties, subtilisins need to be remodeled through protein engineering. Therefore, it is crucial to study the mechanisms of secretion of these enzymes to improve the expression of the homologous/heterologous proteins by the optimization of vectors and host strains.

While subtilisins from mesophilic bacteria are well documented [17], reports on thermophilic and hyperthermophilic subtilisins are still quite limited. In a previous study, the authors describe the purification and biochemical characterization of a novel acido-halotolerant and thermostable endochitinase (called ChiA-Mt45) [18] and a novel extracellular thermostable and haloalkaliphilic serine alkaline protease (termed as subtilisin SAPN) from the *M. thermohalophilus* strain Nari2A^{T} [19]. Its catalytic activity and its stability at alkaline pH and at high temperatures are fundamental for the industrial bioprocesses that involve high pH and temperatures. The genus *Melghiribacillus* is a newly identified member of the *Bacillaceae* family, which integrates the *Bacilli* class as of 2015 [20]. To date, only the species *M. thermohalophilus* Nari2A^{T} (= DSM 25894^{T} = CCUG 62543^{T}) is accessible in the roster of prokaryote names (http://www.bacterio.net/melghiribacillus.html). Additionally, the genome sequence of the *M. thermohalophilus* strain DSM 25894^{T} demonstrated the involvement of many serine proteolytic enzymes (3 proteases and 7 peptidases) as recorded in GenBank (accession no. SMAN00000000). This study constitutes the first molecular investigation targeting the subtilisin synthetized by this new taxon isolated from the Saharan soil. Based on the existing

knowledge, isolation, cloning, and sequencing of the *sapN* encoding gene expressed in *E. coli* and *P. pastoris* strains have been carried out in this work. The expression levels of secreted proteins have been evaluated and optimized; biochemical properties of the produced recombinant enzymes were investigated and compared. Additionally, modeling the 3D structure of subtilisin SAPN has been performed to improve the understanding of the structure-function relationships of the SAPN enzyme.

2. Materials and methods

2.1. Materials

The Wizard® genomic DNA purification kit used was from Promega (Madison, WI, USA). The oligonucleotides used in this work (listed in Supplementary data, Table S1) were from Eurogentec Corporation (Angers, France). The kit and reagents for DNA purification, Phusion® High-fidelity DNA Polymerase (NEB, Ipswich, UK), T4 DNA ligase, and restriction enzymes (EcoRI, HindIII, ClaI [Bsu15], XbaI, and PmeI [MssI]) were purchased from Thermo Scientific™ (Rockford, IL, USA). The LR Clonase® II Enzyme Mix was from Thermo Fisher Scientific, Inc., (Waltham, MA, USA). The GeneJET Genomic DNA Purification kit and DNA GeneJET Plasmid Miniprep kit were supplied by Fisher Scientific SAS (Illkirch-Graffenstaden, France). E.Z.N.A® Yeast DNA isolation kit was from Omega Bio-tek (Norcross, GA, USA). The T100™ Thermal Cycler and DC protein analysis kit, for Bradford-based assay, were purchased from Bio-Rad Laboratories Inc. (Hercules, CA, USA). The ABI DNA Genetic Analyzer Automated Sequencer 3100 was supplied by ABI (PRISM®, USA). The pH-STAT 907 Titrando was from Metrohm AG (Herisau, Switzerland). The mouse-His-probe (H-3) antibody and the anti-mouse IgG-HRP (Horseradish peroxidase; Clontech) were purchased from Santa Cruz Biotechnology, Inc. (CA, USA). The HisTrap™

HP chelating Ni-affinity chromatography column was purchased from GE Healthcare Bio-Sciences AB (Uppsala, Sweden).

Hammerstein casein was provided by Merck (Darmstadt, Germany). Kanamycin sulfate and ampicillin sodium salt USP were purchased from Bio Basic (Markham, Canada Inc.). Isopropyl-thio-β-D-galactopyranoside (IPTG) and lysozyme were bought from Sigma-Aldrich Co (St Louis, MO, USA).

The culture media used were Luria-Bertani (LB) and yeast extract peptone dextrose (YPD), buffered glycerol-complex (BMGY), and buffered methanol-complex (BMMY). The YPD media were composed of yeast extract, 10 g; peptone, 20 g; and dextrose, 20 g per liter. The LB and YPD were complemented with 20 g/L agar to make solid media. The BMGY/BMMY medium was composed of yeast extract, 1%; peptone, 2%; biotin, 4×10^{-5}% (*w*/*v*); and yeast nitrogen base, 1.34% in 100 mM potassium phosphate pH 6 plus glycerol 1% (*v*/*v*) (for BMGY) or methanol 1% (*v*/*v*) (for BMMY). The buffers used were **A**: 25 mM Tris-HCl and 50 mM NaCl at pH 8, **B**: 20 mM sodium phosphate, and 500 mM NaCl at pH 7.4, **C**: 100 mM glycine-NaOH and 2 mM $CaCl_2$ at pH 10, **D**: Tris-buffered saline (TBST) and Tween 20 solution [50 mM Tris-HCl, 150 mM NaCl, and Tween 20 at 0.1% (*v*/*v*] at pH 7.5, and **E:** lysis-buffer [Tris-HCl (30 mM, pH 8), EDTA (1 mM), and sucrose (200 g/L)].

2.2. Vectors and strains

The pCR-Blunt™, pDONR™ 221, pDEST™17, and pENTR™/D-TOPO® vectors and the EasySelect™ *Pichia* Expression kit were purchased from Invitrogen (Carlsbad, CA, USA). The pUT57 derivative from pUC19 and pTrc99A was from CAYLA (Toulouse, France) and Addgene (LGC Standards, Teddington, UK), respectively.

Three *E. coli* strains TOP10, BL21(DE3)pLysS, and BL21-AI™ were obtained from Thermo Fisher Scientific, Inc., (Waltham, MA, USA) and the two *P. pastoris* strains X33 and

147 SMD1168 were obtained from Invitrogen (Carlsbad, CA, USA). The wild-type extracellular
148 subtilisin SAPN isolated from Nari2A^T strain was used as the reference protease [19].

149 ***2.3. PCR protocols***

150 The DNA protocols were executed as illustrated in Sambrook et al. [21]. The *sapN* gene,
151 isolated using the GeneJET Genomic DNA Purification kit, was PCR-amplified by mixing
152 ~200 ng genomic DNA (2 μL) in a thermocycler microtube, including 5× Phusion HF buffer,
153 forward and reverse oligonucleotides (10 μmol), Phusion® High-Fidelity DNA Polymerase
154 (0.2 μL), and dNTPs (10 mM). The reaction mixtures (50 μL) in a T100™ Thermal Cycler
155 were denatured at 94°C for 5 min and underwent 35 cycles of: denaturation at 94°C for 40 s,
156 oligonucleotide annealing at the hybridization temperature of each pair of primers for 30 s,
157 and extension at 72°C for 90 s. This was followed by an extension phase at 72°C for 7 min.
158 The PCR amplifications were subsequently extracted through an appropriate DNA
159 purification kit.

160 ***2.4. sapN gene fishing and cloning protocols***

161 The F-MS15 and R-MS16 primers were elaborated on the basis of the high sequence
162 similarity found in *Bacillus* and homologous genes. They were used to isolate the *sapN* gene
163 with its neighboring regions by amplifying a ~1.4 kb PCR segment from the Nari2A^T strain
164 genomic DNA (Supplementary data, Table S1). The PCR products were then cloned with
165 pCR-Blunt™ into *E. coli* TOP10, using the instructions of the manufacturer (Fig. 1).
166 Recombinant clones were grown in LB agar medium supplemented with 100 μg/mL
167 kanamycin-sulfate, 160 μg/mL X-Gal, and 360 μg/mL IPTG. The colonies were screened by
168 restriction digestion and by performing PCR amplification using F-MS15 and R-MS16
169 primers (Supplementary data, Table S1). Two oligonucleotides (F-MS17 and R-MS18) for
170 nested PCR were used for sequencing the *sapN* gene inner region of ~0.5 kb. Plasmids

harboring DNA inserts of predictable lengths were isolated using the DNA GeneJET Plasmid Miniprep kit. A recombinant plasmid, termed pSM30, containing the open reading frame (ORF) encoding the pre-pro-enzyme SAPN was used for extensive sequencing.

2.5. Expression in E. coli

2.5.1. Expression in BL21(DE3)pLysS E. coli strain

The ORF region that encodes SAPN was amplified with F-MS19 and R-MS20 primers by using pSM30 or strain Nari2A^T genomic DNA (Fig. 1). The PCR fragment containing the *sapN* gene was then double-restricted with EcoRI and HindIII and inserted into the pUT57 and pTrc99A vectors leading to pSM31 and pSM32 plasmids, respectively, containing the pre-pro-enzyme SAPN (Fig. 1). The pSM31 and pSM32 ligations served to transform *E. coli* strain BL21(DE3)pLysS (Table 1). Afterward, two transformants were selected on LB agar with 100 μg/mL ampicillin, 160 μg/mL X-Gal, and 360 μg/mL IPTG. The constructions were verified by sequencing (Fig. 2).

After reaching an absorbance ($A_{600\ nm}$) of 0.8 units, the expression of rSAPN BL21(DE3)pLysS was induced by the addition of IPTG (0.4 mM). Next, *E. coli* BL21(DE3)pLysS/pSM32 was cultivated for 14 h (200 rpm and 37°C) and centrifuged. The culture supernatant was used for rSAPN BL21(DE3)pLysS purification, using the same protocol as for the native enzyme [19].

2.5.2. Expression in BL21-AI™ E. coli strain

For the Gateway™ cloning technology, the ORF region encoding SAPN was amplified by PCR using the F-MS21 and R-MS22 oligonucleotides containing *att*B attachment sites for recombination (Supplementary data, Table S1). The construct pSM30 or the genomic DNA of strain Nari2A^T was used as a template. The amplified products were purified and cloned in the

pDONR™ 221 plasmid in frame with an N-terminal hexahistidine tag by a BP reaction recognized as recombination between the *att*B sequence (PCR product) and *att*P sequence (pDONR™ 221 vector) (Fig. 3A). The pDONR™ 221 vector was used to create *att*L-fringed entry clones, including the *sapN* gene ensuing by recombination with an *att*B PCR product. Then, the ORF encoding SAPN was transferred into pDEST™17 through LR-reaction recognized as recombination between the *att*L sequence (recombinant pENTR™/D-TOPO® vector) and the *att*R sequence (pDEST™17 vector). Subsequently, the recombinant pDEST™17 was transferred to the *E. coli* strain BL21-AI™ leading to the pSM33 plasmid containing the pre-pro-enzyme SAPN. In the next step, the recombinant cells were multiplied in LB broth complemented with 100 μg/mL ampicillin. After reaching a mid-log phase with $A_{600\ nm}$ of 0.4 units, the expression of the recombinant protein (His_6-rSAPN BL21-AI^{TM}) was induced by adding L-arabinose at 2 g/L. Next, *E. coli* BL21-AI™/pSM33 was cultivated for 12 h (200 rpm and 37°C). After centrifugation (10,000×*g*, 30 min, and 4°C), the supernatant was mixed with 5 mL Ni^{2+}-NTA resin in buffer A. The mixture was loaded onto the Ni-affinity column and rinsed with 200 mL buffer A. The adsorbed proteins were eluted using a linear imidazole gradient from 0 to 500 mM in buffer A. The elution profile was monitored at 280 nm for protein concentration and at 660 nm for subtilisin activity. The fractions, containing His_6-rSAPN BL21-AI^{TM} activity, were combined and concentrated (if required) for further investigation.

2.6. Expression in P. pastoris

2.6.1. Design of sapN gene constructions

Three cassettes of *sapN* gene were designed and transformed in *P. pastoris* strains: X33 and SMD1168 (Fig. 3B). These constructs include the *sapN* gene encoding SAPN with a SP and a propeptide (rSAPN/SP-Pro-M), SAPN without its SP (rSAPN/Pro-M), and SAPN free

of both peptides (rSAPN/M). The forward primers with the ClaI restriction site were F-MS23, F-MS24, and F-MS25 (Supplementary data, Table S1). The reverse primer R-MS26 having an XbaI restriction site was selected to generate recombinant C-terminal His-tagged subtilisins. The PCR product of each cassette was digested with ClaI and XbaI, purified, and inserted into the pPICZαC vector. The three ligations were used for the transformation of *E. coli* using competent One Shot® TOP10 cells. Recombinants were selected on LB agar in the presence of 25 μg/mL zeocin. For the validation of DNA insertions, the transformants were checked through direct PCR amplification of colonies and by digestion with ClaI and XbaI. The integrity of each cassette in the different pPICZαC recombinant plasmids pSM34, pSM35, and pSM36 (containing SP-Pro-M, Pro-M, and M, respectively) was validated by DNA sequencing.

2.6.2. Electroporation and selection processes of P. pastoris expression strains

Three pSM34, pSM35, and pSM36 plasmids (3 μg) were linearized with PmeI [MssI], purified with the extraction kit and then transformed into electrocompetent *P. pastoris* X33 and SMD1168 strains. Practically, competent cells (80 μL) were mixed with the corresponding linearized plasmid (10 μL) in an electroporation cuvette of a Bio-Rad Gene Pulser and subjected to a 1.5 KV frequency for 5 ms. Immediately, 500 μL of cold sterile 1 M sorbitol and 500 μL of YPD were added. The mixture was incubated at 30°C for 1 h. Afterward, 150 μL of each transformed *P. pastoris* strain was spread on YPD agar plates containing an increasing amount of zeocin (100 μg/mL, 500 μg/mL, and 1,000 μg/mL). The plates were kept at 30°C for 2-3 days until colonies appeared. The integration of the cassettes into the genomes of both *P. pastoris* strains, X33 and SMD1168, was checked by PCR using the 5′ and 3′ alcohol oxidase operon 1 (AOX1) oligonucleotides.

2.6.3. Heterologous SAPN expression in P. pastoris and screening of subtilisin-producing clones

The expression of the three forms of subtilisin SAPN in *P. pastoris* was performed following methanol induction [0.5% (*v*/*v*)]. Briefly, the growth occurred in two phases. For the first one, an individual transformed colony was used to inoculate 25 mL of BMGY medium in a baffled flask (250 mL) cultivated overnight (30°C and 250 rpm) up to an absorbance $A_{600\ nm}$ from 2 to 6 units. The cells were then recovered by centrifugation (1,500×*g* for 5 min at 23±2°C) and diluted in 500 mL of BMMY medium to an $A_{600\ nm}$ of 1. Methanol [0.5% (*v*/*v*)] was added every 24 h to induce the strong AOX1 promoter. Aliquots were taken every 6 h to follow the growth and the protease activity for 10 days. For protein purifications, the BMMY cultures of *P. pastoris* X33/pSM34 and *P. pastoris* SMD1168/pSM34 were stopped after a 72-h incubation and centrifuged (1,500×*g* for 20 min). The clear supernatants containing the His_6-rSAPN proteins from X33 and SMD1168 strains were concentrated (if required) using an Amicon membrane (10 kDa). The concentrate was loaded onto the 5 mL HiTrap™ chelating HP column equilibrated with buffer B. Unbound proteins were washed out with buffer B. Adsorbed proteins were eluted with a linear imidazole gradient (0-500 mM) in buffer B. The pooled fractions containing subtilisin activity were lyophilized and then used for enzyme characterization.

2.7. Electrophoresis and Edman degradation

The protein concentration in culture supernatants and chromatographic fractions was measured using the Bradford protocol. The expression of tagged recombinant enzymes (rSAPNs) in pDEST™17 and pPICZαC vectors was analyzed by using western blot. The proteins were precipitated with cold trichloroacetic acid (TCA) and acetone (2 *v*/*v*), heat-treated at 90°C for 10 min in the presence of 100 mM DL-dithiothreitol and separated by

sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). The proteins were then transferred to nitrocellulose membranes blocked for 1 h at 23±2°C with buffer D-BSA (25 g/L). The membrane was kept with His-probe (H-3) (dilution 1: 1500) in buffer D-BSA for 10 min, cleaned twice with buffer D, and incubated with anti-mouse IgG-HRP (dilution 1:15000) in buffer D-BSA. The peroxidase activity was detected by an imaging system (PXi, Syngene) using ECL™ reagents. The rSAPNs sample's purity was analyzed by SDS-PAGE. The proteins were visualized by Coomassie brilliant blue R-250 staining. Alternatively, a zymography technique was performed by incorporating 1 mg/mL azo-casein into the separating gel [22]. The specific rSAPNs bands were sliced and transferred to a polyvinylidene difluoride blotting membrane, and their N-terminal residues were identified by Edman degradation with an ABI Automated Protein Sequencer (Model 473A).

2.8. Subtilisin activity assay

Subtilisin activity was measured through the accumulation of free tyrosine resulting from casein degradation using the Folin-Ciocalteu reagent [23]. A suitably diluted enzyme (0.5 mL) was mixed with 0.5 mL of buffer C containing 10 mg/mL casein and incubated for 15 min at 70°C. Then, 2.5 mL of 200 g/L TCA was added to stop the reaction. Thirty minutes later, the cooled mixture was spun down (15,000 rpm for 20 min). A quantity of 0.5 mL of supernatant was mixed with 2.5 mL of 500 mM Na_2CO_3 and 0.5 mL Folin-Ciocalteu reagent and incubated for 30 min at 23±2°C. One unit of subtilisin activity was defined as the amount of enzyme yielding the equivalent of 1 μmol of tyrosine per minute under the defined assay conditions.

2.9. Biochemical and kinetic characterization of rSAPNs subtilisins

The biochemical properties of subtilisin rSAPNs were investigated as previously described [19]. The effect of pH on the activity was analyzed between pH 3 and 13 with

different buffers (100 mM) at 75°C. The pH stability was assessed by incubating every rSAPN at pH 11 in bicarbonate-NaOH buffer (100 mM + 2 mM $CaCl_2$) for 24 h at 40°C. The optimum temperature for rSAPNs activity was determined from 40°C to 100°C in buffer C. The rSAPNs thermostability was explored by incubation at 80°C and determination of the subtilisin activity at 2 h intervals. For analyzing the combined effect of 2 mM Ca^{2+} and 100 g/L polyethylene glycol (PEG) 1000, rSAPNs were kept for 24 h at 80°C. The rSAPNs activities without additives were considered as the control. Kinetic parameters were calculated by Lineweaver–Burk plots with the Hyper 32 software package, using substrates at variable concentrations (0-70 mM) with 1.5 mg/mL of each enzyme.

The degree of hydrolysis (DH) of proteins in crab and shrimp by-products was determined at 75°C and pH 10 for SAPNs. A quantity of 2 g of each by-product in 100 mL of assay buffer was treated with 2000 U of purified enzymes. The pH was maintained constant with the pH-STAT. The NaOH amount that was needed to maintain the pH constant was proportional to the DH. The enzymatic reactions were stopped when the pH became stable. The DH, defined as the percent ratio of the number of peptide bonds broken (h) to the total number of peptide bonds in the assayed substrate (htot), was calculated in each case from the amount of the base (NaOH) added to keep the pH constant during hydrolysis as previously reported [19].

2.10. Sanger sequencing, in silico tools, and comparative modeling

The DNA sequencing of the *sapN* gene cloned in *E. coli* and *P. pastoris* strains was done at least three times on direct and reverse strands with the ABI DNA Genetic Analyzer Automated Sequencer 3100. The protein sequence alignments were accomplished through the Blast tools associated with Clustal Omega program and web portals (ExPASy, SWISS-PROT, and TrEMBL). The sequence similarities and secondary structure representation were

analyzed with the new version 3.0 of Easy Sequencing in Postscript (ESPript) web server [24]. The functional protein analyses were conducted using the InterPro database and MEROPS, which is the peptidase data base from the European Bioinformatics Institute (EMBL-EBI). Molecular enzymatic signatures and conserved motifs were also recognized through the InterProScan tool by scanning the SAPN protein sequence against the PRINTS database. The rSAPNs N-terminal sequences were compared to those in the UniProtKB and SWISS-PROT protein databases. The 3D modeling of Pro-SAPN was achieved with the SWISS-MODEL protein structure prediction tool. The structure of subtilisin E from *B. subtilis* subsp. *subtilis* str. 168 (PDB ID: 3WHI) was used to construct the model for the Pro-SAPN. The final molecular model was built, analyzed, and visualized by using COACH (a meta-server-based approach to protein-ligand binding site prediction) and PyMOL v0.99 software.

2.11. Statistical analyses

The tests were executed at least 3 times independently, and the statistical analyses were performed. The results concerning the influence of pH, temperature, and DH on subtilisin activity and stability were deemed to be statistically significant at $P<0.05$ and are presented with the standard error (±SE).

3. Results and discussion

3.1. The sapN gene molecular cloning and DNA analysis

The sequencing of the 1.4 kb PCR fragment cloned into pSM30 allowed the identification of the *sapN* gene that comprised an 1140 bp-ORF. The deduced putative SAPN protein sequence is shown in Fig. 2: an alkaline serine protease (subtilisin SAPN) of 379 aa. A putative ribosomal binding site (RBS) of 10 bp - 16 bp upstream of the initiation site (ATG)

was predicted as well as the sequences of putative promoter elements at -35 (ATATTA) and -10 (AATTAT). SAPN has a recognized pre-pro-sequence domain of 105 aa, with a theoretical mass of 39 kDa. It contains a 29 aa-SP (from Met to Ala) showing typical features for subtilisins: a short positive N-terminal sequence (one Arg and two Lys) followed by a highly hydrophobic region of about 20 aa and a C-terminal site (Ala-Ala-Gln) that is cleaved by the signal peptidase during maturation. The propeptide sequence corresponds to the next 76 aa with a theoretical mass of 8 kDa. This prodomain is believed to act as a competitive inhibitor for the enzyme in addition to its primordial role in the correct folding of SAPN. Indeed, once the folding is complete, subtilisin begins intramolecular chaperone degradation to finalize the maturation by subsequent autoproteolysis to generate the active mature subtilisin of 274 aa (A1 to of Q274), with a theoretical mass of 27 kDa.

To improve the expression efficacy of heterologous SAPN, we have studied different secretion systems by using several vectors and host strains (*E. coli* and *P. pastoris* strains) and purification procedures.

3.2. Comparison of expression level and activity of recombinant SAPN in E. coli BL21(DE3)pLysS and BL21-AI™ strains

The pSM31, pSM32, and pSM33 constructs were designed to express in *E. coli* the SP-Pro-M protein. Several *E. coli* strains BL21(DE3)pLysS and BL21-AI™ recombinants were selected (Table 1). The culture growth and the subtilisin activity were monitored for 36 h. The optimum for the production of subtilisin activity was obtained 18 h after IPTG and 12 h after L-arabinose inductions for BL21(DE3)pLysS/pSM32 and *E. coli* BL21-AI™/pSM33, respectively (Fig. 4A). While no subtilisin activity was present intracellularly in the recombinant *E. coli* cells, substantial activity was measured in the culture media. In comparison with the previous results, SAPN activity was higher when the protein expression

was under the control of the strong promoter Ptrc (pSM32) than the lac promotor (pSM31) [22, 25, 26]. Nevertheless, the selected expression system BL21(DE3)pLysS/pSM32 exhibited lower productivity than the native Nari2A^{T} strain. The BL21(DE3)pLysS/pSM32 strain was kept for further subtilisin purifications (Table 2).

The higher protease activity produced by *E. coli* BL21-AI™/pSM33 suggested that SAPN was more efficiently expressed in *E. coli* BL21-AI™ under the control of the promoter AraBAD than the pSM31 and pSM 32 expressed in BL21(DE3)pLysS.

The *E. coli* BL21-AI™/pSM33 transformants were also analyzed by protease activity assays and western blotting for their capacity to release His_6-rSAPN. The maximum protease activity reached 5,000 U/mL after 12 h of induction with L-arabinose (Fig. 4A), with an activity of 13,157 U/mg (Table 2). Fig. 5A shows that His_6-rSAPN BL21-AI™ was detected as an intense band with His-tag antibodies at around 33 kDa. A control culture was performed using the same protocol using a non transformed *E. coli* BL21-AI™ strain.

3.3. Expression in P. pastoris X33 and SMD1168 strains

For the production of SAPN in yeast, the pSM34, pSM35, and pSM36 cassettes were designed to express the SP-Pro-M (379 aa), the Pro-M (350 aa), and the M (274 aa) enzymes, respectively, and two *Pichia* strains were compared: X33 and SMD1168 (Table 1). Several *P. pastoris* X33 and SMD1168 recombinants were selected, and the subtilisin activity was monitored for each strain in the culture medium. As shown in Table 2, the constructions SP-Pro-M gave the highest activities, the Pro-M cassette exhibited a lower activity, and the lowest activity was observed for the constructs containing only the mature protein (M). Thus, the presence of both signal peptides (α-factor and the native SP of SAPN) leads to better results that could be possibly attributed to a better secretion level. This is consistent with previous work showing that the presence of dual signal peptides, the native one and that of the

expression vector (pGAPZαB or pPICZαB), enhanced the expression level of the protease from *B. stearothermophilus* FI [27]. Furthermore, the absence of the propeptide signal decreases the SAPN activity; this could be related to its regulatory role in protein secretion and folding.

Thus *P. pastoris* X33/pSM34 and SMD1168/pSM34 clones were used for further studies (Fig. 4B). Fig. 5A shows that His_6-rSAPN X33/pSM34 and His_6-rSAPN SMD1168/pSM34 were detected in the culture supernatants as intense bands with His-tag antibodies at around 45 kDa.

3.4. Purification of the rSAPNs and protease activity comparison

Activities of the crude extracts are presented in Tables 2 and 3. Yeast cells lead to the highest activity with almost 10 times more than strain Nari2A^T [19]. As illustrated in Table 3, purification yields for the different recombinant proteins range between 10 and 12. SDS-PAGE (Fig. 5B) and azo-casein zymography (Fig. 5C) showed the purity and integrity of the purified SAPNs. Similar to the native SAPN, the rSAPN BL21(DE3)pLysS enzyme consists of a single band of ~30 kDa, while His_6-rSAPN BL21-AI^{TM} showed a single band corresponding to ~33 kDa. The difference is attributed to the polylinker extension of hexahistidine. Recombinant SAPNs expressed in *P. pastoris* X33 and SMD1168 strains have a higher molecular mass (~45 kDa). The additional 15 kDa corresponds to the presence of the carboxy-terminal tags (c-myc and hexahistidine). Thus, yeast secretes the mature protein after cutting the signal sequences. This is in line with previous studies showing that a polylinker extension of hexahistidine may improve the yield of active protein [4, 28].

The N-terminal sequences of the recombinant proteins, identified by Edman degradation, confirmed that the recombinant proteins produced either in *E. coli* or in yeast correspond to SAPN proteins (Table 4).

Thus, the *P. pastoris* expression system is more efficient than that of *E. coli* for the production of active enzymes as illustrated in Table 3. We also observed that the presence of the SP (α-factor) provided by the pPICZαC vector, in addition to the native signal peptide, improved the subtilisin activity. These observations can be explained by the presence of the propeptide; indeed, it is an intramolecular chaperone that facilitates the protease domain folding similarly to other subtilisins [29]. The propeptide interferes with two surface helices and the active-site crevice of subtilisin. The C-terminus of the prodomain binds to the catalytic site in a substrate-like approach and competitively eliminates activity during autoprocessing [29].

3.5. Characterization of subtilisins rSAPNs and comparison with the SAPN wild-type

3.5.1. pH and temperature dependencies

The purified native and recombinant subtilisins (SAPN, rSAPN BL21(DE3)pLysS/pSM32, His_6-rSAPN BL21-AI™/pSM33, His_6-rSAPN X33/pSM34, and His_6-rSAPN SMD1168/pSM34) showed the same pH and temperature optima of 10 and 75°C, respectively, with a slightly higher enzymatic activity for rSAPN expressed in both yeast strains (Figs. 6A and B). In addition, we observed that rSAPNs enzymes were more stable at pH 11 and 40°C than SAPN (Fig. 6C). The residual subtilisin activities at 80°C showed for His_6-rSAPN BL21-AI™, His_6-rSAPN X33, and His_6-rSAPN SMD1168 a noteworthy increase in their half-lives versus SAPN and rSAPN BL21(DE3)pLysS (Fig. 6D and Supplementary data, Table S2). Similarly, the subtilisin thermostability was further improved by the addition of PEG 1000 (Fig. 6E and Supplementary data, Table S2). Thus, the rSAPNs expressed in *P. pastoris* strains X33 and SMD1168 proved to be the most stable enzymes, even when compared with native SAPN. The additional tags (c-myc and hexahistidine) may contribute to increased stability. The introduction of an affinity tag has

been suggested to have a positive effect on the biochemical properties of some hydrolases [30-32] but the reverse was sometimes observed [33]. The improved activity and stability of rSAPNs at alkaline pH and elevated temperatures is indeed a highly valued feature, which provides additional support to industrial bioprocesses that involve high pH and elevated temperatures.

3.5.2. Determination of kinetic parameters and degree of hydrolysis (DH)

As shown in Table 5, the closest kinetic parameters to those of the native subtilisin were found for rSAPN expressed in *E. coli*. They were somewhat different for the other rSAPNs enzymes. For all substrates, His_6-rSAPNs secreted by the two yeasts, in particular by SMD1168, showed the best k_{cat}/K_M, which suggested a better catalytic efficiency. A cumulative effect of His-tag on an expression in *P. pastoris* system improved the catalytic parameters. This was also observed in previous research on k_{cat}/K_M values of xylanases, amylases, and lipases expressed in tagged vectors in *E. coli* [34] and *P. pastoris* [30, 31, 35].

The DH of each recombinant protein was measured. The SAPN enzymes have similar hydrolysis curves for the white shrimp (Supplementary data, Fig. S1A) and blue crab by-products (Supplementary data, Fig. S1B). The results showed that while the *E. coli* recombinant enzymes show the same DH as native SAPN, rSAPN secreted by *P. pastoris* SMD1168 was the most effective enzyme with a DH of 61% and 57% on shrimp and crab by-products, respectively. This great potential in the deproteinization of crustacean by-products makes this enzyme a very good candidate for the preparation of protein hydrolysates.

3.6. Amino acid sequence comparison between SAPN and other subtilisins

Fig. 7 illustrates the sequence alignment of mature SAPN with other subtilisins. The comparison of subtilisins amino acid sequences showed that the mature SAPN possesses 96% sequence identity with subtilisin MP1 from *B. licheniformis* MP1 with a 10 aa difference.

SAPN showed 92% sequence identity with subtilisin subC from *B. licheniformis* 11594 with a divergence of 22 aa (Supplementary data, Table S3). In addition, there is 67% identity (91 substitutions) with that of protease SAPA from *Anoxybacillus kamchatkensis* M1V and 41% identity (162 substitutions) with that of SAPV produced by *Virgibacillus natechei* FarD^T [25, 26] (Fig. 7). The conserved catalytic triad (D32, H63, and S220) and three conserved subtilisin motifs at positions 23-42, 59-72, and 217-233 (Fig. 7) are present in mature SAPN as described by Varshney et al. [36] and Ouelhadj et al. [37] and confirmed by the InterProScan tool according to the PRINTS database. Based on the three critical residues in the catalytic site and on statistically significant similarities in amino acid sequence to other alkaline serine proteases, SAPN should be assigned to the S8 subtilisin family.

It is particularly interesting to note that despite the high percentage of identity, the substitutions of some key amino acids (Supplementary data, Table S3) in the mature form of SAPN, which varies between MP1 and other subtilisins, contributed to improving its alkaline pH and thermal stability as well as its catalytic efficiency. Statistical analyses showed more charged (17.67%), hydrophobic (45.1%), and aromatic (9.5%) residues in SAPN than in mesophilic subtilisin. Furthermore, a trend toward substitutions such as G→A and K→R was found. In accordance with previous research showing a higher content of Ala residues in thermophilic enzymes, a higher content of alanine (14.5%) was found in SAPN than in MP1. The properties of alanine make this residue the best helix-forming residue [38].

3.7. Homology modeling

To obtain an overview of the mechanisms by which the SAPN enzyme operates at high pH and temperature conditions, with high substrate specificity and an incredibly efficient DH, a homology modeling has been conducted. As yet unknown, the actual structure of SAPN is assumed to be similar to that of many subtilisins, particularly those with high sequence

identity and whose structures are elucidated. Thus, the available 3D structure of pro-subtilisin E (PDB ID: 3WHI), which displays 65% sequence identity with Pro-SAPN was used for building the 3D model of Pro-SAPN from *M. thermohalophilus* Nari2A^T (Fig. 8).

The root mean square deviation that implies α-carbons between the initial and the optimized model was 0.85 Å, which indicates a good degree of similarity of the superimposed model and template; this suggests the reliability of the predicted structure for its use in further studies. The Ramachandran diagram of the final Pro-SAPN model generated using the MolProbity program revealed that 93% of the amino acids are located in favored regions (Supplementary data, Fig. S2).

The structural model of Pro-SAPN presents a globular form with a characteristic α/β-folded hydrolase (Fig. 8). This structure shows a unique and compact domain having eleven beta-strands and nine alpha-helices. It is arranged in a central structure comprising eight parallel beta-strands—namely β1-β7, β11, and the anti-parallel β10-strand—all bordered by seven alpha-helices—namely α1-α3 and α6-α9—in addition to the antiparallel β8-β9-strands on one side and the α4- and α5-helices on the other side.

In the Pro-SAPN model, two calcium-binding sites were predicted by the COACH tool using the 3D structure of pro-subtilisin E (PDB ID: 3WHI) as a template. The Pro-SAPN model shows the first calcium-binding site (Ca_1 or high-affinity site), involving the oxygen atoms of 6 potential residues—namely Q2, D41, L75, N77, S79, and V81. In the pro-subtilisin E (PDB ID: 3WHI), the Ca_1 is heptacoordinated with the oxygen atoms at the vertices of a pentagonal bipyramid. The second calcium binding site (Ca_2 or low-affinity site) potentially involves 3 residues, which are A169, Y171, and V174. The Ca_2 in subtilisin Carlsberg (PDB ID: 2SEC) is presumably occupied by a K^+ ion in subtilisin BPN′. A third calcium-binding site in subtilisin Carlsberg (Ca_3) has a relatively high occupancy and low *B* factor, but the mean ligand distance of 2.85Å is again more typical of a K^+ ion [7].

4. Conclusions

The current study constitutes to our knowledge the first molecular investigation that targets the subtilisin synthesized by a new thermophilic taxon, *M. thermohalophilus* Nari2A^T. The data highlight the high potential of SAPN, from strain Nari2A^T, isolated, cloned, and overexpressed in bacteria and yeast. This study allowed the production of four new rSAPN enzymes secreted by *E. coli* and by yeast. The yeast strain *P. pastoris* as compared to *E. coli* is a more efficient expression system for the production of large quantities of soluble, correctly folded, and potent subtilisins rSAPNs. Herein, a successful approach was designed for engineering SAPN, which leads to the improvement of activity, catalytic efficiency, pH stability profile, and thermal resistance. Given its novel properties and characteristics, the His_6-rSAPN enzyme produced in SMD1168 *Pichia* strain has proven to be a potential candidate for numerous biotechnological industries, mainly in the bioprocesses of the deproteinization of crustacean by-products. The in-depth study and characterization of the role of SAPN amino acids is the next challenge to understand at the molecular level the involvement of these residues in protease activity and stability, in light of the 3D model.

Abbreviations

aa: amino acids; AOX1, alcohol oxidase operon 1; BSA, bovine serum albumin; BMGY, buffered glycerol-complex, BMMY, buffered methanol-complex; BTEE, *N*-benzyol-L-tyrosine ethyl ester; DH, degree of hydrolysis; IPTG, isopropyl-thio-β-D-galactopyranoside; LB, Luria-Bertani; M, mature protein; MM, molecular mass; PEG, Polyethylene glycol; Pro, propeptide; RBS, ribosomal binding site; SAPN, a alkaline serine protease from *M. thermohalophilus* strain Nari2A^T designated as subtilisin SAPN; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis, SP, signal peptide; ORF, open reading frame; Suc-

FAAF-*p*NA, *N*-succinyl-L-Phe-L-Ala-L-Ala-L-Phe-*p*-nitroanilide; TBST, Tris-Buffered Saline and Tween 20; TCA, trichloroacetic acid; and YPD, yeast extract peptone dextrose.

Acknowledgments

The authors are greatly indebted to Dr. T. Bizouarn and Mr. S.B Owusu (ICP, UMR 8000-CNRS, Université Paris-Saclay) for their helpful suggestions and invaluable technical help. We want to express our great admiration to Prof. A. F. Gargouri (LBME, CBS) and Dr. A. Ezzine (LIP-MB/INSAT, Université de Carthage) for their valuable and constructive suggestions and discussions throughout this work. We extend our thanks to Dr. N. Ammara Addou (LBCM, FSB-USTHB, Algeria) for the gift of the *M. thermohalophilus* Nari2A^{T}. The authors also want to sincerely acknowledge Dr. I. Boudia (Université de M'Sila) for language editing and polishing services as well as constructive proofreading.

Funding information

This work was supported by the Ministère de l'Enseignement Supérieur et de la Recherche Scientifique (MESRS) in Tunisia in the framework of the Contract Programs LBMIE-CBS, code grant no.: LR15CBS06 (2015-2018) and LBMEB-CBS, code grant no.: LR19CBS01 (2019-2022), the Multilateral Project Partenariats Hubert Curien (PHC)-Maghreb 2020 Program (FranMaghZYM 2020-2023, code Campus France: 43791TM & code PHC: 01MAG20), and the Algerian-Tunisian R&I Cooperation for the Mixed Laboratories of Scientific Excellence 2021-2024, code LABEX/TN/DZ/21/01 as well as by the Ph.D. scholarship given by the Ecole Doctorale Sciences Fondamentales, Faculté des Sciences de Sfax (Université de Sfax), code grant no.: ED08FSSf01.

Declaration of Competing Interest

The authors declare that there are no conflicts of interest.

References

[1] C.R. Harwood, Introduction to the Biotechnology of *Bacillus*, in: C.R. Harwood (Ed.) *Bacillus*, Springer US, Boston, MA, 1989, pp. 1-4.

[2] G. Xie, Z. Shao, L. Zong, X. Li, D. Cong, R. Huo, Heterologous expression and characterization of a novel subtilisin-like protease from a thermophilic *Thermus thermophilus* HB8, Int. J. Biol. Macromol. 138 (2019) 528-535.

[3] P. Bryan, Protein engineering of subtilisin, Biochim. Biophys. Acta 1543 (2001) 203-222.

[4] H. Rekik, F. Frikha, N. Zaraî Jaouadi, F. Gargouri, N. Jmal, S. Bejar, B. Jaouadi, Gene cloning, expression, molecular modeling and docking study of the protease SAPRH from *Bacillus safensis* strain RH12, Int. J. Biol. Macromol. 125 (2019) 876-891.

[5] B. Jaouadi, N. Aghajari, R. Haser, S. Bejar, Enhancement of the thermostability and the catalytic efficiency of *Bacillus pumilus* CBS protease by site-directed mutagenesis, Biochimie 92 (2010) 360-369.

[6] M.L. Stahl, E. Ferrari, Replacement of the *Bacillus subtilis* subtilisin structural gene with an in vitro-derived deletion mutation, J. Bacteriol. 158 (1984) 411-418.

[7] P. Carter, B. Nilsson, J.P. Burnier, D. Burdick, J.A. Wells, Engineering subtilisin BPN′ for site-specific proteolysis, Proteins: Struct. Funct. Bioinf. (1989) 240-248.

[8] M. Jacobs, M. Eliasson, M. Uhlén, J.I. Flock, Cloning, sequencing and expression of subtilisin Carlsberg from *Bacillus licheniformis*, Nucleic Acids Res. 13 (1985) 8913-8926.

[9] S. Mechri, M. Ben Elhoul Berrouina, M. Omrane Benmrad, N. Zaraî Jaouadi, H. Rekik, E. Moujehed, A. Chebbi, S. Sayadi, M. Chamkha, S. Bejar, B. Jaouadi, Characterization of a novel protease from *Aeribacillus pallidus* strain VP3 with potential biotechnological interest, Int. J. Biol. Macromol. 94 (2017) 221-232.

[10] A. Goldblum, On the Mechanisms of Proteinases. In: Náray-Szabó G., Warshel A. (eds) Computational Approaches to Biochemical Reactivity. Understanding Chemical Reactivity, vol 19. Springer, Dordrecht, 2002. https://doi.org/10.1007/0-306-46934-0_7

[11] A. Laskar, E. Rodger, A. Chatterjee, C. Mandal, Modeling and structural analysis of evolutionarily diverse S8 family serine proteases, Bioinformation 7 (2011) 239-45.

[12] D.A. Estell, F. Goedegebuur, M. Kolkman, R. Mejldal, K. Augustyn, L.M. Babe, R.R. Bott, J. Yao, R. Ghirnikar, AprL-clade protease variants and uses thereof, 2016, WO/2016/183509, PCT/US2016/032514.

[13] R. Daly, M.T. Hearn, Expression of heterologous proteins in *Pichia pastoris*: a useful experimental tool in protein engineering and production, J. Mol. Recognit.18 (2005) 119-138.

[14] J. Eder, A.R. Fersht, Pro-sequence-assisted protein folding, Mol. Microbiol. 16 (1995) 609-614.

[15] R. Singh, M. Kumar, A. Mittal, P.K. Mehta, Microbial enzymes: industrial progress in 21st century, 3 Biotech 6 (2016) 174.

[16] S. Mechri, I. Sellem, K. Bouacem, F. Jabeur, M. Chamkha, H. Hacene, A. Bouanane-Darenfed, B. Jaouadi, Antioxidant and enzyme inhibitory activities of *Metapenaeus monoceros* by-product hydrolysates elaborated by purified alkaline proteases, Waste Biomass Valori. 11 (2020) 6741-6755.

[17] H.Y. Zhao, H. Feng, Engineering *Bacillus pumilus* alkaline serine protease to increase its low-temperature proteolytic activity by directed evolution, BMC Biotechnol. 18 (2018) 34.

[18] S. Mohamed, K. Bouacem, S. Mechri, N.A. Addou, H. Laribi-Habchi, M.-L. Fardeau, B. Jaouadi, A. Bouanane-Darenfed, H. Hacène, Purification and biochemical

characterization of a novel acido-halotolerant and thermostable endochitinase from *Melghiribacillus thermohalophilus* strain Nari2A^T, Carbohydr. Res. 473 (2019) 46-56.

[19] S. Mechri, K. Bouacem, F. Jabeur, S. Mohamed, N. Ammara Addou, A. Dab, A. Bouraoui, A. Bouanane-Darenfed, S. Bejar, H. Hacène, L. Baciou, F. Lederer, B. Jaouadi, Purification and biochemical characterization of a novel thermostable and halotolerant subtilisin SAPN, a serine protease from *Melghiribacillus thermohalophilus* Nari2A^T for chitin extraction from crab and shrimp shell by-products, Extremophiles 23 (2019) 529-547.

[20] N. Addou, P. Schumann, C. Spröer, W. Ben Hania, H. Hacene, G. Fauque, J. L. Cayol, M. L. Fardeau, *Melghiribacillus thermohalophilus* gen. nov., sp. nov., a novel filamentous, endospore-forming, thermophilic and halophilic bacterium, Int. J. Syst. Evol. Microbiol. 65 (2015) 1172-1179.

[21] J. Sambrook, E. Fritsch, T. Maniatis, Molecular Cloning: A Laboratory Manual, 2nd edn, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA., 1989, pp. 23-38.

[22] B. Jaouadi, S. Ellouz-Chaabouni, M. Rhimi, S. Bejar, Biochemical and molecular characterization of a detergent-stable serine alkaline protease from *Bacillus pumilus* CBS with high catalytic efficiency, Biochimie 90 (2008) 1291-305.

[23] N. Zaraî Jaouadi, B. Jaouadi, N. Aghajari, S. Bejar, The overexpression of the SAPB of *Bacillus pumilus* CBS and mutated *sapB*-L31I/T33S/N99Y alkaline proteases in *Bacillus subtilis* DB430: new attractive properties for the mutant enzyme, Bioresour. Technol. 105 (2012) 142-151.

[24] P. Gouet, X. Robert, E. Courcelle, ESPript/ENDscript: extracting and rendering sequence and 3D information from atomic structures of proteins, Nucleic Acids Res 31(13) (2003) 3320-3.

[25] S. Mechri, K. Bouacem, N. Zaraî Jaouadi, H. Rekik, M. Ben Elhoul, M. Omrane Benmrad, H. Hacene, S. Bejar, A. Bouanane-Darenfed, B. Jaouadi, Identification of a novel protease from the thermophilic *Anoxybacillus kamchatkensis* M1V and its application as laundry detergent additive, Extremophiles 23 (2019) 687-706.

[26] S. Mechri, K. Bouacem, M. Amziane, A. Dab, F. Nateche, B. Jaouadi, Identification of a new serine alkaline peptidase from the moderately halophilic *Virgibacillus natechei* sp. nov., strain FarD^T and its application as bioadditive for peptide synthesis and laundry detergent formulations, BioMed Res. Int. 2019 (2019) Article ID 6470897, 17 pages.

[27] A.A. Latiffi, A.B. Salleh, R.N.Z.R.A. Rahman, S.N. Oslan, M. Basri, Secretory expression of thermostable alkaline protease from *Bacillus stearothermophilus* FI by using native signal peptide and α-factor secretion signal in *Pichia pastoris*, Genes Genet. Syst. 88 (2013) 85-91.

[28] A. Hmida-Sayari, F. Elgharbi, A. Farhat, H. Rekik, K. Blondeau, S. Bejar, Overexpression and biochemical characterization of a thermostable phytase from *Bacillus subtilis* US417 in *Pichia pastoris*, Mol. Biol. Rep. 56 (2014) 839-848.

[29] Z. Ignatova, F. Wischnewski, H. Notbohm, V. Kasche, Pro-sequence and Ca^{2+}-binding: implications for folding and maturation of Ntn-hydrolase penicillin amidase from *E.coli*, J. Mol. Biol. 348 (2005) 999-1014.

[30] S. Trabelsi, M. Sahnoun, F. Elgharbi, R. Ameri, S. Ben Mabrouk, M. Mezghani, A. Hmida-Sayari, S. Bejar, *Aspergillus oryzae* S2 AmyA amylase expression in *Pichia pastoris*: production, purification and novel properties, Mol. Biol. Rep. 46 (2019) 921-932.

[31] F. Elgharbi, A. Hmida-Sayari, Y. Zaafouri, S. Bejar, Expression of an *Aspergillus niger* xylanase in yeast: application in breadmaking and in vitro digestion, Int. J. Biol. Macromol. 79 (2015) 103-109.

[32] N. Hmidet, A. Bayoudh, J.G. Berrin, S. Kanoun, N. Juge, M. Nasri, Purification and biochemical characterization of a novel α-amylase from *Bacillus licheniformis* NH1: cloning, nucleotide sequence and expression of *amyN* gene in *Escherichia coli*, Process Biochem. 43 (2008) 499-510.

[33] H. Horchani, S. Ouertani, Y. Gargouri, A. Sayari, The N-terminal His-tag and the recombination process affect the biochemical properties of *Staphylococcus aureus* lipase produced in *Escherichia coli*, J. Mol. Catal. B Enzym. 61 (2009) 194-201.

[34] F. Elgharbi, H. Ben Hlima, A. Farhat-Khemakhem, D. Ayadi-Zouari, S. Bejar, A. Hmida-Sayari, Expression of *A. niger* US368 xylanase in *E. coli*: purification, characterization and copper activation, Int. J. Biol. Macromol. 74 (2015) 263-270.

[35] X. F. Zhang, Y. H. Ai, Y. Xu, X. W. Yu, High-level expression of *Aspergillus niger* lipase in *Pichia pastoris*: characterization and gastric digestion in vitro, Food Chem. 274 (2019) 305-313.

[36] D. Varshney, A. Jaiswar, A. Adholeya, P. Prasad, Phylogenetic analyses reveal molecular signatures associated with functional divergence among subtilisin like serine proteases are linked to lifestyle transitions in Hypocreales, BMC Evol. Biol. 16 (2016) 220.

[37] A. Ouelhadj, K. Bouacem, K. L. Asmani, F. Allala, S. Mechri, M. Yahiaoui, B. Jaouadi, Identification and homology modeling of a new biotechnologically compatible serine alkaline protease from moderately halotolerant *Gracilibacillus boraciitolerans* strain LO15, Int. J. Biol. Macromol. 161 (2020) 1456-1469.

[38] C. Vieille, G.J. Zeikus, Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability, Microbiol. Mol. Biol. Rev. 65 (2001) 1-43.

[39] A. Riffel, A. Brandelli, Isolation and characterization of a feather-degrading bacterium from the poultry processing industry, J. Ind. Microbiol. Biotechnol. 29 (2002) 255-258.

[40] K.A. Walsh, Trypsinogens and trypsins of varoius species, Methods Enzymol. 19 (1970) 41-63.

[41] E.G. DelMar, C. Largman, J.W. Brodrick, M.C. Geokas, A sensitive new substrate for chymotrypsin, Anal. Biochem. 99 (1979) 316-320.

Table 1

List of host strains and plasmids used in this study

Plasmids	**Cloned fragments**	**Host strains**	**Generated plasmids names**	**Recombinant enzyme (His$_6$-tagged/ Untagged)**
pCR-Blunt™	SAPN with its signal peptide + propeptide	*E. coli* One Shot® TOP10	pSM30	Untagged
pUT57	SAPN with its signal peptide + propeptide	BL21(DE3)pLysS	pSM31	Untagged
pTrc99A	SAPN with its signal peptide + propeptide	BL21(DE3)pLysS	pSM32	Untagged
pDEST™17	SAPN with its signal peptide + propeptide	*E. coli* BL21-AI™	pSM33	His$_6$-tagged
pPICZαC	SAPN with its signal peptide + propeptide	*E. coli* One Shot® TOP10	**pSM34**	His$_6$-tagged
pPICZαC	SAPN with its propeptide (without its signal peptide)	*E. coli* One Shot® TOP10	**pSM35**	His$_6$-tagged
pPICZαC	SAPN free of both peptides	*E. coli* One Shot® TOP10	**pSM36**	His$_6$-tagged

The names of the plasmids that were linearized and then transformed into *P. pastoris* X33 and SMD1168 strains are underlined and in bold.

Table 2

List of selected recombinant strains and their extracellular protease activities

Cloned fragments	Selected recombinant strains/Generated plasmids (constructions)	Protease activity (U/mL) [a]	Activity (U/mg of protein)	Molecular mass of predicted subtilisin (kDa)
SAPN with a signal peptide + propeptide	*E. coli* BL21(DE3)pLysS/pSM31	755 ± 8	2,517	30
	E. coli BL21(DE3)pLysS/pSM32	3,000 ± 32	9,375	30
	E. coli BL21-AI™/pSM33	5,000 ± 50	13,157	33
	P. pastoris X33/pSM34	8,200 ± 50	14,386	45
	P. pastoris SMD1168/pSM34	16,000 ± 83	21,476	45
SAPN with its propeptide (without its signal peptide)	*P. pastoris* X33/pSM35	3,200 ± 32	4,105	45
	P. pastoris SMD1168/pSM35	5,310 ± 51	6,900	45
SAPN free of both peptides	*P. pastoris* X33/pSM36	400 ± 4	715	45
	P. pastoris SMD1168/pSM36	720 ± 7	1,010	45

[a] The production was started from a 200 mL culture.

SAPN: Serine alkaline protease from *M. thermohalophilus* Nari2A^T or subtilisin SAPN.

Table 3

Summary of activities of target rSAPNs in the crude extract and after purification

Wild-type and recombinant SAPN enzyme [a]		**Crude extract** [b]	**Purified enzyme**			
		Total activity (units [c]**) × 10^3**	**Total activity (units** [c]**) × 10^3**	**Specific activity (U/mg of protein)**	**Activity yield (%)**	**Purification factor (*n*-fold)**
Nari2A^T	SAPN	3160 ± 36	948 ± 5	94,800	17.5	10.2
E. coli	rSAPN BL21(DE3)pLysS/pSM32	600 ± 7	105 ± 1	95,454	17.5	10.2
	His_6-rSAPN BL21-AI^{TM}/pSM33	1,000 ± 12	279 ± 4	126,818	27.9	9.6
P. pastoris	His_6-rSAPN X33/pSM34	1,640 ± 19	1,023 ± 12	170,500	62.3	11.8
	His_6-rSAPN SMD1168/pSM34	3,200 ± 32	1,960 ± 24	257,894	61.2	12.0

[a] These data are the average of at least 3 assays.

[b] The purification was started from a 200 mL culture.

[c] The unit definition is given in § 2.7.

SAPN: Serine alkaline protease from *M. thermohalophilus* Nari2A^T or subtilisin SAPN.

Table 4

N-terminal sequences of rSAPNs determined by Edman degradation

rSAPN enzyme	Amino acid sequence length	N-terminal sequences 1 5 10 15 20 25
Native SAPN [19]	25	AQTVPYGIPLIKAFKVPAQGFKVAN
rSAPN BL21(DE3)pLysS/pSM32	27	AQTVPYGIPLIKAFKVPAQGFKVANVK
His_6-rSAPN BL21-AI™/pSM33	30	MSYYHHHHHHLESTSLYKKAGLAQTVPYGI
His_6-rSAPN X33/pSM34	22	AQTVPYGIPLIKAFKVPAQGFK
His_6-rSAPN SMD1168/pSM34	20	AQTVPYGIPLIKAFKVPAQG

The amino acid sequences for comparison were obtained using the program BLASTP (NCBI, NIH, USA) database.

The underlined sequence corresponds to the N-terminal hexahistidine tag.

SAPN: Serine alkaline protease from *M. thermohalophilus* Nari2A^T or subtilisin SAPN.

Table 5

Kinetic parameters of target rSAPN enzymes

Substrate	Enzyme	K_M (mM)	k_{cat} (min^{-1})	k_{cat}/K_M (min^{-1} mM^{-1})	Catalytic efficiency relative to His_6-rSAPN SMD1168
Casein [a]	SAPN	0.390 ± 0.03	63,200	162,051	0.16
	rSAPN BL21(DE3)pLysS/pSM32	0.387 ± 0.03	63,434	164,010	0.17
	His_6-rSAPN BL21-AI™/pSM33	0.340 ± 0.03	84,545	248,661	0.25
	His_6-rSAPN X33/pSM34	0.190 ± 0.01	113,667	598,247	0.61
	His_6-rSAPN SMD1168/pSM34	0.175 ± 0.01	171,929	982,451	1.00
Azo-casein [b]	SAPN	0.489 ± 0.04	44,167	90,321	0.18
	rSAPN BL21(DE3)pLysS/pSM32	0.471 ± 0.04	44,051	93,526	0.19
	His_6-rSAPN BL21-AI™/pSM33	0.317 ± 0.02	58,711	185,208	0.38
	His_6-rSAPN X33/pSM34	0.260 ± 0.02	79,487	305,719	0.63
	His_6-rSAPN SMD1168/pSM34	0.247 ± 0.02	119,395	483,380	1.00
BTEE [c]	SAPN	0.572 ± 0.06	26,154	45,723	0.26
	rSAPN BL21(DE3)pLysS/pSM32	0.554 ± 0.05	26,697	48,189	0.27
	His_6-rSAPN BL21-AI™/pSM33	0.428 ± 0.04	34,535	80,689	0.46
	His_6-rSAPN X33/pSM34	0.411 ± 0.03	48,173	117,209	0.67
	His_6-rSAPN SMD1168/pSM34	0.402 ± 0.03	70,322	174,706	1.00
Suc-FAAF-*p*NA [d]	SAPN	0.635 ± 0.07	116,814	187,959	0.24
	rSAPN BL21(DE3)pLysS/pSM32	0.620 ± 0.07	117,352	189,277	0. 25
	His_6-rSAPN BL21-AI™/pSM33	0.573 ± 0.06	155,140	270,750	0.36
	His_6-rSAPN X33/pSM34	0.471 ± 0.05	209,147	444,048	0.60
	His_6-rSAPN SMD1168/pSM34	0.438 ± 0.04	323,618	738,853	1.00

These data are the average of at least 3 assays.

[a] The activity on casein was measured at $A_{660\ nm}$ [23].

[b] The activity on azo-casein was determined at $A_{440\ nm}$ [39].

[c] Esterase activities on BTEE were measured at $A_{253\ nm}$ [40].

[d] The activity on Suc-FAAF-*p*NA was determined at $A_{410\ nm}$ [41].

SAPN: Serine alkaline protease from *M. thermohalophilus* Nari2A^T or subtilisin SAPN.

Figure legends

Fig. 1. Schematic diagram of the recombinant SM30-32 plasmids. Construction of the recombinant plasmids pSM30, pSM31, and pSM32 designed to overexpress serine alkaline protease (subtilisin SAPN) in competent *E. coli* strain BL21(DE3)pLysS.

Fig. 2. Nucleotide sequence with putative amino acid sequence of the *sapN* gene encoding SAPN. The *sapN* gene has 1140-bp encoding a protein of 379-aa. The position of PCR primers (F-MS15, R-MS16, F-MS17, and R-MS18) is underlined. The black box shows the N-terminal residues of mature serine alkaline protease (subtilisin SAPN). This sequence has the accession no. **MT292372** in the GenBank public repository.

Fig. 3. Schematic diagram of the recombinant SM33-36 plasmids. (A) Gateway™ cloning of the plasmid pSM33 for recombinant serine alkaline protease (subtilisin SAPN) expression in *E. coli* BL21-AI™. **(B)** Expression cassettes of *sapN* (plasmids pSM34, pSM35, and pSM36) introduced into *P. pastoris* X33 and SMD1168 genomes.

Fig. 4. Growth kinetics and subtilisin activity in *E. coli* and yeast recombinant strains. Time course of growth and subtilisin activity in the culture supernatants for four clones in *E. coli* BL21(DE3)pLysS and BL21-AI™ strains and *P. pastoris* X33 and SMD1168 strains. **(A)** The IPTG (0.4 mM) and arabinose (2 g/L) were added after 8 h and 10 h to induce the Ptrc and AraBAD promoters, respectively. **(B)** The methanol [0.5% (*v*/*v*)] was added every 24 h to induce the strong AOX1 promoter. Aliquots were taken every 2 h **(A)** and 6 h **(B)** to follow the growth and the protease activity for 36 h and 240 h, respectively. Bars represent ± SE.

Fig. 5. Electrophoretic analysis of target SAPN subtilisins. (A) Western blot of total proteins (50 μg) produced by recombinant clones using the mouse-His-probe antibody: Lane no. 1: non-transformed *P. pastoris* X33, Lane no. 2: recombinant *P. pastoris* X33, Lane no. 3:

recombinant *P. pastoris* SMD1168, Lane no. 4: non-transformed *P. pastoris* SMD1168, Lane no. 5: recombinant *E. coli* BL21-AI™, and Lane no. 6: nontransformed *E. coli* BL21-AI™. **(B)** SDS-PAGE (10%) of purified native (serine alkaline protease or subtilisin SAPN) and recombinant (rSAPNs) enzymes (50 μg). Lanes nos. 1 - 6: SAPN, rSAPN BL21(DE3)pLysS, His_6-rSAPN BL21-AI™, His_6-rSAPN X33, His_6-rSAPN SMD1168, and molecular mass marker, respectively. **(C)** Zymography of subtilisin activity of each purified wild-type (SAPN) and recombinant enzymes (rSAPNs) (50 μg). Lanes nos. 1 - 5: SAPN, rSAPN BL21(DE3)pLysS, His_6-rSAPN BL21-AI™, His_6-rSAPN X33, and His_6-rSAPN SMD1168, respectively. In the zymogram, the activity corresponds to a clear band on blue background due to the hydrolysis of azo-casein. The 3 kDa difference between His_6-rSAPN BL21-AI™ with SAPN and rSAPN BL21(DE3)pLysS is attributed to the polylinker extension of hexahistidine, with a theoretical mass of 2.6 kDa. The higher mass (~45 kDa) observed for the recombinant SAPN expressed in *P. pastoris* X33 and SMD1168 strains as compared to the native one and recombinant enzymes expressed in *E. coli* is due to the presence of the carboxy-terminal tags present on the pPICZαC vector: c-myc followed by hexahistidine. These tags add 15 kDa to the molecular weight of the rSAPN.

Fig. 6. Physicochemical properties of target serine alkaline protease SAPN subtilisins. pH **(A)** and temperature **(B)** dependences of rSAPN activity. The activity for each rSAPN at the optimum pH was considered as a control. Influence of pH **(C)** and temperature **(D)** on the stability of the protease activity. Each subtilisin was kept at 80°C (in the absence or presence of 2 mM $CaCl_2$). Residual protease activity was assayed every 2 h for 24 h. **(E)** Combined effect of PEG 1000 (100 g/L) and Ca^{2+} (2 mM) on rSAPNs activity. Remaining activity was assayed at 2 h intervals for 24 h. The activity of subtilisins without additives was taken as 100%. Bars represent ± SE.

Fig. 7. Structure-based multiple sequence alignment of serine alkaline protease subtilisin SAPN amino acid sequences with several subtilisins. Multiple sequence alignment of mature SAPN from *M. thermohalophilus* Nari2A^T with subtilisin E from *B. subtilis* subsp. *subtilis* str. 168 (PDB ID: 3WHI, UniProtKB/Swiss-Prot or GenBank accession no.: P04189), MP1 from *B. licheniformis* MP1 (UniProtKB/Swiss-Prot: Q6PNN5, GenBank accession no.: ADJ80171), subC from *B. licheniformis* 11594 (UniProtKB/Swiss-Prot: Q45299, GenBank accession no.: CAA62666), and SAPA from *Anoxybacillus kamchatkensis* (UniProtKB/Swiss-Prot: A0A482EWF5, GenBank accession no.: MH979641). Identical residues are in red boxes and conserved groups are framed in blue. Highly conserved residues inside each conserved group are presented as red letters. The subtilisin molecular signatures are framed in green. Green stars indicate the typical catalytic triad residues D32, H63, and S220. The ten different residues between SAPN and MP1 are indicated with magenta triangles. 2D structure elements of the subtilisin E 3D structure are given above the alignment. "α" (α-helices), "β" (β-strands), and "T" (β-turns).

Fig. 8. Predicted 3D structure of Pro-SAPN. The 3D model generated from the 3D structure of pro-subtilisin E (PDB ID: 3WHI). The propeptide chain is represented as a ribbon. The catalytic triad residues (D32, H63, and S220) are presented as red sticks. The 10 residues (F14D, P17Q, V23G, R77T, S78T, I109G, E147V, I148V, M233I, and G234L) which are different from those in subtilisin MP1 are indicated as blue sticks. The serine alkaline protease subtilisin SAPN propeptide structure is shown as a yellow surface. Ca_1 and Ca_2 are both represented as orange spheres in their respective predicted locations using the COACH server.

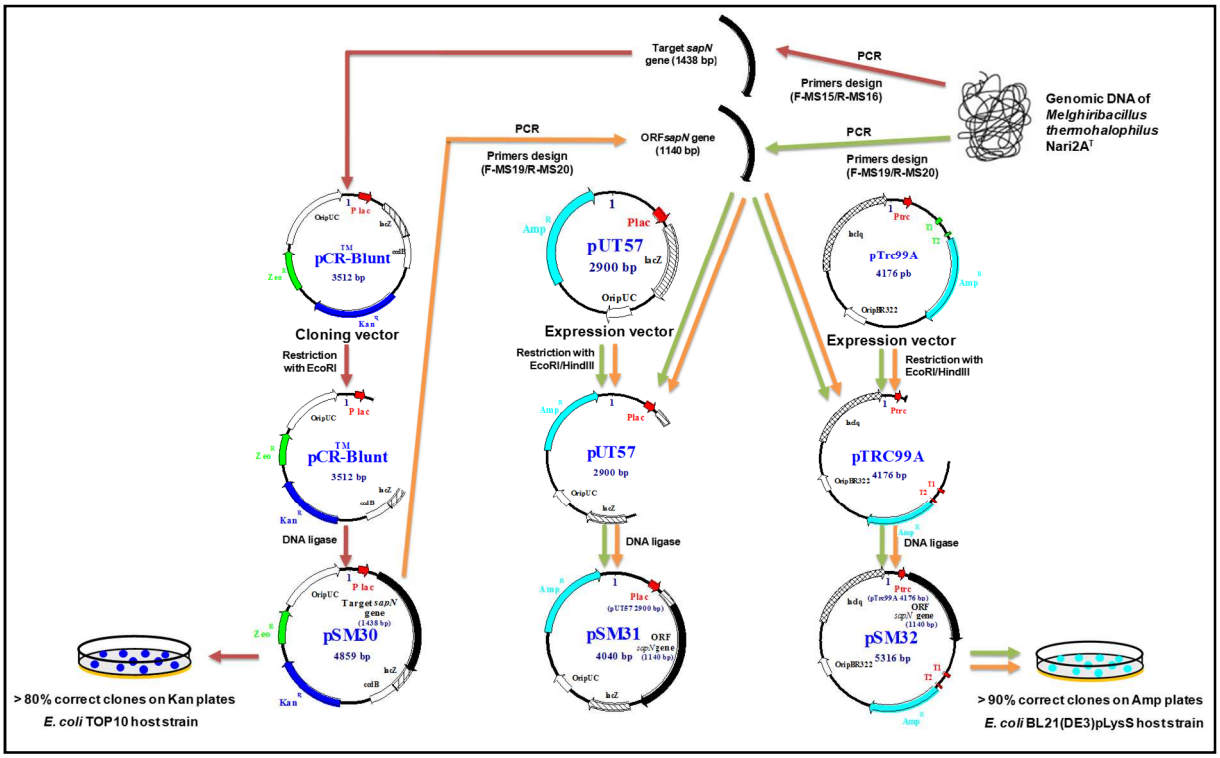

Target sapN gene (1438 bp)
PCR
Primers design (F-MS15/R-MS16)
Genomic DNA of Melghiribacillus thermohalophilus Nari2A^T
PCR
Primers design (F-MS19/R-MS20)
ORFsapN gene (1140 bp)
PCR
Primers design (F-MS19/R-MS20)
pCR-Blunt 3512 bp
Cloning vector
Restriction with EcoRI
pUT57 2900 bp
Expression vector
Restriction with EcoRI/HindIII
pTrc99A 4176 pb
Expression vector
Restriction with EcoRI/HindIII
pCR-Blunt 3512 bp
DNA ligase
pUT57 2900 bp
DNA ligase
pTRC99A 4176 bp
DNA ligase
pSM30 4859 bp
pSM31 4040 bp
pSM32 5316 bp
> 80% correct clones on Kan plates
E. coli TOP10 host strain
> 90% correct clones on Amp plates
E. coli BL21(DE3)pLysS host strain

Figure 1

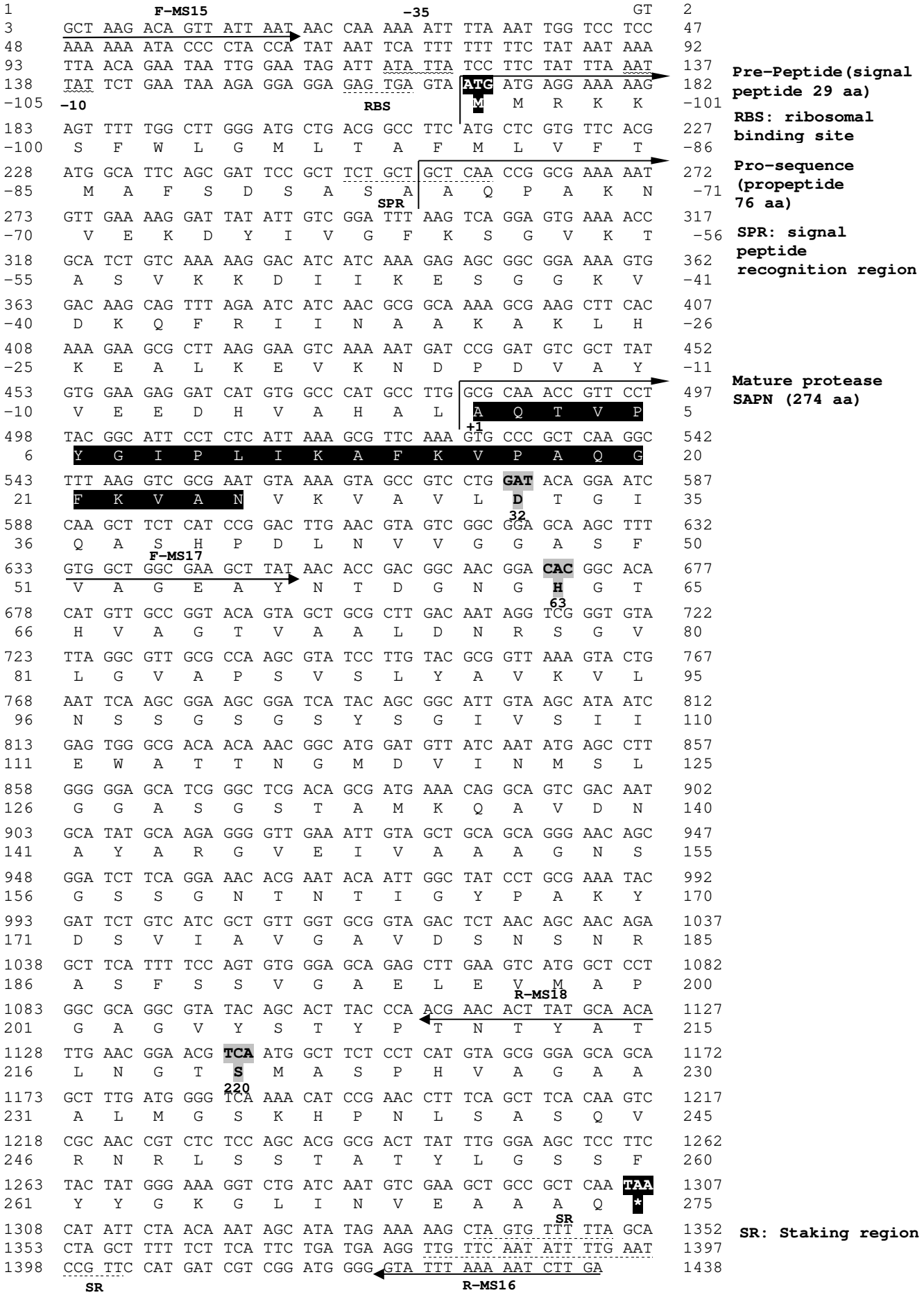
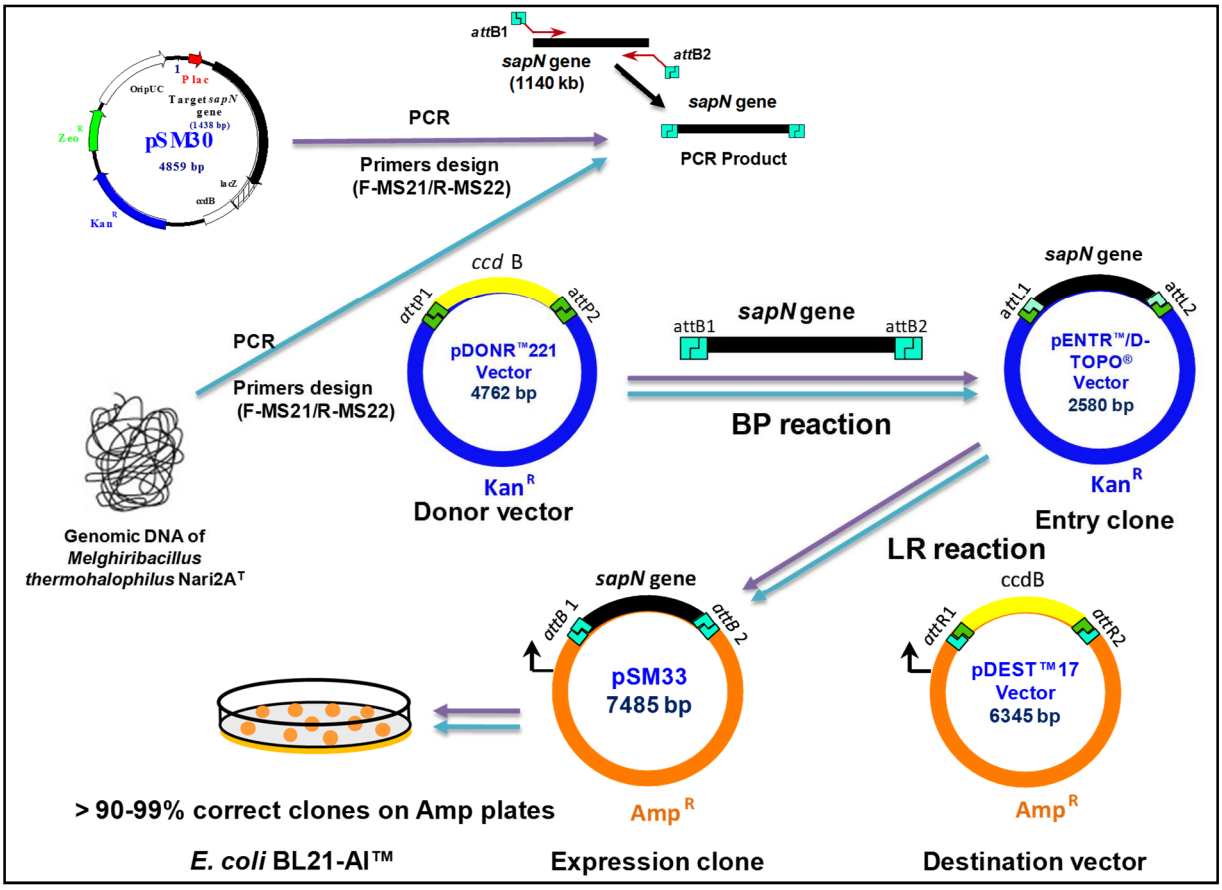

F-MS15
-35
-10
RBS
SPR
+1
F-MS17
32
63
R-MS18
220
R-MS16
SR
SR
Pre-Peptide(signal peptide 29 aa)
RBS: ribosomal binding site
Pro-sequence (propeptide 76 aa)
SPR: signal peptide recognition region
Mature protease SAPN (274 aa)
SR: Staking region

Figure 2

A


attB1
sapN gene
(1140 kb)
attB2
sapN gene
PCR Product
PCR
Primers design
(F-MS21/R-MS22)
pSM30
4859 bp
Target sapN gene (1438 bp)
P lac
OriUC
Zeo R
Kan R
lacZ
ccdB
ccd B
attP1
attP2
pDONR™221 Vector 4762 bp
Kan R
Donor vector
sapN gene
attB1
attB2
BP reaction
attL1
attL2
pENTR™/D-TOPO® Vector 2580 bp
Kan R
Entry clone
LR reaction
Genomic DNA of Melghiribacillus thermohalophilus Nari2AT
sapN gene
attB 1
attB 2
pSM33
7485 bp
Amp R
ccdB
attR1
attR2
pDEST™17 Vector 6345 bp
Amp R
> 90-99% correct clones on Amp plates
E. coli BL21-AI™
Expression clone
Destination vector

B

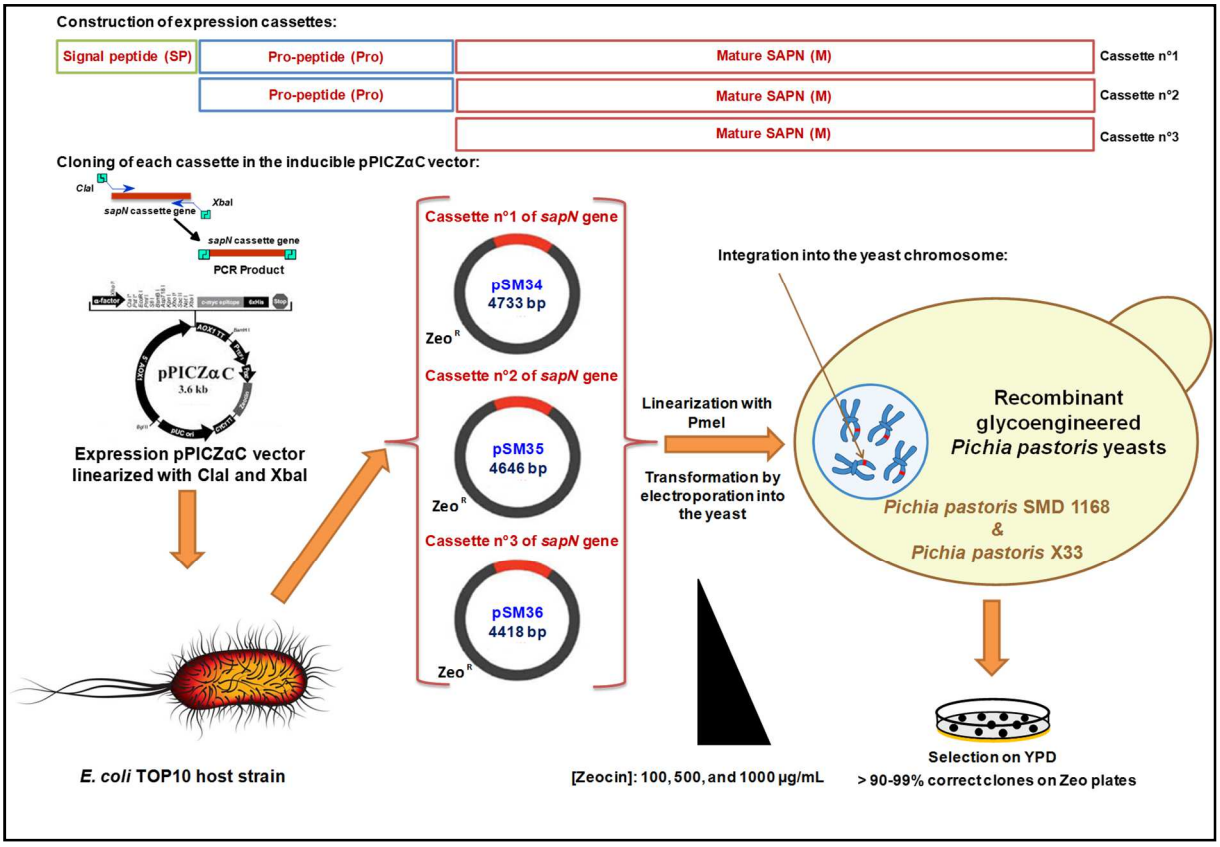

Construction of expression cassettes:
Signal peptide (SP)
Pro-peptide (Pro)
Mature SAPN (M)
Cassette n°1
Pro-peptide (Pro)
Mature SAPN (M)
Cassette n°2
Mature SAPN (M)
Cassette n°3
Cloning of each cassette in the inducible pPICZαC vector:
ClaI
XbaI
sapN cassette gene
sapN cassette gene
PCR Product
pPICZαC
3.6 kb
Expression pPICZαC vector linearized with ClaI and XbaI
Cassette n°1 of sapN gene
pSM34
4733 bp
Zeo R
Cassette n°2 of sapN gene
pSM35
4646 bp
Zeo R
Cassette n°3 of sapN gene
pSM36
4418 bp
Zeo R
Linearization with PmeI
Transformation by electroporation into the yeast
Integration into the yeast chromosome:
Recombinant glycoengineered Pichia pastoris yeasts
Pichia pastoris SMD 1168
&
Pichia pastoris X33
E. coli TOP10 host strain
[Zeocin]: 100, 500, and 1000 µg/mL
Selection on YPD
> 90-99% correct clones on Zeo plates

Figure 3

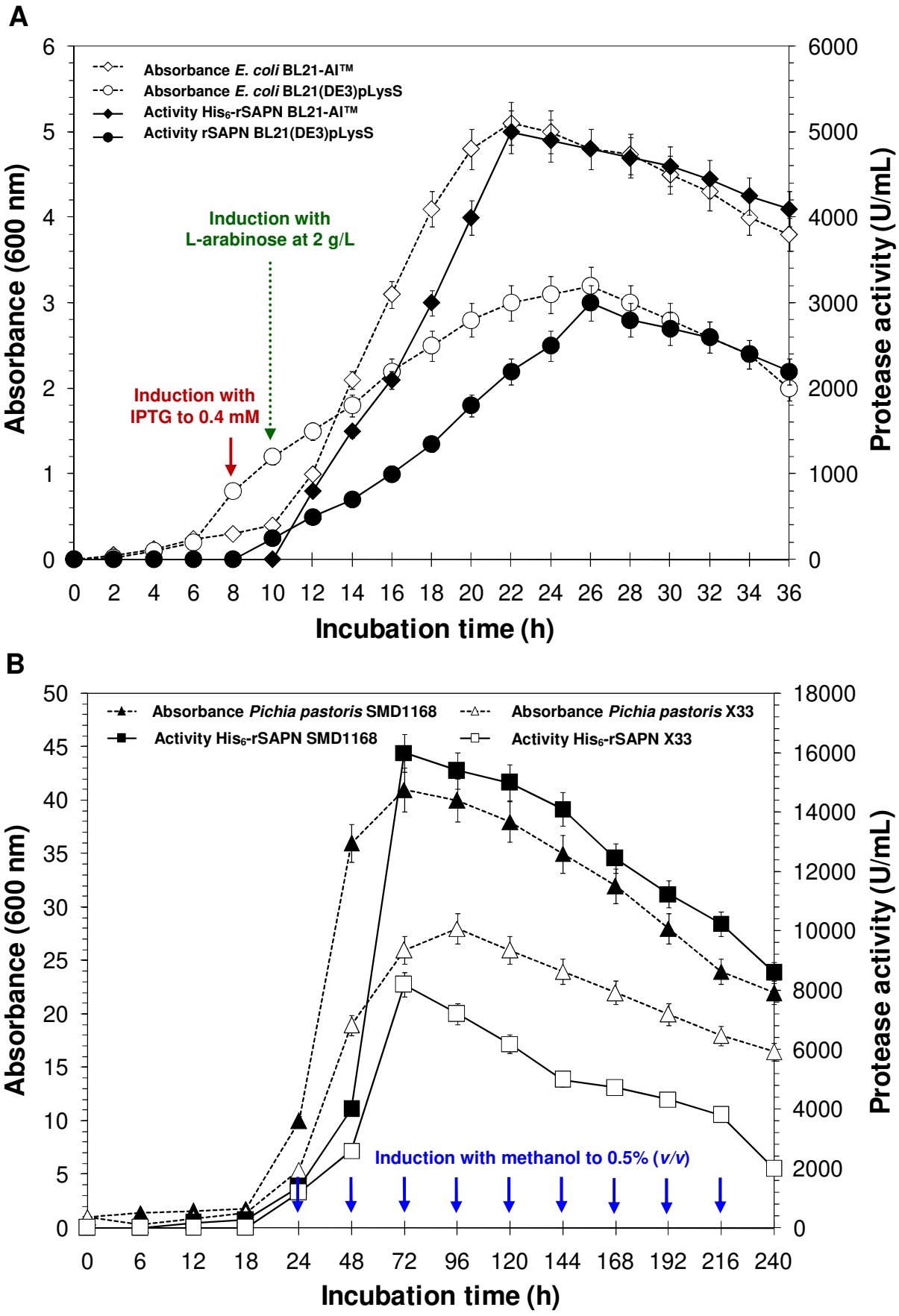

A
Absorbance E. coli BL21-AI™
Absorbance E. coli BL21(DE3)pLysS
Activity His6-rSAPN BL21-AI™
Activity rSAPN BL21(DE3)pLysS
Induction with L-arabinose at 2 g/L
Induction with IPTG to 0.4 mM
Absorbance (600 nm)
Protease activity (U/mL)
Incubation time (h)
B
Absorbance Pichia pastoris SMD1168
Absorbance Pichia pastoris X33
Activity His6-rSAPN SMD1168
Activity His6-rSAPN X33
Induction with methanol to 0.5% (v/v)
Absorbance (600 nm)
Protease activity (U/mL)
Incubation time (h)

Figure 4

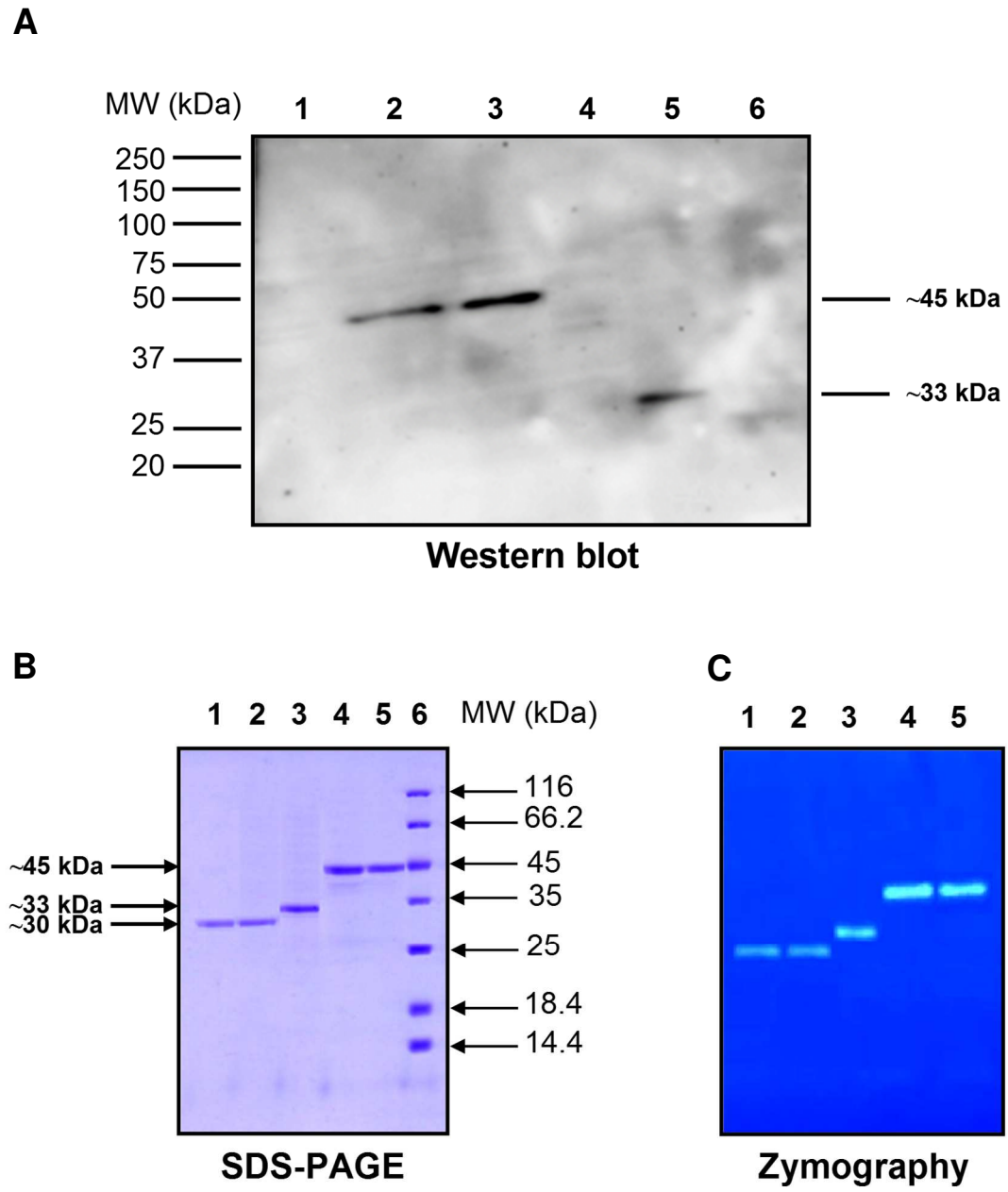

A
MW (kDa)
1
2
3
4
5
6
250
150
100
75
50
37
25
20
~45 kDa
~33 kDa
Western blot
B
1 2 3 4 5 6 MW (kDa)
~45 kDa
~33 kDa
~30 kDa
116
66.2
45
35
25
18.4
14.4
SDS-PAGE
C
1 2 3 4 5
Zymography

Figure 5

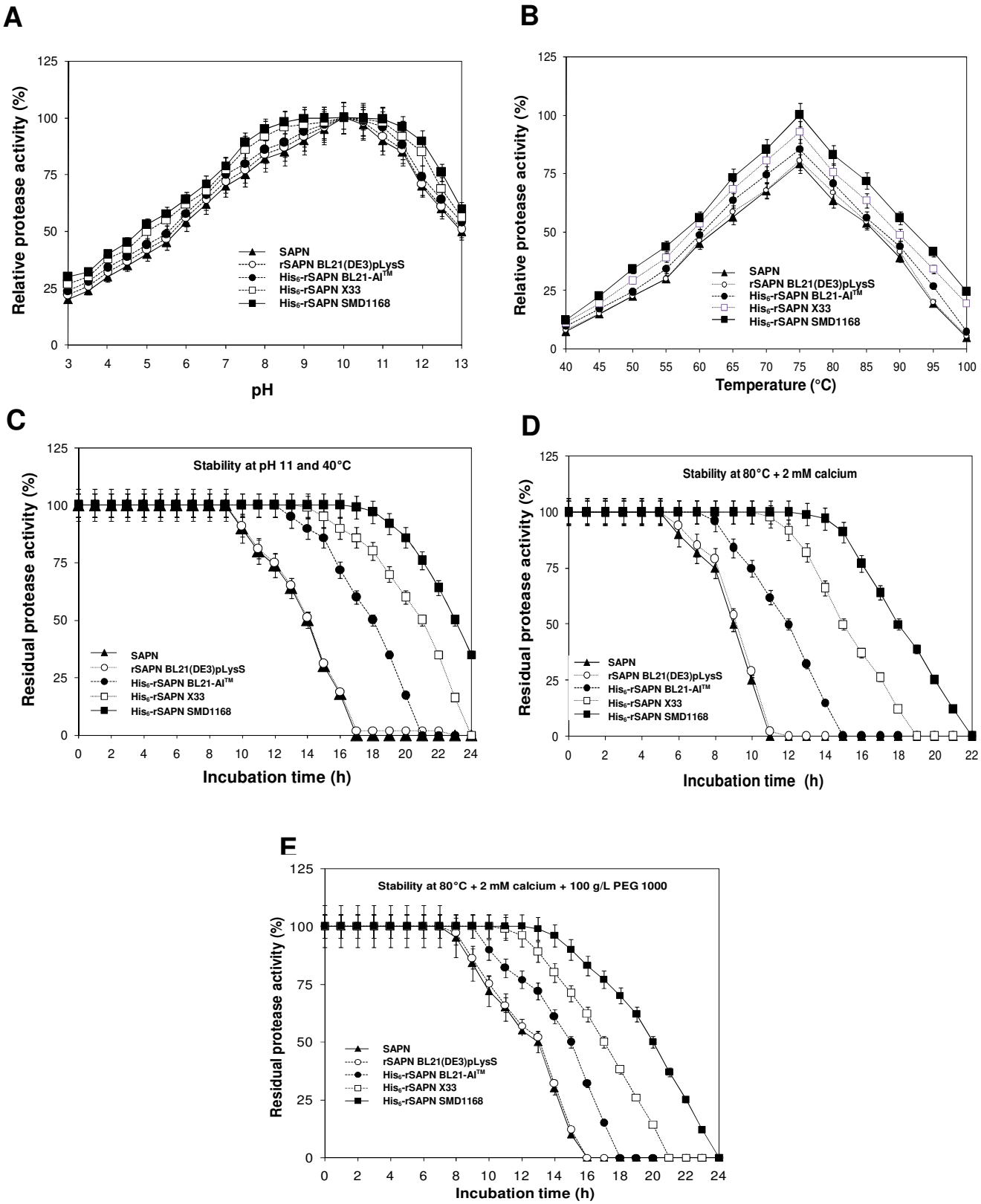

A
Relative protease activity (%)
pH
SAPN
rSAPN BL21(DE3)pLysS
His6-rSAPN BL21-AI™
His6-rSAPN X33
His6-rSAPN SMD1168
B
Relative protease activity (%)
Temperature (°C)
C
Stability at pH 11 and 40°C
Residual protease activity (%)
Incubation time (h)
D
Stability at 80°C + 2 mM calcium
Residual protease activity (%)
Incubation time (h)
E
Stability at 80°C + 2 mM calcium + 100 g/L PEG 1000
Residual protease activity (%)
Incubation time (h)

Figure 6

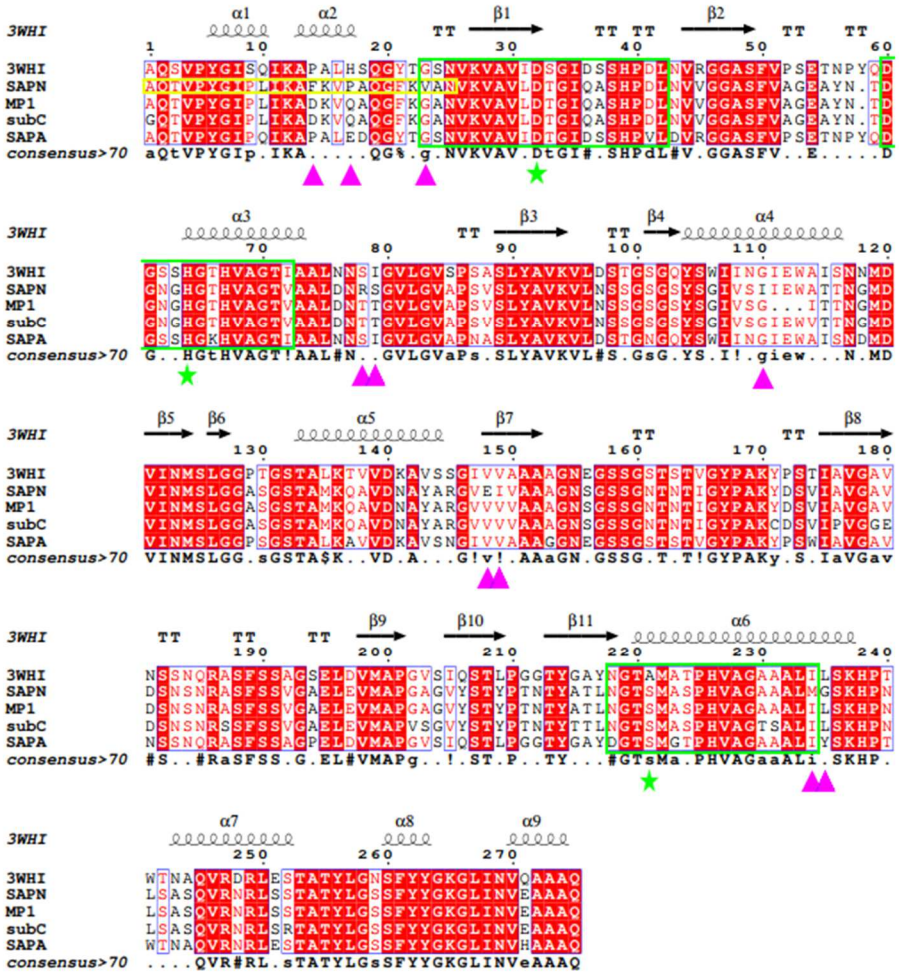


Figure 7

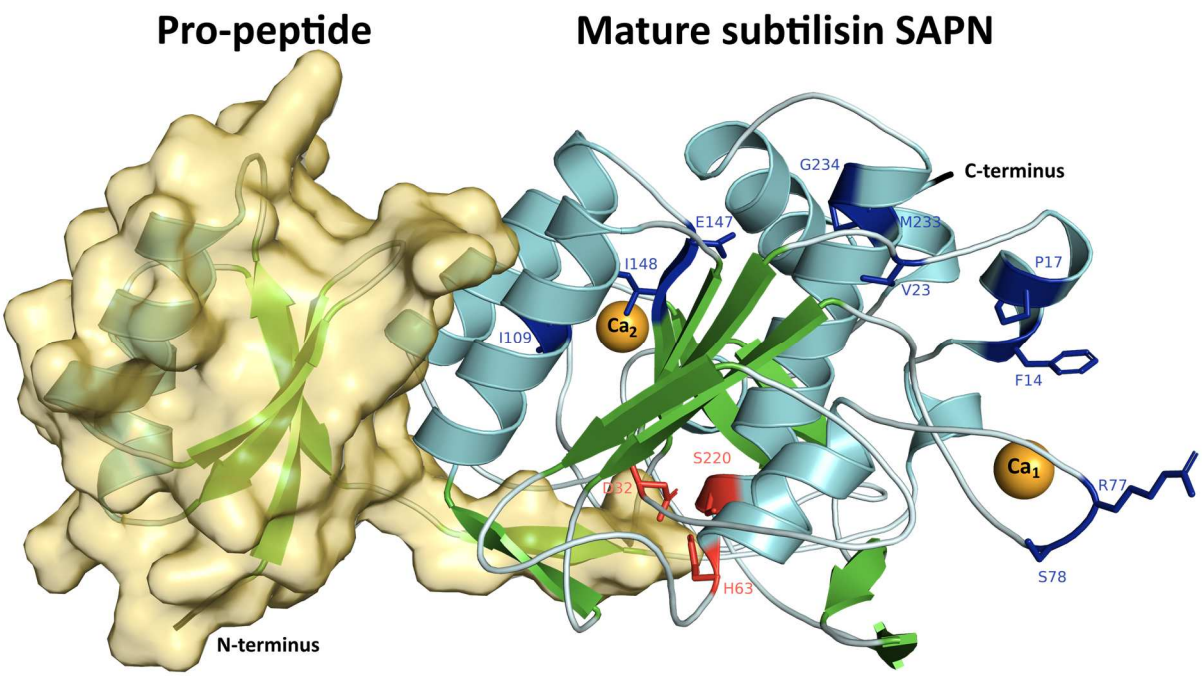

Pro-peptide
Mature subtilisin SAPN
G234
C-terminus
E147
M233
I148
P17
V23
Ca2
I109
F14
S220
Ca1
R77
D32
H63
S78
N-terminus

Figure 8

Graphical Abstract