KINETIC AND PROTEOMIC IDENTIFICATION OF PROTEASE INHIBITORS IN MARINE INVERTEBRATES. CHARACTERIZATION OF A CARBOXYPEPTIDASE INHIBITOR ISOLATED FROM THE MOLLUSC *Nerita versicolor*

Giovanny Covaleda Cortés

October 2012

KINETIC AND PROTEOMIC IDENTIFICATION OF PROTEASE INHIBITORS IN MARINE INVERTEBRATES. CHARACTERIZATION OF A CARBOXYPEPTIDASE INHIBITOR ISOLATED FROM THE MOLLUSC *Nerita versicolor*

Doctoral thesis presented by Giovanny Covaleda Cortés to obtain the PhD degree in "Structure and Function of Proteins" from the Universitat Autònoma de Barcelona

This work has been performed in the Protein Engineering and Enzymology Laboratory, Institut de Biotecnologia i Biomedicina "Vicent Villar i Palasi". Thesis work supervised by Profs. Francesc Xavier Avilés Puigvert and María de los Ángeles Chávez Planes

Giovanny Covaleda Cortés

Francesc Xavier Avilés Puigvert

María de los Ángeles Chávez Planes

Bellaterra, October 2012

Les 7 jours de la création

« 1.1 Au commencement, Dieu créa les cieux et la terre. 1.2 La terre était informe et vide: il y avait des ténèbres à la surface de l'abîme, et l'esprit de Dieu se mouvait au-dessus des eaux. 1.3 Dieu dit: Que la lumière soit! Et la lumière fut. 1.4 Dieu vit que la lumière était bonne; et Dieu sépara la lumière d'avec les ténèbres. 1.5 Dieu appela la lumière jour, et il appela les ténèbres nuit. Ainsi, il y eut un soir, et il y eut un matin: ce fut le premier jour. 1.6 Dieu dit: Qu'il y ait une étendue entre les eaux, et qu'elle sépare les eaux d'avec les eaux. 1.7 Et Dieu fit l'étendue, et il sépara les eaux qui sont au-dessous de l'étendue d'avec les eaux qui sont au-dessus de l'étendue. Et cela fut ainsi. 1.8 Dieu appela l'étendue ciel. Ainsi, il y eut un soir, et il y eut un matin: ce fut le second jour. 1.9 Dieu dit: Que les eaux qui sont au-dessous du ciel se rassemblent en un seul lieu, et que le sec paraisse. Et cela fut ainsi. 1.10 Dieu appela le sec terre, et il appela l'amas des eaux mers. Dieu vit que cela était bon. 1.11 Puis Dieu dit: Que la terre produise de la verdure, de l'herbe portant de la semence, des arbres fruitiers donnant du fruit selon leur espèce et ayant en eux leur semence sur la terre. Et cela fut ainsi. 1.12 La terre produisit de la verdure, de l'herbe portant de la semence selon son espèce, et des arbres donnant du fruit et ayant en eux leur semence selon leur espèce. Dieu vit que cela était bon. 1.13 Ainsi, il y eut un soir, et il y eut un matin: ce fut le troisième jour. 1.14 Dieu dit: Qu'il y ait des luminaires dans l'étendue du ciel, pour séparer le jour d'avec la nuit; que ce soient des signes pour marquer les époques, les jours et les années; 1.15 et qu'ils servent de luminaires dans l'étendue du ciel, pour éclairer la terre. Et cela fut ainsi. 1.16 Dieu fit les deux grands luminaires, le plus grand luminaire pour présider au jour, et le plus petit luminaire pour présider à la nuit; il fit aussi les étoiles. 1.17 Dieu les plaça dans l'étendue du ciel, pour éclairer la terre, 1.18 pour présider au jour et à la nuit, et pour séparer la lumière d'avec les ténèbres. Dieu vit que cela était bon. 1.19 Ainsi, il y eut un soir, et il y eut un matin: ce fut le quatrième jour. 1.20 Dieu dit: Que les eaux produisent en abondance des animaux vivants, et que des oiseaux volent sur la terre vers l'étendue du ciel. 1.21 Dieu créa les grands poissons et tous les animaux vivants qui se meuvent, et que les eaux produisirent en abondance selon leur espèce; il créa aussi tout oiseau ailé selon son espèce. Dieu vit que cela était bon. 1.22 Dieu les bénit, en disant: Soyez féconds, multipliez, et remplissez les eaux des mers; et que les oiseaux multiplient sur la terre. 1.23 Ainsi, il y eut un soir, et il y eut un matin: ce fut le cinquième jour. 1.24 Dieu dit: Que la terre produise des animaux vivants selon leur espèce, du bétail, des reptiles et des animaux terrestres, selon leur espèce. Et cela fut ainsi. 1.25 Dieu fit les animaux de la terre selon leur espèce, le bétail selon son espèce, et tous les reptiles de la terre selon leur espèce. Dieu vit que cela était bon. 1.26 Puis Dieu dit: Faisons l'homme à notre image, selon notre ressemblance, et qu'il domine sur les poissons de la mer, sur les oiseaux du ciel, sur le bétail, sur toute la terre, et sur tous les reptiles qui rampent sur la terre. 1.27 Dieu créa l'homme à son image, il le créa à l'image de Dieu, il créa l'homme et la femme. 1.28 Dieu les bénit, et Dieu leur dit: Soyez féconds, multipliez, remplissez la terre, et l'assujettissez; et dominez sur les poissons de la mer, sur les oiseaux du ciel, et sur tout animal qui se meut sur la terre. 1.29 Et Dieu dit: Voici, je vous donne toute herbe portant de la semence et qui est à la surface de toute la terre, et tout arbre ayant en lui du fruit d'arbre et portant de la semence: ce sera votre nourriture. 1.30 Et à tout animal de la terre, à tout oiseau du ciel, et à tout ce qui se meut sur la terre, ayant en soi un souffle de vie, je donne toute herbe verte pour nourriture. Et cela fut ainsi. 1.31 Dieu vit tout ce qu'il avait fait et voici, cela était très bon. Ainsi, il y eut un soir, et il y eut un matin: ce fut le sixième jour. 2.1 Ainsi furent achevés les cieux et la terre, et toute leur armée. 2.2 Dieu acheva au septième jour son oeuvre, qu'il avait faite: et il se reposa au septième jour de toute son oeuvre, qu'il avait faite. 2.3 Dieu bénit le septième jour, et il le sanctifia, parce qu'en ce jour il se reposa de toute son oeuvre qu'il avait créée en la faisant. 2.4 Voici les origines des cieux et de la terre, quand ils furent créés. »

Genèse, chapitre 1, versets 1-31 et chapitre 2, versets 1-2.

SUMMARY

Approximately 2% of the genes in most organisms are proteases and proteolytic activity is involved in practically all biochemical and physiological processes in living organisms, as well as in the generation/transmission of different pathologies. Consequently, proteolytic enzymes are of great interest for both biotechnological and pharmaceutical industry as potential drugs targets. Therefore, protease inhibitors have already become into very attractive and promising molecules as potential drugs and diagnostic tools.

The present work describes the enzymatic kinetic identification of inhibitory activities against metallocarboxypeptidases A and B of the M14 family (MCPs, of the CPA and CPB type), pepsin, papain, trypsin and subtilisin in 30 Caribbean marine invertebrate extracts. The screening was performed on aqueous crude extracts from species belonging to the Phyla *Annelida, Bryozoa, Chordata, Cnidaria, Echinodermata, Mollusca,* and *Porifera* collected on the northern coast of Cuba. The qualitative results showed that fourteen extracts exhibited trypsin inhibitory activity, nine extracts displayed CPA, CPB, papain and subtilisin inhibitory activities, whilst only four extracts showed pepsin inhibitory activity. The most promising extracts showing inhibitory activities in terms of specific inhibitory activity (the highest values), dose-response relationships, IC_{50} (the lowest values) and bioavailability of the species were *N. versicolor*, *P. homomalla* and *S. helianthus* for CPA; *N. versicolor*, *C. muricatus* and *P. homomalla* for CPB; *B. granulifera*, *H. carunculata* and *D. listerianum* for pepsin; *P. homomalla*, *P. physalis* and *C. muricatus* for papain; *S. helianthus*, *N. versicolor*, *C. muricatus* and *N. peloronta* for trypsin; and *L. isodyctialis*, *I. badionotus* and *C. muricatus* for subtilisin.

This was complemented with the use of Intensity Fading (IF) MALDI-TOF MS, a medium-throughput proteomic-based strategy for the quick identification of those inhibitors in the marine extracts. It proved to be very useful to selectively identify, in some extracts, proteinaceous molecules (inhibitors) able to interact with the immobilized target enzymes, such as in *N. versicolor* against CPA and CPB, *H. carunculata* with pepsin, *S. helianthus* with papain, *N. peloronta* and *S. helianthus* with trypsin and *L. isodyctialis* for subtilisin. This work also describes the application of MALDI-TOF MS fragmentation and characterization approach using the resolved and most promising molecular species obtained in the elution fraction of IF MALDI-TOF MS. Two examples are described for this strategy: in the first one is the major peak resulting from the elution fraction of microaffinity interaction between *H. carunculata* crude extract and pepsin-NHS activated Sepharose™, which was fragmented by CID MALDI-TOF/TOF and subsequently analyzed by *de novo* sequencing. Another example was about ISD MALDI-TOF MS fragmentation and sequence analysis approach of the elution fraction of microaffinity interaction between *S. helianthus* crude extract and trypsin-glyoxal Sepharose®. Both approaches show an excellent complementarity with the better established IF MALDI-TOF MS.

Another important goal of this work was the purification and structural–functional characterization of NvCI, a proteinaceous carboxypeptidase inhibitor isolated from the marine snail *N. versicolor*. The combination of techniques such as automated Edman degradation and *de* novo *sequencing* by MALDI-TOF MS allowed the establishment of its primary structure. This strategy included a prior chemical modification of cysteine residues, followed by hydrolysis with endoproteinases such as trypsin, Lys-C and Glu-C. The synthesis and cloning of cDNA encoding NVCI allowed to confirm the amino acid sequence of this molecule. The amino acid sequence of NvCI showed that this protein consists of 53 amino acid residues with a molecular mass of 5944 Da, and three intrachain disulfide bridges (UniProt code: P86912).

Recombinant NvCI was produced in *Pichia pastoris* system with a high yield in terms of protein and activity. The kinetic characterization of natural and recombinant NvCI was performed according to the strategy described for tight-binding inhibitors. Both molecules displayed K_i values in the picomolar range against carboxypeptidases such as bCPA1, hCPA1 and hCPA4. Other K_i values are in the order of $1x10^{-10}$ M (hCPB1, pCPB1, hTAFI and bTAFI), except for hCPA2 ($1x10^{-9}$ M). NvCI is a competitive reversible tight binding inhibitor which represents the most potent carboxypeptidase inhibitor from proteinaceous nature so far described.

Summary

The X-ray structure of the inhibitor, in complex with the enzyme, displays a new NvCI extended protein-folding small motif, which is basically formed by a central anti-parallel β-two-stranded sheet connected by three major loops. The β-strands are stabilized by three disulfide bridges, a small hydrophobic core located next to the C-terminus of the protein and a few bulky exposed hydrophobic residues to the solvent. The C-terminal tail of NvCI is shorter in comparison to the other known carboxypeptidase inhibitors, only formed by two residues, Tyr52 and Ala53. This short extension is sufficient to interact with the active site residues and zinc atom of the carboxypeptidase (PDB code: 4A94).

NvCI interacts extensively with hCPA4, with a total contact area of 1875.1 $Å^2$. The inhibition mechanism of NvCI is due to a competitive interaction with the active site of the carboxypeptidase, by occlusion of the active-site subsites S1', S1, S2 and S3, as it has been observed for most of the reported proteinaceous carboxypeptidase inhibitors. These sites are occupied by the C-terminal tail of NvCI and constitute the primary contact region of the inhibitor. The secondary contact region is more extended and covers almost one face of the inhibitor.

The most relevant structural difference in the NvCI-hCPA4 complex is the main-chain conformation of the P3 residue in comparison to the other known carboxypeptidase inhibitors, which favours the formation of two extra hydrogen bonds with Glu163 and presumably induces a reduction of the inhibitory constant by stabilization of the product formation. However, the lower K_i values observed for NvCI in comparison to the other inhibitors can be attributed to both, the "primary" and "secondary" interaction regions, which create an extended interface with the carboxypeptidase that minimizes the product release from the catalytic reaction.

ABBREVIATIONS

1,5-DAN	1,5-Diaminonaphthalene
α-CHCA	α-Cyano-4-hydroxycinnamic acid
AAFA	N-(4-Methoxyphenylazoformyl)-Arginine-OH
AAFP	N-(4-Methoxyphenylazoformyl)-Phenylalanine-OH
ACN	Acetonitrile
ACI	Carboxypeptidase inhibitor from *Ascaris suum*
ANOVA	Analysis of variance
AOX	Alcohol oxidase
BAPNA	Nα-Benzoyl-L-arginine 4-nitroanilide hydrochloride
BCA	Bicinchoninic acid
bCPA	Bovine carboxypeptidase A1
BIRD	Blackbody infrared radiative dissociation
BPTI	Bovine pancreatic trypsin inhibitor
BLAST	Basic local alignment search tool
BSA	Bovine serum albumin
BSM	Basal salts medium for *Pichia pastoris*
CCPs	Cytosolic carboxypeptidases
CD	Circular dichroism
cDNA	Complementary Deoxyribonucleic acid
CID	Collision-induced dissociation
CP	Carboxypeptidase
CPD-I	Carboxypeptidase D domain I from *Drosophila melanogaster*
CPY	carboxypeptidase Y from *Saccharomyces cerevisiae*
Da	Daltons
D_{AB}	Diffusion coefficient of solute A in solvent B
D_{eff}	Effective diffusion coefficient
DHAP	2,6-Dihydroxyacetophenone
DHB	2,5-dihydroxy-benzoic acid
DI	Degree of immobilization
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
EA	Enzymatic activity
ECD	Electron capture dissociation
EDD	Electron-detachment dissociation
ETD	Electron transfer dissociation
E-64	*trans*-Epoxysuccinyl-L-leucylamido(4-guanidino)butane
EDTA	Ethylenediaminetetraacetic acid
GGLPNA	Benzylocarbonyl-glycyl-glycyl-L-Leucine 4-nitroanilide
Glu-C	Glutamyl endopeptidase from *Staphylococcus aureus*
hCPA1	Human carboxypeptidase A-1
hCPA2	Human carboxypeptidase A-2
hCPA4	Human carboxypeptidase A-4
hCPB	Human carboxypeptidase B
HPLC	High performance liquid chromatography
IF MALDI-TOF MS	Intensity Fading MALDI-TOF mass spectrometry
IRMPD	Infrared multiphoton dissociation
ISD	In-source decay
kDa	Kilodaltons
K_M	Michaelis-Menten constant
K_i	Equilibrium dissociation constant
K_{iapp}	Apparent equilibrium dissociation constant
kV	Kilovolts
LAP	Leucine aminopeptidase microsomal from porcine kidney

Abbreviations

LC	Liquid chromatography
LCI	Leech carboxypeptidase inhibitor
LSPNPLAL	H-Leu-Ser-p-nitro-Phe-Nle-Ala-Leu-OMe.TFA
Lys-C	Lysyl endopeptidase from *Lysobacter enzymogenes*
m/z	Mass/charge
MALDI-TOF	Matrix assisted laser-desorption ionization time-of-flight
MCP	Metallocarboxypeptidase
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
NCBI	National Center for Biotechnology Information
NHS	N- hydroxysuccinimide
NvCI	*Nerita versicolor* carboxypeptidase inhibitor
PAGE	Polyacrylamide gel electrophoresis
pCPB	Porcine carboxypeptidase B
PCI	Potato carboxypeptidase inhibitor
PCR	Polymerase Chain Reaction
PD	Photodissociation
PDB	Protein Data Bank
PI	Protease inhibitor
PMF	Peptide mass fingerprinting
PFLNA	L-Pyroglutamyl-L-phenylalanyl-L-leucine-*p*-nitroanilide
PSD	Post-source decay
PTM	Trace minerals solution for *Pichia pastoris* growth
RACE	Rapid Amplification of cDNA Ends
RNA	Ribonucleic acid
rNvCI	Recombinant *Nerita versicolor* carboxypeptidase inhibitor
RP-HPLC	Reversed-phase high performance liquid chromatography
SA	Sinapic acid
SD	Standard deviation
SDS	Sodium dodecyl sulfate
SID	Surface-induced dissociation
ShPI	*Stichodactyla helianthus* protease inhibitor
SmCI	*Sabellastarte magnifica* carboxypeptidase inhibitor
super-DHB	Mixture of 2,5-dihydroxy-benzoic acid and 2-hydroxy-5-methoxy -benzoic acid
TAFI	Thrombin-activable fibrinolysis inhibitor
TCA	Trichloroacetic acid
TCI	Tick carboxypeptidase inhibitor
TDS	Top-down sequencing
TFA	Trifluoroacetic acid
UV	Ultraviolet

TABLE OF CONTENTS

Table of contents

Table of contents

I. INTRODUCTION

I.1. General Considerations

I.1.1. Proteolytic enzymes. General characteristics and classification

Proteolytic enzymes, also known as peptidases or proteases comprise a group of proteins with the ability to catalyze the hydrolysis of peptide bonds. These enzymes account for approximately 2% of the genes in humans and other forms of life López-Otín and Matrisian, 2007, including infectious organisms and the replication/transmission of different pathogens. Proteases constitute one of the largest functional groups of proteins (Barrett, A.J., 1994), found in all living organisms, from viruses, bacteria and archaea, protozoans, metazoan, fungi to plants and animals (Hedstrom, L., 2002a; Barrett *et al.*, 2004).

Proteases play important roles in all biological functions. They are involved directly or indirectly in practically all biochemical and physiological processes in living organisms, such as cell growth, differentiation and death, that is apoptosis, cell nutrition, intra and extra cellular protein turnover (housekeeping and repair), cell migration and invasion, fertilization and implantation, among the most important. These functions extend from the molecular and cellular level to the organ and organism level to produce cascade systems, such as homeostasis and inflammation, and many other complex processes at all levels of physiology and pathophysiology (Abbenante and Fairlie, 2005)

Therefore, proteases are implicated in the whole life cycle of an organism: conception, birth, growing, maduration, ageing, disease and death. Deficiences or alterations in the regulation of these enzymes leads to abnormal development, disease and death, such as arthritis, cancer, inflammatory, immunological, respiratory, neurodegenerative and cardiovascular diseases, including parasitic, fungal, and viral infectious (Abbenante and Fairlie, 2005). Consequently, proteases are of great interest for the pharmaceutical industry as potential drug targets.

Classification: Proteases are subdivided into two major groups, such as, exopeptidases and endopeptidases, depending on their site of action. Exopeptidases cleave the peptide bond proximal to the amino or carboxy terminus of the substrate, whereas endopeptidases cleave internal sicissile bonds of the substrate. According to the functional group present at the active site and their mechanism of action, proteases can be classified as serine proteases, aspartic proteases, cysteine proteases and metalloproteases (Barrett *et al.*, 2004).

Based on the amino acid sequences, proteases are classified into different families (Rawlings and Barrett, 1993). A family is a group of peptidases that are evolutionarily related with similar primary structure. Each family of peptidases are assigned with a code letter denoting the type of catalysis, i.e., A, C, G, M, N, S, T or U for aspartic, cysteine, glutamic, metallo, asparagine, serine, threonine or unknown type, respectively. These families are further subdivided into "clans" to accommodate sets of peptidases that have a common ancestor (Rawlings and Barrett, 1993). In June 2012 there 52 clans and 228 families of proteases listed in the MEROPS database.

On the other hand, proteases have been classified according to their catalytic activity by the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (NC-IUBMB) (http://www.chem.qmul.ac.uk/iubmb/enzyme/), which has assigned a numerical classification for each enzyme, known as Enzyme Commission Number (EC number). The first number determines the enzyme class, whereas the second and third numbers indicate the subclass and sub-subclass, respectively. An example of the nomenclature for proteases is listed as follows:

Table 1. Classification of proteases according to the NC-IUBMB

EC number	Enzyme
3.4.1	a-Amino-acyl-peptide hydrolases (now in EC 3.4.11)
3.4.2	Peptidyl-amino-acid hydrolases (now in EC 3.4.17)
3.4.3	Dipeptide hydrolases (now in EC 3.4.13)
3.4.4	Peptidyl peptide hydrolases (now reclassified within EC 3.4)
3.4.11	Aminopeptidases
3.4.12	Peptidylamino-acid hydrolases or acylamino-acid hydrolases (now reclassified within EC 3.4)
3.4.13	Dipeptidases
3.4.14	Dipeptidyl-peptidases and tripeptidyl-peptidases
3.4.15	Peptidyl-dipeptidases
3.4.16	Serine-type carboxypeptidases
3.4.17	Metallocarboxypeptidases
3.4.18	Cysteine-type carboxypeptidases
3.4.19	Omega peptidases
3.4.21	Serine endopeptidases
3.4.22	Cysteine endopeptidases
3.4.23	Aspartic endopeptidases
3.4.24	Metalloendopeptidases
3.4.25	Threonine endopeptidases
3.4.99	Endopeptidases of unknown catalytic mechanism

A conceptual model was proposed by Schechter and Berger, 1997, in order to designate the interaction sites between enzyme and substrate. In this model, amino acid residues corresponding to the substrate are labelled with the symbol P, which interacts with the subsites of the active-site in the enzyme, identified as S. These residues are listed from the scissile peptide bond towards the N-terminal tail, as P1, P2, P3, etc., while those residues located towards the C-terminal tail are tagged as P1', P2', P3', etc. The subsites in the proteases interacting with the amino acid residues of the substrate are listed as S1, S2, S3 and S1', S2', S3', respectively (figure 1).

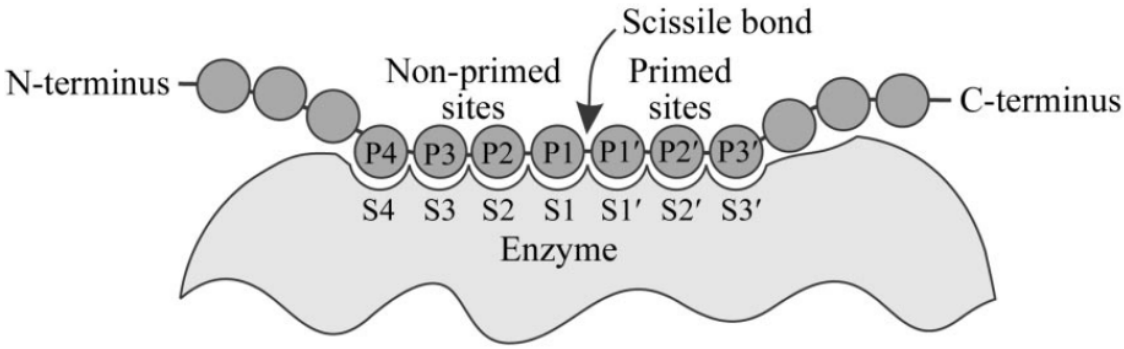


Figure 1. Schechter and Berger's definition of substrate-binding sites (figure from Turk and Guncar, 2003)

I.1.2. Protease inhibitors. General characteristics and classification

Protease inhibitors (PI) are usually polypeptides that inhibit the action of proteases. They are widely distributed in various tissues of animals, plants and microorganisms (Laskowski and Kato, 1980). Protease inhibitors are divided into 4 major mechanistic classes: inhibitors of serine proteases, inhibitors of cysteine proteases, inhibitors of aspartic proteases and inhibitors of metalloproteases, excepting α 2-macroglobulins which inhibit proteases of all classes (Hartley *et al.*, 1960).

Protease inhibitors belonging to these mechanistic classes have been further clasiffied into 48 families (Rawlings *et al.*, 2004) based on amino acid sequence homologies, topological relationships between disulfide bonds and the location of reactive sites (previosusly described by Laskowski and Kato, 1980).

Individual families are identified by a prominent member of that family without considering functional and biological source. According to this classification, inhibitor families include members that may inhibit proteases of different mechanistic classes, so called "cross class" inhibitors. The specificity of inhibitors largely relies on the exposed reactive-site loop (Laskowski and Kato, 1980; Laskowski *et al.*, 1987). In august 2012, 51 clans and 77 families of protease inhibitors have been listed in MEROPS database.

Generally, these molecules are transition state analogs and act as competitive inhibitors, having usually good class specificity, although it has been increased in the last years the number of PI described, able to inhibit proteases belonging to different mechanistic classes, but not all with the same efficiency (Rawlings *et al.*, 2004). However, the most striking kinetic feature of these inhibitors, derived from their biological function, is their capacity of tight binding to their target proteases, ensuring a high efficiency as inhibitors with values of inhibition constants (K_i) in the nanomolar range or even lower, such as BPTI, which has a K_i value against its target protease trypsin in the fentomolar range, considered one of the strongest interactions between proteins in nature (Lazdunski *et al.*, 1974; Fritz *et al.*,, 1983). Consequently, the inhibitors are efficient at concentrations within the enzyme concentration range and some basic principles of Michaelis Menten kinetics, such as those of the classical competitive inhibition, are not obeyed (Bieth, J.G., 1995).

Protease inhibitors play important basic roles, such as: to prevent proteolysis in sites where this activity must not occur and its regulation, in order to guarantee partial proteolysis as a physiological event, among those most important (Bode and Hubber, 2000). It should be mentioned that some protease inhibitors might play an important role in defense against prey and predator's proteases, such as in the case of marine invertebrates (Mebs and Gebauer, 1980; Shiomi *et al.*, 1985; Sencic and Macek, 1990; Castañeda *et al.*, 1995).

These functions have converted protease inhibitors into very attractive molecules, not only as valuable tools in the establishment of structure-function relationships, which have been their traditional application, but also due to their multiple biotechnological and biomedical applications (Zingali *et al.*, 1993; Hugli, 1996; Van Noorden, 1997; Birk, 2003). The former, allows increasing yields during the production processes of proteins, improving also their operational and storage stabilities. However, the demonstrations that the control of proteolysis constitutes a valid pharmacological tool and the success of some protease inhibitors in human therapy have determined a renewed interest in these molecules (Groneberg *et al.*, 1999; Hayashi *et al.*, 2001; Aghajanian *et al.*, 2002; Eriksson *et al.*, 2003; Lin *et al.*, 2003; Lorrain *et al.*, 2003; Braddock and Quinn, 2004; Quan *et al.*, 2004; Schiff *et al.*, 2004; Visconte *et al.*, 2004).

I.2. Proteases and their inhibitors

I.2.1. Metallocarboxypeptidases and their inhibitors

I.2.1.1. Metallocarboxypeptidases

Natural evolution frequently generated a large adaptative variety of forms among protein functional families, and metallocarboxypeptidases (MCPs) followed this trend. Such enzymes are exopeptidases that catalyze the hydrolysis of peptide bonds at the C-terminus of peptides and proteins. They belong to the catalytic classes of either the most MCPs (clan MC, family M14) or some serine proteases (clan SC, family S10) (Barrett *et al.*, 2004; Rawlings *et al.*, 2010) and their action causes strong effects in the biological activity of their peptide and protein substrates (Skidgel R., 1996).

MCPs from the family M14, including those from animals, plants and bacteria, have been divided into four subfamilies based on structural similarity and sequence homology: M14A (also termed as A/B subfamily), M14B (N/E subfamily), M14C, and the M14D subfamily, also named as Nna-like proteins or cytosolic carboxypeptidases (CCPs) and described by Rodriguez de la Vega *et al.*, 2007.

M14A subfamily: Members of the A/B subfamily appear to be soluble proteins that are normally synthesized as inactive precursors containing a preceding signal peptide of 15-22 residues. The pro-enzymes of the A/B subfamily are known as procarboxypeptidases (ProCPs) and contain a 90-95 residues at their N-terminal pro-region that folds in a globular independent unit, the "activation segment", linked to the enzyme through a connecting part (Coll *et al.*, 1991; Guasch *et al.*, 1992). Furthermore, these carboxypeptidases contain a catalytic domain of approximately 300 amino acid residues. Carboxypeptidase A1 (CPA1), Carboxypeptidase A2 (CPA2), and Carboxypeptidase B (ProCPB) are synthesized as inactive zymogens (ProCPA1, ProCPA2, and ProCPB) in the pancreatic cells until they are secreted into the digestive tract and subsequently limited tryptic digestion in the duodenum (Arolas *et al.*, 2007). The dramatic difference in specificity between CPB-like and CPA-like enzymes has been attributed to the residue in position 255.In the A-like enzymes, the presence of a hydrophobic residue in this position (Ile, Val, Leu, and Met) produces the specificity for cleavage of aliphatic/aromatic residues, while the presence of an acidic residue (Asp) in CPB and CPU produces the specificity for cleavage of basic residues (Vendrell *et al.*, 2004).

- **Carboxypeptidase A1:** CPA1 displays cleavage specificity for aliphatic and smaller aromatic residues (Gardell *et al.*, 1988; Pascual *et al.*, 1989; Arolas *et al.*, 2007). Three-dimensional structure of hCPA1 revealed that this enzyme, like its other non-human homologues, exhibits a tertiary fold corresponding to the α/β hydrolase fold and is formed by a central mixed paraller/antiparallel 8-stranded β-sheet over which 8 α-helices pack on both sides to form a globular molecule (Pallarès *et al.*, 2008).
- **Carboxypeptidase A2:** This enzyme shows a preference by the bulky side chain of Trp (Gardell *et al.*, 1988; Pascual *et al.*, 1989; Arolas *et al.*, 2007). Three forms of monomeric procarboxypeptidases, A1, A2 and B occur in human pancreas, together with a binary complex between the A1 form and proproteinase E (Reverter *et al.*, 1997). CPA2 isoforms have also been reported in rat extra-pancreatic tissues such as brain, testis, and lung; these CPA2 isoforms are shorter and have a distinct role from the pancreatic isoform (Normant *et al.*, 1995b; Reverter *et al.*, 1997).

Other relevant members of this subfamily found in humans have been described by Arolas *et al.*, 2007 and they are listed as follows:

- **Carboxypeptidase A3:** This enzyme, also known as mast cell CPA is found in the secretory granules of mast cells mainly in its active form, in complex with proteoglycans (Springman *et al.*, 1995). The precise role of this carboxypeptidase still needs to be defined, although it presumably functions in anaphylactic and inflammatory processes following the action of chymase (Springman *et al.*, 1995; Lundequist *et al.*, 2004; Metz *et al.*, 2006).

- **Carboxypeptidase A4:** CPA4 in its pro-form (hPCPA4) has been found as a gene product, involved in prostate cancer (Huang *et al.*, 1999). The hPCPA-4 gene is imprinted and may be responsible for prostate cancer aggressiveness (Kayashima *et al.*, 2003). A recent study of substrate specificity performed by Tanco *et al.*, 2010 showed that this enzyme was able to hydrolyse hydrophobic C-terminal residues with a preference for Phe, Leu, Ile, Met, Tyr, and Val. Some peptides such as neurotesin, granin and opioid peptides were transformed by the CPA4. These peptides were previously described to function in cell proliferation and differentiation. Thus, a potential link between CPA4 and neuropeptide processing, cancer aggressiveness as well as regulation in the extracellular environment was suggested by these authors.
- **Carboxypeptidase A5:** It has been suggested that this enzyme play a role in the processing of pro-opiomelanocortin-derived peptides, and possibly other peptides that undergo C-terminal removal of aliphatic or aromatic residues, which are specifically hydrolyzed by the CPA-5 (Wei *et al.*, 2002; Wei *et al.*, 2003).
- **Carboxypeptidase A6:** CPA6 gene has been linked to Duane syndrome, a disorder of the visual system that results from abnormal innervation of the optic muscle during embryogenesis (Pizzuti *et al.*, 2002). In addition, this enzyme is also found in several other brain regions as well as in a variety of tissues during embryogenesis (Fontenele *et al.*, 2005). Kinetic parameters were determined using a panel of synthetic carboxypeptidase substrates, indicating a preference of CPA6 for large hydrophobic C-terminal amino acids and only very weak activity toward small amino acids and histidine (Lyons and Fricker, 2010).
- **Carboxypeptidase O:** CPO is membrane attached via a phosphatidylinositol glycan (GPI) anchor and is found on the apical surface of intestinal epithelial cells. This enzyme is able to cleave proteins and synthetic peptides with greatest activity toward acidic C-terminal amino acids, unlike other CPA-like enzymes. CPO displays a neutral pH optimum and is inhibited by common metallocarboxypeptidase inhibitors as well as citrate (Lyons and Fricker, 2011).
- **Carboxypeptidase U:** CPU, commonly named as TAFI (Thrombin-Activable Fibrinolysis Inhibitor) is a relevant protease belonging to the A/B subfamily, which cleaves the basic amino acids lysine and arginine from the C terminus of a variety of peptides and proteins (Skidgel R.A., 1996). Activated TAFI plays a key role in the regulation of fibrinolysis, thus contributing to hemostatis (Bouma and Mosnier, 2003; Hendriks D.F., 2004; Willemse and Hendriks, 2007; Rijken and Lijnen, 2009; Marx P.F. 2004). It removes C-terminal lysine residues from the fibrin clot following the endoproteolytic action of plasmin at late stages of coagulation, thus preventing fibrinolysis from entering the propagation phase. Therefore, this enzyme has become in an important target for thrombotic diseases (Sanglas *et al.*, 2008).

M14B subfamily: Proteases of M14B subfamily are produced as active enzymes and rely on subcellular localization in order to prevent inappropriate cleavages that would otherwise damage the cell. Instead of the pro-segment, these proteins contain a transthyretin-like subdomain (80-90 residues) at the C-terminus of the catalytic domain. Members of this subfamily contain other domains and even repeats of the carboxypeptidase domain (Vendrell *et al.*, 2004; Rodriguez de la Vega *et al.*, 2007). Carboxypeptidases E, N, M, Z, and D represent the five active enzymes described so far in this subfamily.

- **Carboxypeptidase E:** CPE, also known as enkephalin convertase or Carboxypeptidase H, was found in the biosynthesis pathway of a wide range of peptide hormones and neurotransmitters. For instance, CPE removes C-terminal Lys and Arg residues from intermediates formed by the action of prohormone convertases 1 and 2 on the peptide precursor (Fricker and Leiter, 1999). CPE has been proposed to have a function in the trafficking of pro-hormones into the regulated secretory pathway, although this function remains controversial (Cool *et al.*, 1997).

- **Carboxypeptidase N:** CPN circulates in plasma together with an 83 kDa binding protein, and removes basic residues from the C-terminus of various blood proteins/peptides. CPN removes the C-terminal Arg residue from bradykinin, which alters the affinity of this peptide for the two distinct bradykinin receptors. In addition, CPN has been proposed to cleave C-terminal basic amino acids from anaphylatoxin C3a, C4a, and C5a (Skidgel and Erdos, 1998).
- **Carboxypeptidase M:** CPM is a glycosylated monomer consisting of a single amino acid chain, which is bound on the cell surface through a GPI anchor to the plasma cell membrane (Shimamori *et al*., 1990; Deddish *et al*., 1990; Skidgel *et al*., 1996; Tan *et al*., 2003). Release of CPM from the membrane by phosphatidylinositol phospholipase C upregulated the enzyme synthesis maintaining constant levels of CPM on the cell surface (McGwire *et al*., 1999). In agreement with the expected apical localization of GPI anchored and N-substituted glycoproteins, CPM was localized on the apical surface of Madine–Darby canine kidney cells (Deddish *et al*., 1990). However, up to approximately half of the CPM protein was found on the basolateral domain of the polarized cells (Li *et al*., 1999).
- **Carbozypeptidase Z:** CPZ was discovered during a search for novel carboxypeptidases that could compensate for defective carboxypeptidase E in*Cpe* ^{fat} */Cpe* ^{fat} mice (Song and Fricker, 1997; Fricker and Leiter, 1999). CPZ contains a cysteine-rich domain that has amino acid sequence identity with proteins that bind Wnt/wingless, such as the frizzled receptors, frzb, sizzled, and several other proteins. In addition, this enzyme contains an N-terminal domain that has amino acid sequence similarity to Wnt-binding proteins (Novikova *et al*., 2000).
- **Carboxypeptidase D:** CPD is a 180-kDa protein that contains three carboxypeptidase-like domains, a transmembrane domain, and a cytosolic tail, which functions in the processing of proteins that transit the secretory pathway (Sidyelyeva and Fricker, 2002). Putative substrates include growth factors and receptors that are produced from larger precursors by cleavage at basic amino acid-containing sites (Fricker, L.D., 1998). Human, rat, mouse, and duck CPD contain three carboxypeptidase-like domains followed by a transmembrane domain and a 58-residue cytosolic tail (Kuroki *et al*., 1995; Tan *et al*., 1997; Ishikawa *et al*., 1998). Of the three carboxypeptidase-like domains, only the first two have enzyme activity toward standard substrates (Novikova *et al*., 1998; Novikova *et al*., 1999).

M14C subfamily: The M14C subfamily comprises the bacterial orthologs, which includes -D-glutamyl-(L)-meso-diaminopimelate peptidase I from *Bacillus sphaericus* (Hourdou *et al*., 1993).

M14D subfamily: This subfamily was recently classified (Rodriguez de la Vega *et al*., 2007), and their members are also known as cytosolic carboxypeptidases (CCPs). Nna1 (nervous system nuclear protein induced by axotomy) or CCP1 was the first gene product described. Five additional Nna1-like genes were identified in the mouse genome and named CCP2 through 6. All CCPs are abundant in testis and also expressed in brain, pituitary, eye, and other mouse tissues. In brain, Nna1/CCP1, CCP5, and CCP6 are broadly distributed, whereas CCP2 and 3 exhibit restricted patterns of expression. Nna1/CCP1, CCP2, CCP5, and CCP6 were found to exhibit a cytosolic distribution, with a slight accumulation of CCP5 in the nucleus (Kalinina *et al*., 2007).

Catalytic mechanism of MCPs– Bovine CPA1 is one of the most thoroughly studied enzymes in Biochemistry, and it has been used in the study of the catalytic mechanism for MCPs. However, its pathway of catalysis has been a controversial topic during the last 50 years. Most crystallographic and kinetic data currently support the promoted-water model, in which the zinc ion of CPA-1 is a classical electrophylic catalyst that provides electrostatic stabilization for the negatively charged intermediates generated during the hydrolysis and promotes a water molecule as a potent nucleophile to attack the scissile peptide bond of the substrate, leading to a tetrahedral transition state (Fife and Przystas, 1986; Alvarez-Santos *et al*., 1998).

A detailed description of the catalytic mechanism for MCPs has been made by Vendrell *et* al., 2004, in which the nucleophilicity of the metal-bound water is enhanced by a hydrogen bond with Glu270 which acts as the general acid-base catalyst. While Glu270 abstracts a proton from the water molecule, Arg127, which polarizes the peptide bond prior to hydrolysis, also stabilizes the tetrahedral transition state (Christianson and Lipscomb, 1989). The zinc binding residues are His69, Glu72 and His196, with the catalytic water molecule also affording a coordination bond (figure 2). From crystallographic studies, amino acids Asn144, Arg145, Tyr248 in S1' and Arg127 and Glu270 in S1 have been defined as important for substrate binding and catalysis and are shared by all enzymatically active members of the metallocarboxypeptidases family. The terminal carboxylate group of the peptide substrate is fixed by Asn144, Arg145 and Tyr248, while the carbonyl group of the scissile peptide bond becomes positioned near Glu270, Arg127 and zinc. Differences are restricted to the positions that define subsites S2, S3 and S4 and to amino acids defining specificity for a C-terminal side-chain. The amino acid residue at position 255 appears to be the main determinant of specificity. It is an Ile in CPA1 and A2 and an Asp in CPB, allowing for the binding of hydrophobic and basic residues, respectively.

The enzimatically active members of the N/E subfamily do not use the residue in the position equivalent to residue 255 of bovine CPA to provide the specificity for basic C-terminal residues (Gomis-Rüth *et al.*, 1999). This residue is either a Gln (CPE, CPD domains 1 and 2, CPM, CPN) or Ser (CPZ). Based on the crystal structure, the role of binding to the positive charge of the substrate's side chain in CPD domain 2 is carried out by Asp192, at a position equivalent to Gly207 in CPA (Gomis-Rüth *et al.*, 1999; Aloy *et al.*, 2001).

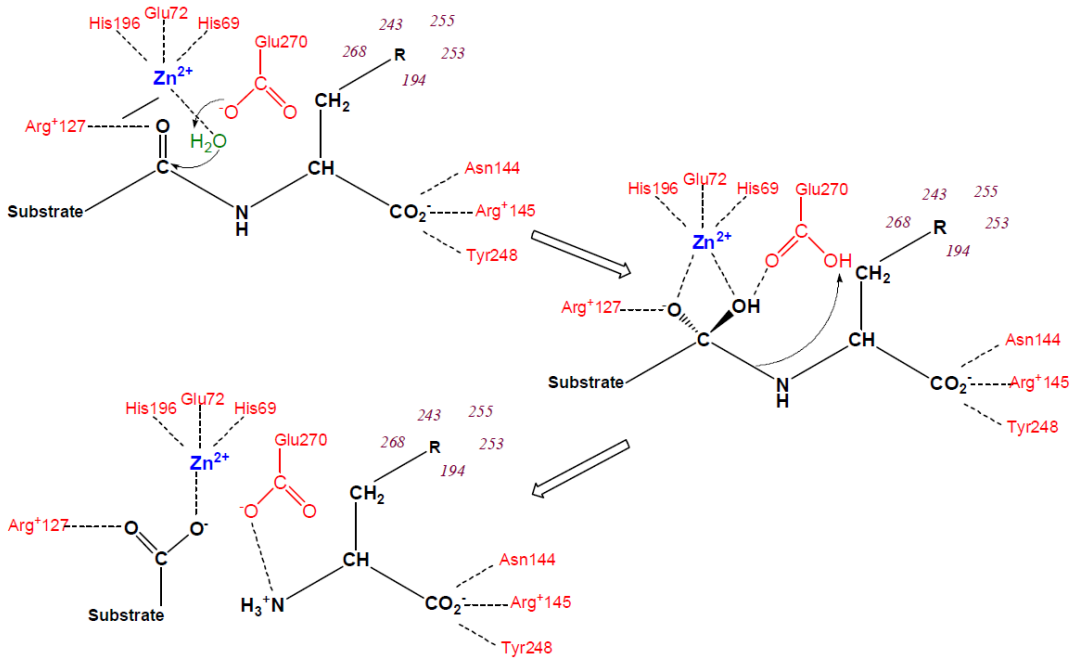


Figure 2. Catalytic mechanism of metalloproteases as exemplified by bovine CPA-1 (figure from Vendrell *et al.*, 2004)

Metallocarboxypeptidase like enzymes have been isolated from different sources, mainly from vertebrates, but a few of them from marine invertebrate organisms (Arolas *et al.*, 2007; Vendrell *et al.*, 2004; Rawlings *et al.*, 2004), thus, the digestive crayfish carboxypeptidase (CPB) (Zwilling *et al.*, 1979), the CPE like enzyme from the sea hare *Aplysia californica,* that carries out important regulatory function in this organism (Juvvadi *et al.*, 1997), two carboxypeptidases (A & B types) from the hepatopancreas of the crab *Paralithodes camtschatica* (Sakharov and Prieto, 2000), the CPA like protease from squid hepatopancreas of *Illex illecebrosus* (Raksakulthai and Haard, 2001), the CPB like isolated from the pyloric ceca of the starfish *Asterias amurensis* (Kishimura and Hayashi, 2002) and the CP like isolated from the marine annelid *Sabellastarte magnifica* (Alonso del Rivero *et al.*, 2009).

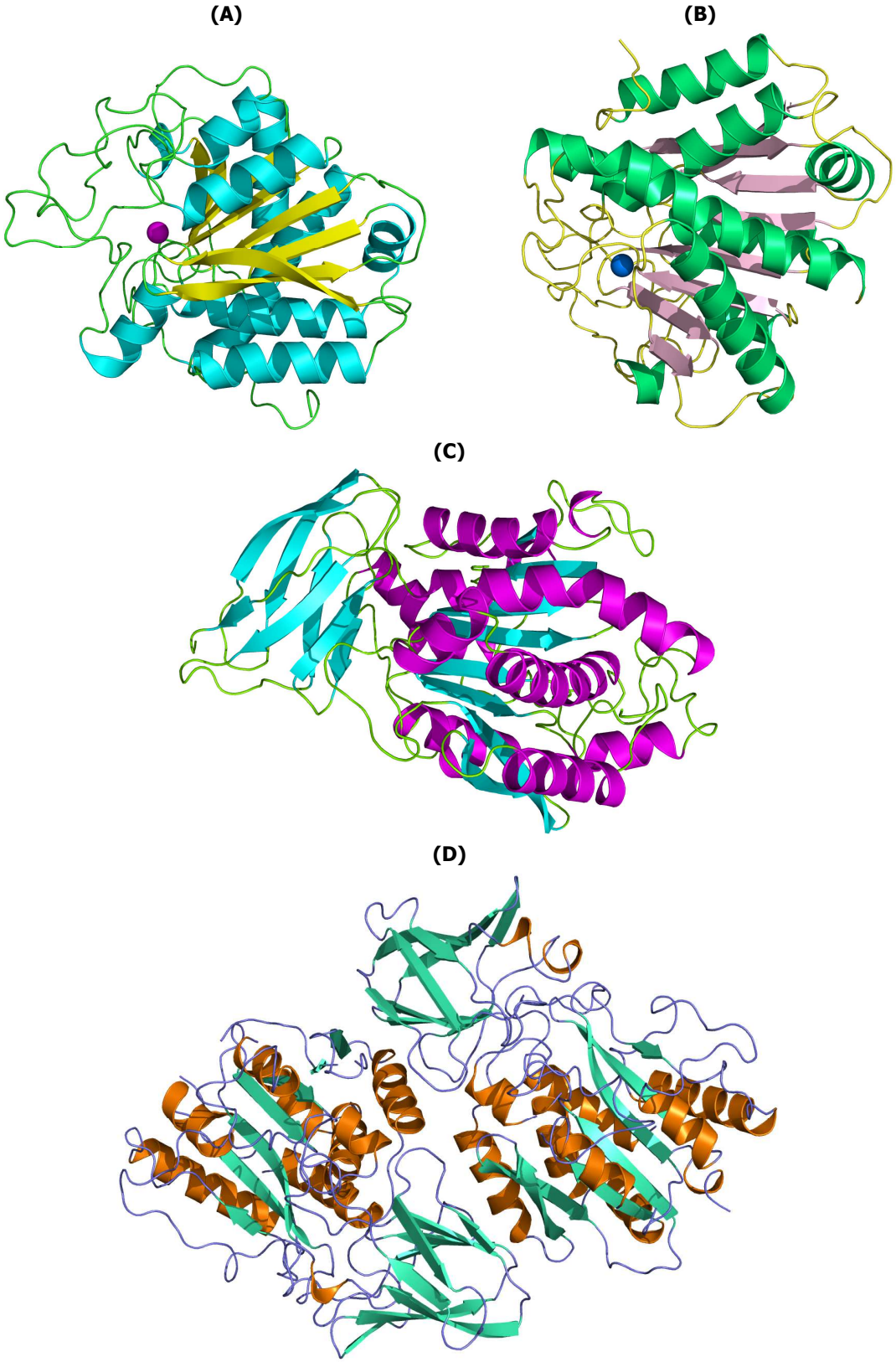


Figure 3. Three-dimensional structure of various MCPs.
(A) Bovine CPA. PDB code: 3I1U (Fernández *et al.*, 2010). **(B) Porcine CPB.** PDB code: 2PJ9) (Adler *et al.*, 2008). **(C) Human CPN.** PDB code: 2NSM (Keil *et al.*, 2007). **(D) Putative CPA from *Shewanella denitrificans os-217*.** PDB code: 3L2N (Joint Center for Structural Genomics, JCSG, 2009)

Matrix metalloproteinases (MMPs) play a key role in remodeling and extracellular matrix degradation. MMPs are important enzymes involved in cancer formation, development and metastasis. Therefore, MMPs inhibitors are potencial antimetastatic and antiangiogenic agents (Fujita *et* al., 2003; Krizkova *et al*., 2011). Membrane type 1 matrix metalloproteinase (MT1-MMP) is one of the key enzymes involved in tumor growth, migration, angiogenesis, invasion, and metastasis (Cussens and Werb, 1996; Woessner, J.F.Jr., 1999). This enzyme plays important roles in degradation of basal membranes and extracellular matrices (ECM) by digesting type IV collagen, a major component of ECM.

Metallocarboxypeptidases represent new potential drug targets due to their diversity of functions, involved in allergic and inflammatory responses, cardiovascular diseases and cancer, among others (Reynolds *et al.,* 1989). One of the most attractive is TAFI. Inhibitors of this enzyme could be valuable fibrynolytic agents (Eaton *et al.,* 1991; Bajzar *et al.,* 1995). Recently, a novel CP-like was detected in *Plasmodium falciparum*, the most aggressive causative agent of malaria (Rodrgiguez de la Vega *et al*., 2007), converting this enzyme in a potential drug target for malaria.

I.2.1.2. Metallocarboxypeptidase inhibitors

Unlike endoproteases for which numerous examples of proteinaceous inhibitors have been reported, metallocarboxypeptidase inhibitors are somewhat limited. They have only been found in *Solanacea,* tomato and potato (PCI) (Ryan *et al*., 1974; Hass *et al*., 1975; Hass and Hermodson, 1981), the intestinal parasite *Ascaris suum* (ACI) (Homandberg *et al*., 1989), the medical leech *Hirudo medicinalis* (LCI) (Reverter *et al*., 1998), the tick *Rhipicephalus bursa* (TCI) [Arolas *et al.* (a), 2005]*, Haemaphysalis longicornis* (HITCI) (Haiyan *et al*., 2007) and in rat and human tissues (latexin) (Normant *et al*., 1995; Lui *et al*., 2000; *Pallares et al*., *2005*). These carboxypeptidase inhibitors differ in their target proteases; whereas the inhibitor from mammalian tissues apparently binds to endogenous carboxypeptidases, the rest interact with exogenous enzymes, i.e. carboxypeptidases of a distinct organism. No proteinaceous N/E MCP inhibitors have been found to date (Arolas *et al*., 2007).

Potato carboxypeptidase inhibitor (PCI): PCI, a polypeptide consisting of 39 residues (molecular mass of 4.2 kDa), is the most extensively studied proteinaceous MCP inhibitor. This molecule is a tight-binding inhibitor against hCPA1 (K_i = 1.6 nM), hCPA2 (K_i = 8.8 nM), hCPB (K_i = 1.8 nM), hTAFI (K_i = 5.3 nM), among other A/B-type MCPs (Hass *et al*., 1975; Arolas *et al*., 2005a). PCI is organized in a 27-residue globular core, 7 residues at its N-terminal tail as well as 5 residues at its C-terminus sequence. The core is stabilized by the presence of three disulfide bridges (Cys8-Cys24, Cys12-Cys27, and Cys18-Cys34). PCI displays a scaffold called T-knot and this molecule contains few regular secondary structure elements such as a short 3/10 helix, and a very small antiparalell β-sheet (figure 4A).

Tomatos (*S. lycopersicum*) contain a metallocarboxypeptidase inhibitor with an amino acid sequence highly similar to that of PCI (70% identity at the protein level, 85% at DNA level, and strong immunological crossreactivity) (Blanco *et al*., 1998). The function of potato and tomato inhibitors is probably related to plant defense mechanisms (Ryan C.A., 1989).

Carboxypeptidase inhibitor from *Ascaris suum* (ACI): ACI was initially isolated from roundworms of the genus *Ascaris* by Homandberg *et al*., 1989. Later, an isoform of ACI was isolated and characterized by Sanglas *et al*., 2009. This isoform consists of 67 residues and did not show significant homology with any known protein. These domains are linked by an α-helical segment and a fifth disulfide bond. Recombinant ACI displays inhibition constants in the nanomolar range against A/B-type MCPs such as hCPA1 (K_i = 1.6 nM), hCPA2 (K_i = 2.5 nM), hCPA4 (K_i = 23.9 nM), hTAFI (K_i = 42.0 nM), among others (Sanglas *et al*., 2009). The crystal structure of ACI showed a protein with a fold consisting of two tandem homologous domains, each containing a β-ribbon and two disulfide bonds (Sanglas *et al*., 2009).

Tick carboxypeptidase inhibitor (TCI): TCI is a 7.9 kDa protein that contains 75 amino acid residues (Arolas *et al.*, 2005a). A 22-residue hydrophobic signal peptide has been identified in this protein. TCI, as the MCP inhibitors described above, is a tight-binding and competitive inhibitor of A/B-type MCPs such as bCPA (K_i = 1.1 nM), hCPA1 (K_i = 1.2 nM), hCPA2 (K_i = 3.6 nM), pCPB (K_i = 1.6 nM), among others (Arolas *et al.*, 2005b). Three-dimensional structure of TCI revealed that this molecule consists of a short α-helix followed by a small twisted antiparallel β-sheet, showing high structural homology to proteins of the β-defensin-fold family (Arolas *et al.*, 2005b) (figure 4B).

Leech carboxypeptidase inhibitor (LCI): LCI is a cysteine-rich polypeptide composed of 66 amino acid residues. It does not show sequence similarity to any other protein except at its C-terminal tail. In this region, the inhibitor shares the amino acid sequence -Thr-Cys-*X*-Pro-Tyr-Val-*X* with *Solanacea* carboxypeptidase inhibitors, suggesting a similar mechanism of inhibition where the C-terminal tail of the inhibitor interacts with the active-site of Mcps in a substrate-like manner. LCI is a tight-binding and competitive inhibitor of A/B-type MCPs such as bCPA (K_i = 0.25–0.48 nM), pCPB (K_i = 0.27–0.51 nM), hCPA2 (K_i = 0.17–0.78 nM) and hTAFI (K_i = 0.1–0.2 nM) (Reverter *et al.*, 1998). In addition, the structure of LCI defines a new protein motif that comprises a five-stranded antiparallel β-sheet and one short α-helix (Reverter *et al.*, 2000) (figure 4C).

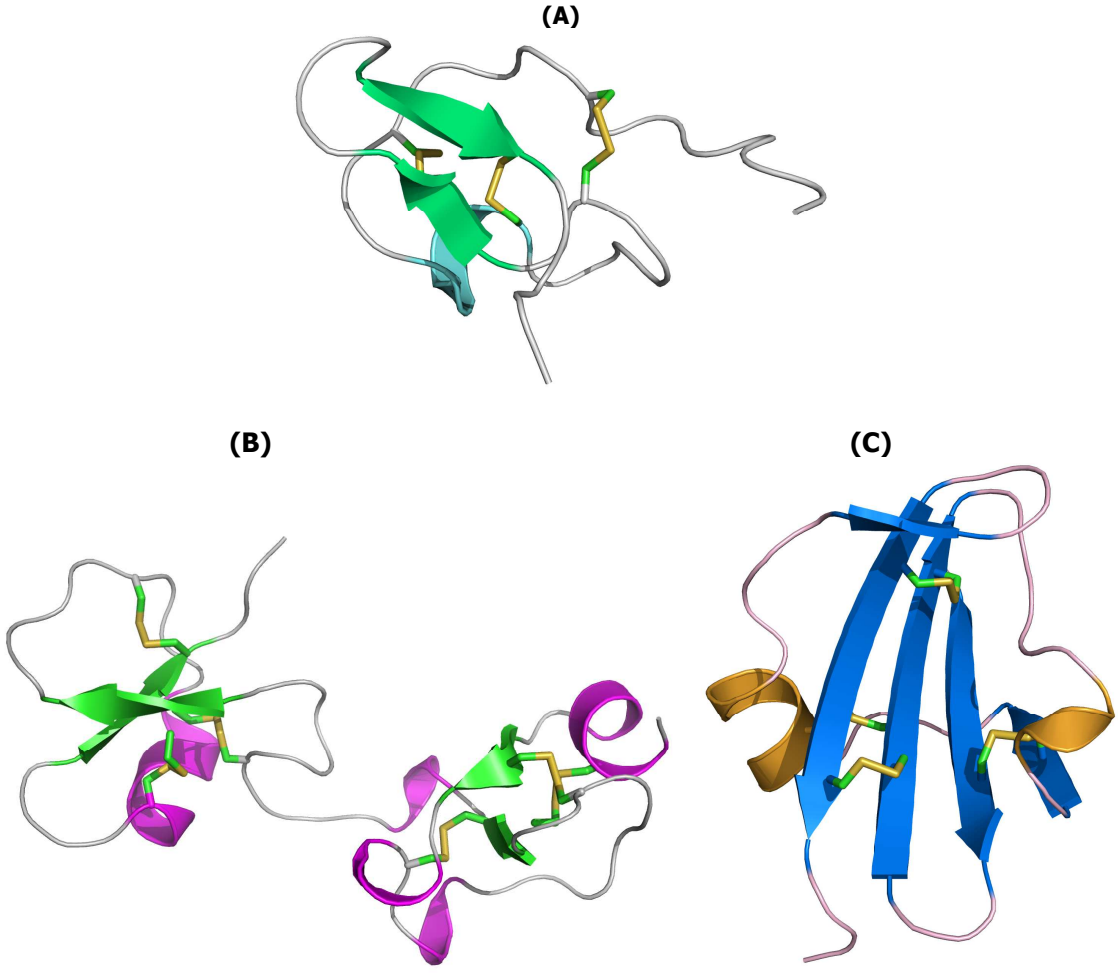
Bifunctional inhibitor of carboxypeptidases A and serine proteases from *Sabellastarte magnifica* (SmCI): this bivalent inhibitor was isolated from the tentacle crown of the sea annelide *S. magnifica*. SmCI is a 165 residues glycoprotein with a molecular mass of 19.7 kDa and three BPTI Kunitz domains. The three-dimensional models of the 3 domains are typical for members of this family. This molecule is able to inhibit several A-type MCPs with a K_i value of 1.3×10^{-8} M against bovine CPA but not B-type MCPs. In addition, SmCI displays a strong inhibitory activity against bovine trypsin and pancreatic elastase with K_i values of 3.6 and 8.8 nM, respectively (Alonso del Rivero *et al.*, 2012).

Endogenous carboxypeptidase inhibitor: This inhibitor also known as tissue carboxypeptidase inhibitor or latexin is the largest proteinaceous carboxypeptidase inhibitor reported. Its cDNA encodes a 233-amino acid residues (MW 26 kDa) protein present in nonpancreatic tissues (e.g., brain, lung, digestive tract). Its sequence does not present significant homology with the rest of reported carboxypeptidase inhibitors other than a limited degree of similarity between a short segment of its sequence and the activation segment of porcine procarboxypeptidase B. Latexin is a hardly reversible, non-competitive and potent inhibitor of hCPA1 (K_i = 1.6 nM), hCPA2 (K_i = 3.5 nM), hCPB (K_i = 1.1 nM), hTAFI (K_i = 1.8 nM), among other A/B-type MCPs (Pallarès *et al.*, 2005). It lacks a signal peptide, a fact that suggests a cytosolic localization. Therefore a rather general functional role, such as the control of cytosolic protein degradation was suggested (Normant *et al.*, 1995). A latexin homologue has been found in humans, which consists of 222 amino acids (Liu *et al.*, 2000). Three-dimensional structure of latexin shows two structurally related domains linked by a connecting helix. Each domain comprises an extended α-helix followed by a strongly twisted four-stranded antiparallel β-sheet of simple up-and-down connectivity that embraces the helix establishing hydrophobic contacts (Pallarès *et al.*, 2005) (figure 4D).

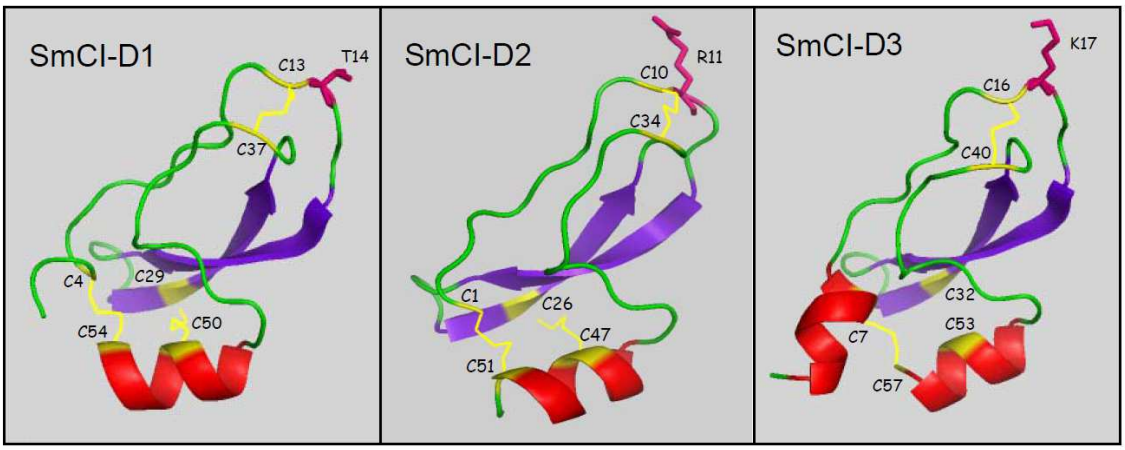
Mechanism of MCPs inhibition– The mechanism of inhibition of PCI, ACI, LCI and TCI relies on the interaction of their C terminus tail with the active site groove of the carboxypeptidase in a way that mimics substrate binding (Vendrell *et al.*, 2000). Although there is no sequence similarity between all these inhibitors, structural comparisons indicates that the C-terminal tail of the inhibitor displays a similar conformation and the last residue is coordinated with the zinc atom of the active-site of the carboxypeptidase. Thus, even though these exogenous carboxypeptidase inhibitors are isolated from evolutionary distant organisms, this is a good example of convergent evolution dictated by the architecture of the active site of the enzyme.

Additionally, TCI anchors to the surface of A/B-type MCPs in a double-headed manner not seem for the other inhibitors, where the last 3 residues of the C-terminal tail in the inhibitor interact with the active-site of the enzyme in a similar manner displayed for the above mentioned inhibitors, and its N-terminal domain binds to an exosite distinct from the active-site groove (Arolas *et al.*, 2005b).

On the other hand, the C terminal tail of latexin does not seem to be a suitable substrate for carboxypeptidases (Reverter *et al.*, 1998). This endogenous inhibitor interacts with the enzyme through a large surface, but exhibits only a few intermolecular contacts. The main interaction area includes the lower barrel surface of latexin around the β-sheet located at its C-terminal domain (figure 4D). The active-site blocking involves an inhibitory loop located in the central part of the mentioned β-sheet (Pallarès *et al.*, 2005). In addition, the propeptides of procarboxypeptidases, which fold as independent globular domains, position their inhibitory loop on the top of the active site cleft of the enzyme keeping the enzymes inhibited (Guasch *et al.*, 1992; Chen *et al.*, 1996; Muller *et al.*, 2002).



(D)



(E)

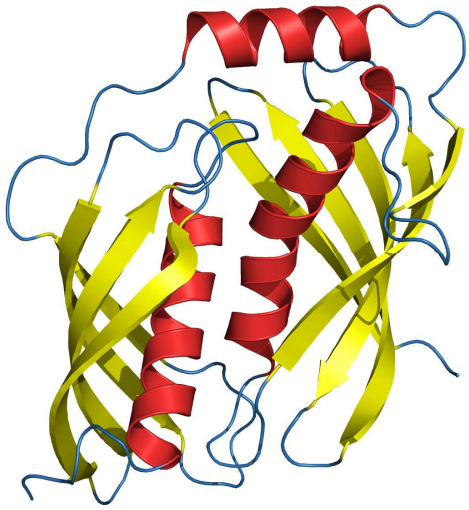


Figure 4. Three-dimensional structure of various proteinaceous MCPs inhibitors. **(A) PCI.** PDB code: 1H2O (González *et al.*, 2003). **(B) TCI.** PDB code: 1ZLH (Arolas *et al.*, 2005b). **(C) LCI.** PDB code: 1DTD (Reverter *et al.*, 2000). **(D) SmCI.** There-dimensional modelling performed by Alonso del Rivero *et al.*, 2012. **(E) Human latexin.** PDB code: 2BO9 (Pallarès *et al.*, 2005).

I.2.2. Aspartic proteases and their inhibitors

I.2.2.1. Aspartic proteases

Aspartic proteases, also known as acid proteases are a group of enzymes that involve two highly-conserved aspartic acid residues at the active site. Members of this type of proteases are widely distributed in different organisms: plants, fungi, invertebrates, vertebrates, bacteria, and retroviruses (Davies, D.R., 1990; Dunn, M.B. 2002). Four major groups of aspartic proteases have been identified in vertebrates, namely pepsins, cathepsins D, cathepsins E and renins (Barrett, A.J., 1979). Pepsins have not been detected in invertebrates, and apparently cathepsin D, as well as other cathepsins act both as digestive and lysosomal enzymes in many of these animals (Gildberg, A., 1988).

According to the MEROPS database (http://www.merops.ac.uk), created by Rawlings, N.D., and Barrett, A.J., 1999, 16 different families of aspartic proteases has been grouped, based on the similarity on their amino acid sequences. Furthermore, six clans for this type of proteases have been assembled based on their evolutionary relationship and tertiary structure.

Most of the mammalian, plant, and fungal aspartic proteases are single chain enzymes with a molecular weight of approximately 35 kDa. Porcine pepsin was the first to be sequenced (Moravek, L., and Kostka, V. 1974; Tang *et al.*, 1973) and it is the prototype enzyme of the family A1 (figure 6A). These enzymes are approximately 327 amino acids long, with approximately 5% sequence identity between all members of the family (Davies, D.R., 1990). The aspartic proteinases have characteristic sequences in the region of the two catalytic aspartyl residues: (hydrophobic, generally Phe) Asp32-Thr-Gly-Ser in the N-terminal domain, and a corresponding (hydrophobic) Asp215-Thr-Gly-Ser/Thr in the C-terminal domain (pepsin numbering) (Davies, D.R., 1990).

In the mechanism of aspartic proteases described by Polgar L., 1987, it has been proposed that the two catalytically competent carboxyl groups of aspartic proteases constitute a functional unit which mediates the proton from the attacking water molecule to the leaving nitrogen atom of the substrate. Protonation of this nitrogen atom has been the main issue of the previous mechanistic proposals. The first step of the present mechanism involves proton transfer from the water to the aspartic diad and concurrently another proton transfer from the diad to the carbonyl oxygen of the scissile peptide bond. These proton transfers provide the driving force for the bond formation between the substrate and water, which leads to the formation of a tetrahedral intermediate. The intermediate breaks down to products by a similar facilitation, i.e. by concerted general acid-base catalysis, which involves simultaneous proton transfers from the intermediate to the diad and from the diad to the leaving nitrogen of the substrate. The symmetrical mechanism of the formation and decomposition of the tetrahedral adduct resembles that found in the serine protease catalysis (figure 5).

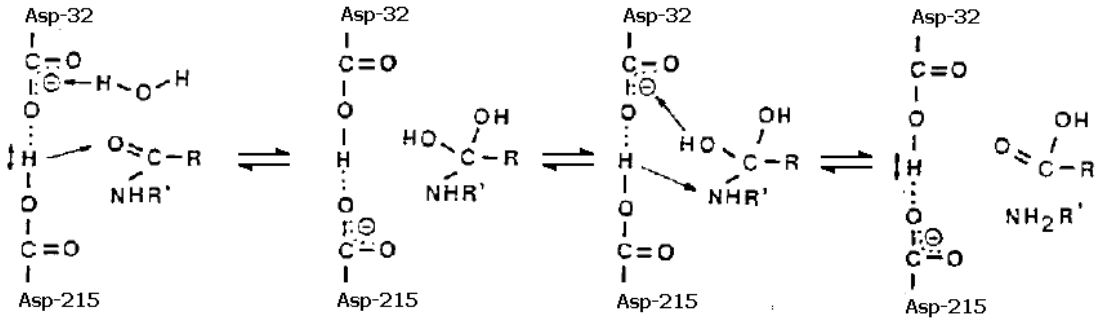


Figure 5. Proposed catalytic mechanism of aspartic proteases (figure from Polgar L., 1987)

Aspartic proteases are involved in cancer, Alzheimer's disease, hypertension, AIDS, etc. (Abbenante and Fairlie, 2005). These enzymes primarily function in the degradation of intracellular and endocytosed proteins. Among these, cathepsins D and E have general proteolytic activities and have a wide tissue distribution (Boldbaatar *et al.*, 2006).

Aspartic proteases function as digestive enzymes in some pathogens including *Plasmodium falciparum* (Goldberg *et al.*, 1991). In this pathogen, these enzymes are known as Plasmepsins (Plm), which have high sequence homology with human cathepsin D (Abbenante, and Fairlie, 2005). Sequencing of the *P. falciparum* genome has identified 10 plasmepsins encoding genes, numbered Plm-I to Plm-X (Coombs *et al.*, 2001; Gardner *et al.*, 2002). Among these, only Plm-I, Plm-II, HAP (histoaspartic protease, or Plm-III), and Plm-IV are active in the acidic food vacuole, where take place the hemoglobin degradation by the parasite (Gil *et al.*, 2011). These enzymes degrade human hemoglobin, the major source of nutrients for the parasite, cleaving Phe33-Leu34 of the α-chain of hemoglobin before degrading the polypeptides to smaller fragments that are processed by other enzymes such as falcipain (Francis *et al.*, 1997).

Another relevant member of the aspartic proteases is the HIV-1 retropepsin (human immunodeficiency virus 1), which is an important drug target for human immunodeficiency virus infection (Wlodawer, A., and Vondrasek, J., 1998). HIV-1 retropepsin cleaves the *gag-pol* polyprotein contained in the retrovirus, at eight sites, mostly with hydrophobic residues in P1 and P1' (Dunn and Rao, 2004). Retropepsins contain mainly β-secondary structural elements and they are active only as their homodimers (figure 6B). Taking into account this catalytic feature, in addition to the conventional inhibitors, there are experimental compounds that interfere with the dimerization that is necessary to form the active peptidase (Lee and Chmielewski, 2006).

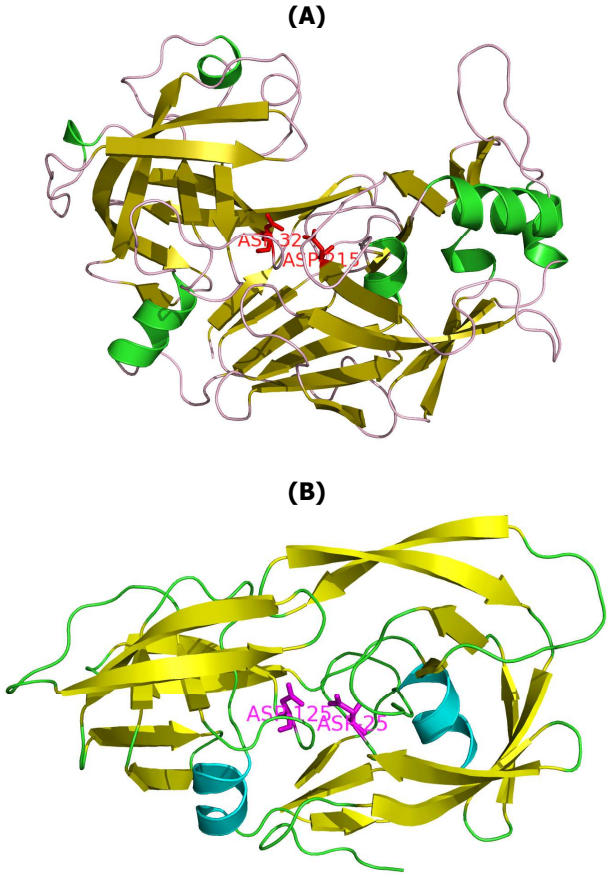


Figure 6. Three-dimensional structure of various aspartic proteases.
(A). Porcine pepsin. PDB code: 4PEP (Sielecki *et al.*, 1990). **(B) HIV-1 protease.** PDB code: 1EBZ (Andersson *et al.*, 2003)

I.2.2.2. Aspartic protease inhibitors

The classic aspartic protease inhibitor pepstatin displays a low molecular weight and is a potent reversible and specific inhibitor for this type of proteases. This molecule was discovered by Umezawa *et* al., 1970 in the culture filtrates of various species of *Actinomyces*. Pepstatin is able to inhibit aspartic proteases of families A1 and A2 (Fujinaga *et al.*, 1995; Prashar, and Hosur, 2004) including porcine pepsin, renin, cathepsin D (Aoyagi *et al.*, 1972), bovine chymosin (Takahashi, K., and Chang, W.J., 1973), human pepsin, human gastricsin, protease B from *Aspergillus niger*, and several acid proteases of microbial origin (Aoyagi *et al.*, 1971). The stringent specificity of pepstatin toward acid proteases clearly has been demonstrated by the lack of inhibition of the neutral and alkaline proteases (Umezawa, H., 1972). Pepstatin is a tight-binding, reversible inhibitor, acting as an analogue of the tetrahedral intermediate in catalysis. For example, a K_i value of 4.5×10^{-11} M was obtained for this inhibitor against porcine pepsin (Schmidt *et al.*, 1982)

Known proteinaceous aspartic protease inhibitors are rare and unevenly distributed among classes of organisms in contrast to proteinaceous inhibitors of serine and cysteine proteinases (Laing *et al.*, 2002). Among them, the proteinaceous inhibitor known as IA_3 protein has been characterized as a potent endogenous inhibitor of 68 amino acid residues against the aspartic protease saccharopepsin from *Saccharomyces cerevisiae* (Biedermann *et al.*, 1980). While it shows no detectable secondary structure in solution, upon complexation with the enzyme, residues 2–32 adopt an almost perfectly helical conformation revealing that the protease body serves as a folding template (figure 7A) (Ng *et al.*, 2000).

The major aspartic protease inhibitor from the intestinal parasitic nematode *Ascaris suum*, denominated pepsin inhibitor-3 or aspin, was isolated and characterized by Martzen *et al.*, 1990. The inhibitor consists of 149 amino acid residues and is able to inhibit porcine pepsin (K_i=2.6 nM), porcine renin, and monkey cathepsin D and E (K_i = 3.7 nM) as described by kageyama, T., 1998. This molecule displays an unrelated mode of inhibition. Its N-terminal β-strand pairs with one strand of the active-site flap, forming an extensive eight-stranded β-sheet spanning both proteins (figure 7B) (Janin and Chothia, 1990).

On the other hand, SQAPI (squach aspartic proteinase inhibitor) from the squash phloem exudates (*Cucurbita maxima* Duchesne) was purified and characterized by Christeller *et al.*, 1998. This inhibitor displayed a molecular mass of 10.5 kDa and it was capable of inhibiting pepsin (K_i = 2 nM) and a secreted aspartic proteinase from the fungus *Glomerella cingulata* (K_i = 20 nM).

Equistatin, a polypeptide isolated from the sea anemone *Actinia equina*, demonstrated that not only interacts with papain (K_i = 0.57 nM) and other members of this family but also inhibited an aspartic protease cathepsin D (K_i = 0.3 nM) (family A1) (Lenarčič *et al.*, 1997; Lenarčič and Turk, 1999). In addition, cathepsin D inhibitors have been described in potato (*Solanum tuberosum*) (Mares *et al.*, 1989) and tomato (*Solanum lycopersicum*) (Werner *et al.*, 1993).

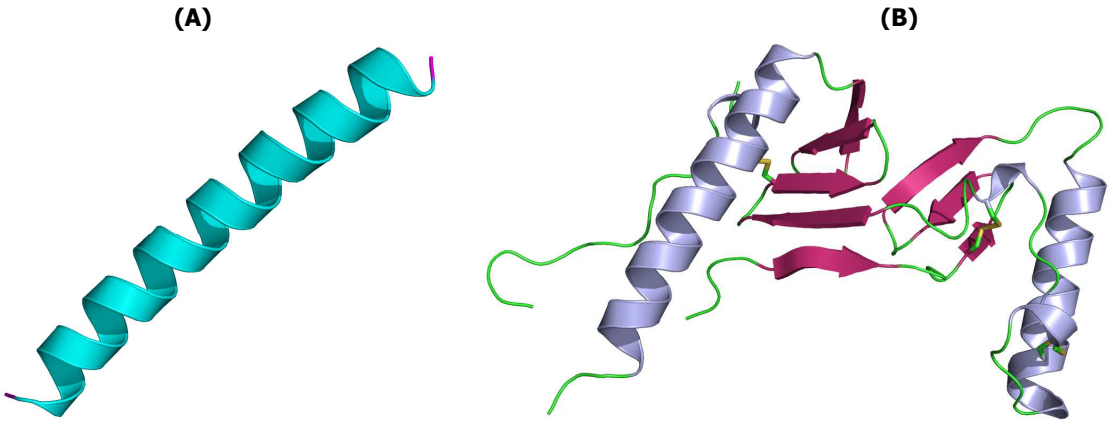


Figure 7. Three-dimensional structure of various proteinaceous aspartic protease inhibitors. (A) IA3 protein from *S. cerevisiae*. PDB code: 1DPJ (Li *et al.*, 2000). **(B) Pepsin inhibitor-3 from *A. suum*.** PDB code: 1F34 (Ng *et al.*, 2000).

I.2.3. Cysteine proteases and their inhibitors

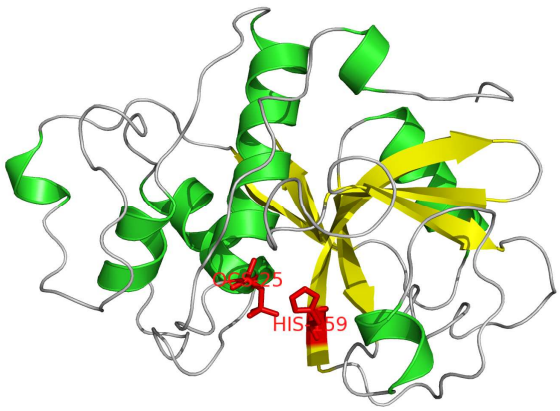
I.2.3.1. Cysteine proteases

The cysteine proteases are divided into at least 21 families (C1-C21) on the basis of the sequences or tertiary structures known for these enzymes, being Clan CA family C1 the most abundant and best characterized (Otto and Schirmeister, 1997; Rawlings and Barrett, 1999). Nearly half of the known families are represented in viruses (C3-C9, C16, C18, C21). Among these, the picornain family (C3) is the best characterized. Most of the enzymes known at present belong to the papain family (C1): proteases from protozoa, plant proteases, and lysosomal cathepsins. The families C1, C2 (calpain family), and C10 (streptopain family) can be described as "papain-like" and form clan CA (Otto and Schirmeister, 1997).

The papain superfamily consists of three members: papain group, calpain group, and bleomycin hydrolase (Berti and Storer, 1995). Cysteine proteases from the family C1 have been described in plants (papain, bromelain, etc), mammals (such as lysosomal cathepsins B, H, L, S, C and K, etc), bacteria (e.g. clostripain in *Clostridium histolyticum*, gingipain in *Porhyromonas gingivalis*), fungi (cathepsin B in *Aspergillus flavus*, proteinease yscF in *yeast*), parasitic organisms as protozoan (cruzipain in *Trypanosoma cruzi*, amoebopain in *Entamoeba histolytica*, and falcipain in *Plasmodium falciparum*) (Otto and Schirmeister, 1997; McGrath, 1999).

Independently of their origin, all C1 cysteine proteases share similarities in gene structure, sequence and folding with papain [3.4.22.2], the representative enzyme of the family (Turk *et al.,* 1997; Sajid and McKerrow, 2002). They are monomeric enzymes (except for tetrameric cathepsin C), consisting of a signal peptide, a pro-peptide and a catalytic domain of 220-260 amino acids. They are very similar proteins with bilobulated 3D-structure (R and L domains) as a result of the same arrange of secondary structure elements. Most of these enzymes exhibit both endo- and exopeptidase activity, and require acidic-to-neutral pH and a reducing environment for optimal proteolytic activity. Additionally, most of the enzymes of the family are reversible inhibited by divalent metallic ions and protein inhibitors belonging to the cystatin superfamily.

(A)



(B)

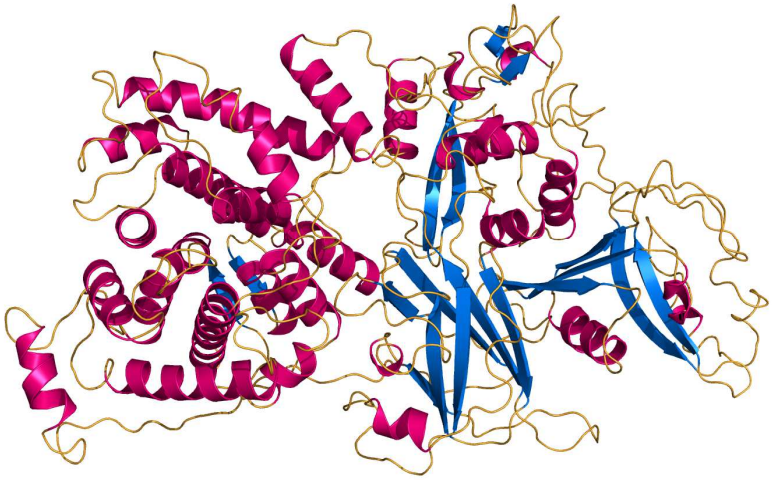


Figure 8. Three-dimensional structure of various cysteine proteases.
(A) Papain from *C. papaya*. PDB code: 9PAP (Kamphuis *et al.*, 1984). **(B) Human calpain-2.** PDB code: 1KFU (Strobl *et al.*, 2000)

The catalytic mechanism of cysteine proteases is very similar to that described for serine and threonine proteases. However, a remarkable difference with respect to the two latter is the presence of a nucleophilic cysteine thiol in the active-site of cysteine proteases (Polgar and Halasz, 1982). The proteolytic activity of all cysteine proteases arises from the presence of the catalytic Cys and His residues in the enzyme active centre. In the case of papain-like cysteine proteinases, the catalytic centre is complemented with Asn that ensures an orientation of the His imidazole ring optimal for successive stages of hydrolysis.

The crucial step of the catalytic process involves formation of a reactive thiolate/imidazolium ion pair (Cys-S–/His-Im+), which results from proton transfer between Cys-25 and His-159 (papain numbering) (figure 8A). In principle, the thiolate anion attacks the carbonyl carbon of the scissile peptide bond and the double bond between the carbon and the oxygen converts into a single one (step A in figure 9). The oxygen assumes a negative net charge allowing formation of the first tetrahedral transition state. The oxyanion is stabilized by hydrogen bonding to the NH groups of Gln-19 side chain and Cys-25 backbone, which is likely to result in the formation of an oxyanion hole (step B in figure 9) (Menard *et al.*, 1991; Menard *et al.*, 1995; Harrison *et al.*, 1997; Otto and Schirmeister, 1997). Subsequent rotation of the His residue enables proton transfer from the imidazolium cation to the nitrogen of the peptide bond being hydrolyzed, and cleavage occurs. The newly formed substrate amine is hydrogen bonded to His-159, whereas the substrate carboxylic part is linked to Cys-25 *via* a thioester bond, forming acyl enzyme (step C in figure 9). The next reaction step involves dissociation of the aminic part of the substrate and its replacement with a water molecule. The imidazole nitrogen contributes to polarization of the water molecule that in turn attacks the carbonyl carbon of acyl enzyme (step D in figure 9). This is followed by formation of the second tetrahedral intermediate (step E in figure 9). In the final step, thioester deacylation leads to reconstruction of the carboxyl group in the hydrolyzed substrate, which is concerted with the release of an active enzyme (step F in figure 9) (Menard *et al.*, 1991; Otto and Schirmeister, 1997).

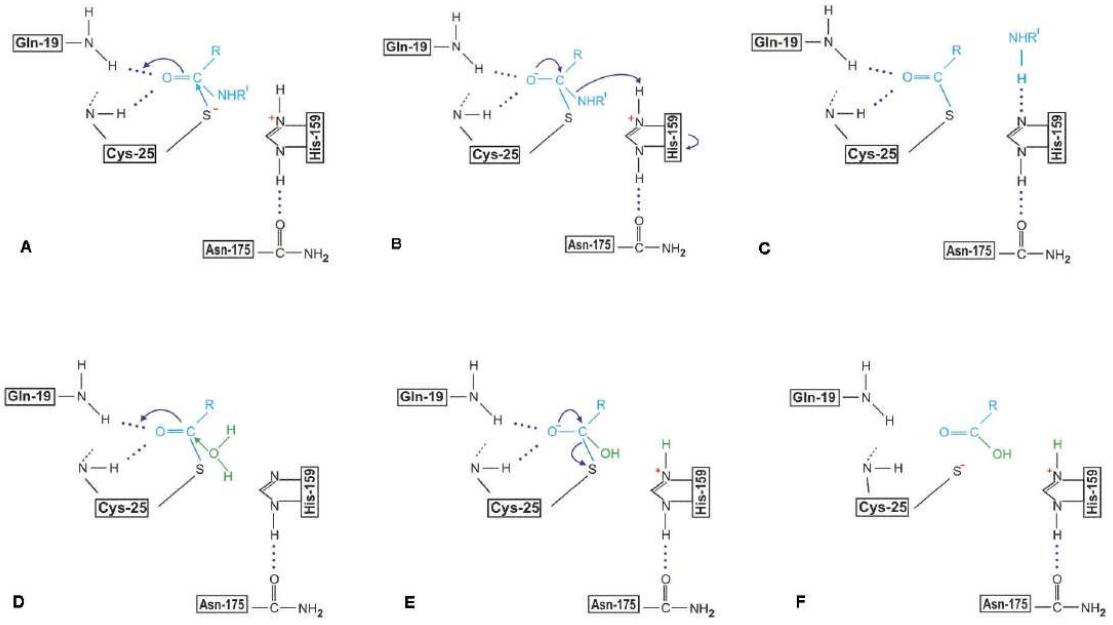


Figure 9. Catalytic mechanism of cysteine proteases as exemplified by papain (figure from Rzychon *et al.*, 2004)

Cysteine proteases are responsible for many biochemical processes occurring in living organisms and they are implicated in systemic diseases as myocardial infarction, osteoporosis, among others (Rosenthal, P.J., 1998; Abbenante and Fairlie, 2005). The main physiological role of cysteine proteases is the metabolic degradation of peptides and proteins. For example, mammalian cystein proteases have been implicated in the development and progression of many diseases that involve abnormal protein turnover (Kirschke *et al.*, 1995; Grubb, 2000), such as cathepsin B, which plays various digestive and processing roles inside cell to maintain normal cellular metabolism and associated with several pathophysiological conditions such as tumor metastasis, inflammation, bone resorption, and myocardial infarction (Leung *et al.*, 2000). Falcipains (Fp1, Fp2, Fp2' and Fp3) are papain like cysteine proteases from *Plasmodium falciparum* and they are very attractive targets for the development of antimalarial drugs (Rosenthal, P.J., 1998).

Therefore, papain-like cysteine proteases are enzymes with high biomedical potential. They are involved in many important aspects in the physiology of human cells, as well as in critical events of the life cycle of multiple human parasites (Sajid and McKerrow, 2002; Dubois *et al.*, 2006).

I.2.3.2. Cysteine proteases inhibitors

Many natural protein inhibitors of cysteine proteases, called cystatins, have been isolated and characterized. They act both intra- and extra-cellularly forming complexes with their target enzymes. Maintenance of appropriate equilibrium between free cysteine proteases and their complexes with inhibitors is critical for proper functioning of all living systems. In this role, cystatins are general regulators of harmful cysteine protease activities. The roles of cystatins in health and disease have been reviewed by Henskens *et al.*, 1996 and Grubb, A., 2000.

The human super family of cystatins is divided into four groups based on structural similarities: Family I, called stefins comprises intracellular cystatins A and B (figure 10A). Family II includes extra-cellular and/or trans-cellular cystatins (cystatins: C, D, E/M, F, S, SA, and SN). Kininogens, the intravascular cystatins, form the family III of cystatins, and family IV, corresponding to phyto-cystatins. Both stefins as cystatins are constituted by monomeric proteins with molecular masses between 11 to 13 kDa.

Stefins are synthesized in free ribosomes, containing a signal peptide, disulfide bridges as well as glycosylations, and they are confined into the intracellular matrix (Anastasi *et al.*, 1983). Cystatins belonging to the family II have approximately 120 amino acid residues, and a presence of a signal peptide for extracellular targeting (Dubin, G., 2005). These proteins present two conserved disulfide bridges in the C-terminal tail of the protein, with 10-20 amino acid residues between the respective cysteine residues forming the disulfide bond, and a characteristic Pro-Trp pair in their C-terminal segments (Abrahamson *et al.*, 2006). Furthermore, these proteins may be modified by phosphorylation (Laber *et al.*, 1989).

Kininogens are multidomain proteins, displaying quite high molecular mass (60-120 kDa). These proteins are constituted by three tandemly repeated cystatin domains, with a total of eight disulfude bridges and they are glycosylated proteins (Abrahamson *et al.*, 2006). Two kininogens are found in mammalian serum: HK (high molecular weight kininogen) and LK (low molecular weight kininogen) with the exception of the rat which encompasses a third kininogen, T-Kininogen (TK) (Abrahamson *et al.*, 2006; Lalmanach *et al.*, 2010). Human kininogens are intravascular proteins of blood plasma and play a role in cell and vascular biology. They are also the precursor peptides for vasoactive kinins (Dickinson, D.P., 2002).

The phyto-cystatin family is constituted by a series of cysteine protease inhibitors from plant sources within the cystatin superfamily (Kondo *et al.*, 1991). While in some respects these proteins resemble the cystatins, they are stefin-like in having no disulfide bonds or cysteine residues. The sequence revealed homology to the cystatins including conservation of the sequence phe-ala-Val-asn-glu-his-asn**,** but no disulfide bonds, as seen in the stefins (Brown and Dziegielewska, 1997).

Thyropins, also known as thyroglobulin type-1 (thyr-1) domain proteinase inhibitors, have demonstrated to inhibit cysteine proteases of the papain family (C1). Similarly to cystatins, the thyr-1 domain fold of some thyropins adapted to inhibition of proteases outside C1 family. A representative member of this family is the polyvalent inhibitor called equistatin.

Chagasin (apparent molecular mass, 12 kDa) is an endogenous cysteine inhibitor found in the parasitic protozoa *Trypanosoma cruz*, which is able to inhibit cruzipain and papain (Monteiro *et al.*, 2001). This inhibitor was the first described member of a novel inhibitor family distinct from cystatins and other known groups, called the chagasin family (figure 10B) (Dubin, G., 2005).

Inhibitors of apoptosis (IAPs) represents a family of proteins distinguished by encompassing one or more characteristic, approximately 70-residue zinc binding BIR (baculovirus IAP repeat) domains. They were primarily characterized as inhibitors of apoptosis, although currently some BIR-containing proteins are known which do not seem to confer such a function ((Dubin, G., 2005). The central mechanisms of IAP apoptotic suppression appear to be through direct caspase and pro-caspase inhibition (primarily caspase 3 and 7) and modulation of the transcription factor NF-kB (LaCasse *et al.*, 1998). The first IAPs were identified in the baculoviruses *Cydia pomonella* granulosis virus (CpGV) and *Orgyia pseudotsugata* nuclear polyhedrosis virus (OpMNPV) (Crook *et al.*, 1993; Birnbaum *et al.*, 1994). Since the discovery of the baculoviral IAPs, numerous cellular homologues have been identified in a range of species from *Drosophila* to vertebrates (Hunter *et al.*, 2007).

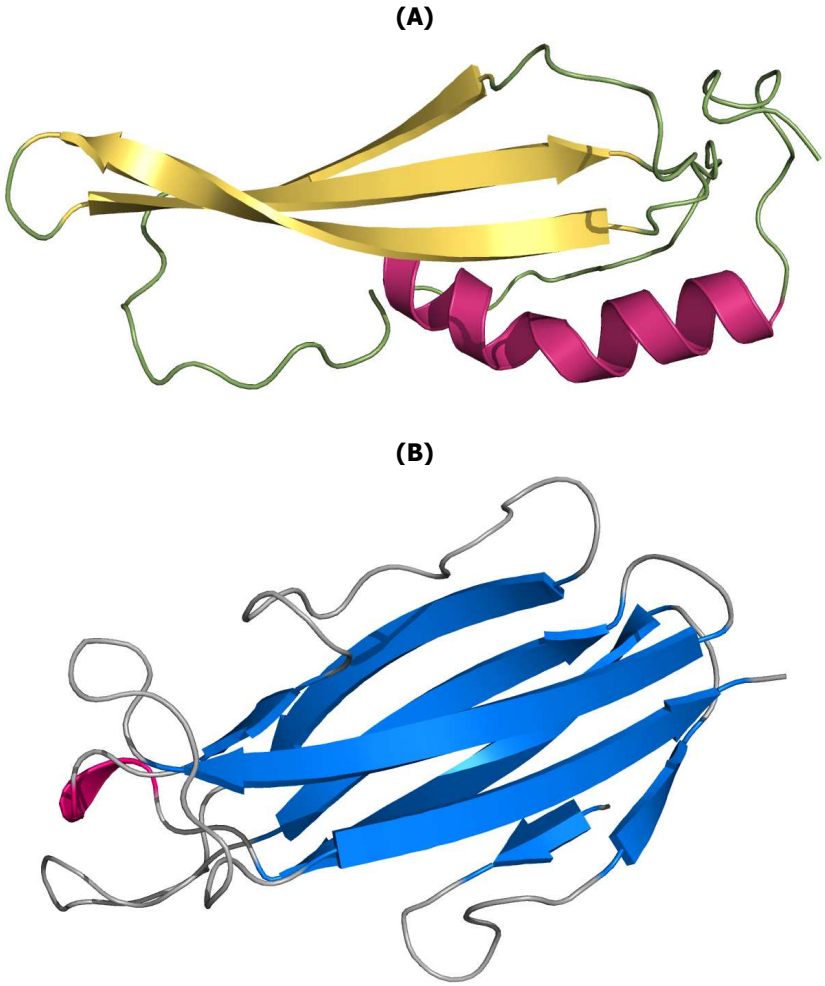


Figure 10. Three-dimensional structure of various proteinaceous cysteine protease inhibitors. (A) Human cystatin B. PDB code: 2OCT (Jenko *et al.*, 2007). **(B) Chagasin.** PDB code: 2NQD (Ljunggren *et al.*, 2007)

A wide variety of synthetic compounds have been described as cysteine protease inhibitors. Among them, aldehydes, semicarbazones, α-keto acids, nitriles, diazomethanes, epoxysuccinyl amino acid derivatives, azobenzenes, etc. (Ettari *et al.*, 2009). L-*trans*-Epoxysuccinyl-leucylamido-(4-guanidino)-butane, also known as E-64 is a potent cysteine inhibitor widely used as an active-site titrant of papain, cathepsins B, H, and L, among others (Barret *et al.*, 1982).

I.2.4. Serine proteases and their inhibitors

I.2.4.1. Serine proteases

Serine proteases are so named due to the presence of a serine residue inside their active-site, which is found to be essential for the catalytic activity (Neurath, H., 1989). This mechanistic class was firstly distinguished by the presence of the sequence Asp-His-Ser, also referred as the catalytic triad, where at least four clans have been identified. Representative enzymes for these clans are chymotrypsin (clan SA), subtilisin A (clan SB) (figure 11B), carboxypeptidase Y (clan SC) and Clp protease (clan SE) (Rawlings and Barrett, 2000; Rawlings *et al.*, 2002, Hedstrom, L., 2002; Tyndall *et al.*, 2005).

Serine proteases from the clan SA are proteins consisting of two-domains formed by β-strand structures and a cleft between domains, where the active-site is located. They have the catalytic triad His-Asp-Ser and all the members of this clan (approximately 300 proteins) are endopeptidases. Trypsin is a typical example of serine protease belonging to this clan (figure 11A).

Clan SB has about 91 proteins, which consist of both α-helix as β-sheet structures. They could be exo- or endopeptidases, and posses the catalytic triad Asp-His-Ser, where subtilisin is a representative enzyme (figure 11B). Proteins grouped into the clan SC (approximately 64 members) are formed by α-helices as β-sheet and present the catalytic triad Ser-Asp-His. In this clan are included oligopeptidases, as well as exopeptidases, which may be amino- and carboxypeptidases.

Furthermore, the clan SE (about 14 members), where the D-Ala-D-Ala carboxypeptidase A (*Geobacillus stearothermophilus*) is the representative enzyme, is composed by proteins displaying both α-helix as β-strand structures. They are grouped into two lobes with the Ser and Lys residues located in a α-helix that crosses the cleft between the lobes. They have the catalytic residues in the order Ser, Lys (within the motif SXXK) in sequence and there is a closely-spaced catalytic dyad Ser66, Lys69. A third residue, Ser/Tyr130, also plays a part in catalysis (Rhazi *et al.*, 2003). The MEROPS database provides detailed information about these clans (http://merops.sanger.ac.uk).

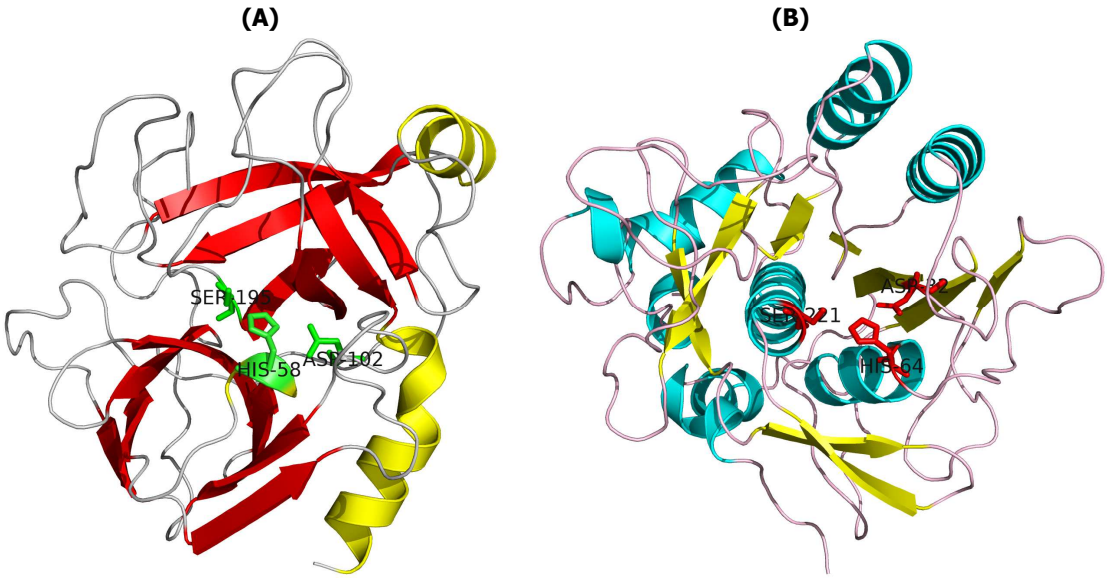


Figure 11. Three-dimensional structure of various serine proteases.
(A) Bovine trypsin. PDB code: 3ATL (Yamane *et al.*, 2011). **(B) Subtilisin from *Bacillus licheniformis*.** PDB code: 1YU6 (Maynes *et al.*, 2005)

The catalytic mechanism of serine proteases consists of a first acylation step, where an acyl-enzyme intermediate and tetrahedral intermediate are formed, followed by a decomposition of the acyl intermediate, in which the enzyme is regenerated (figure 12).

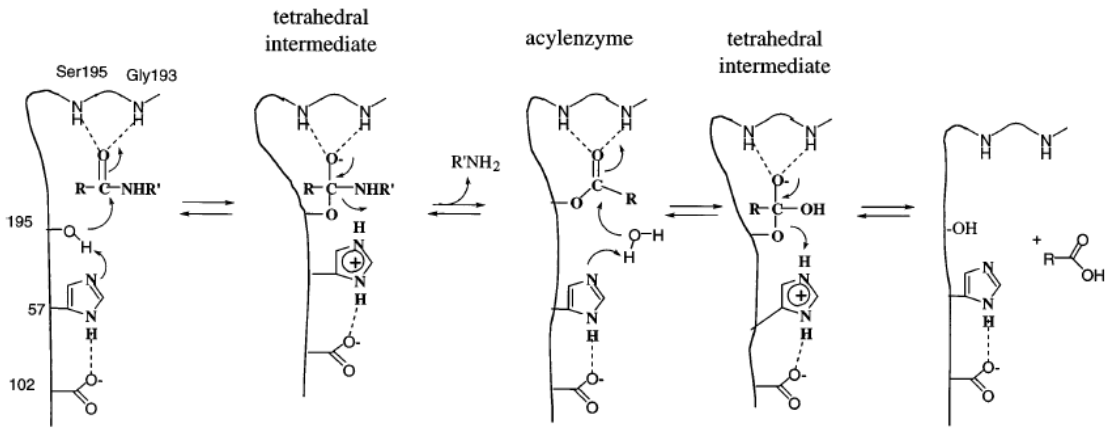


Figure 12. The generally accepted mechanism for serine proteases (figure from Hedstrom, L., 2002b)

Serine proteases are involved in critical physiologic processes in mammals such as digestion, homeostasis, signal transduction, immune response, growing, matrix repair, cell differentiation, as well as numerous proteolytic cascades such as blood clotting, complement fixation and fibrinolysis (Neurath, H., 1989; Barret *et al.*, 1998; Coughlin, S.R., 2000). In prokaryotes and viruses, serine proteases have been ascribed to play important roles in replication and spread cycles (Hedstrom, L., 2002).

The diversity of physiological functions in which the serine proteases are implicated, requires a broad specificity into the hydrolysis sites, as well as distinctive features in these enzymes for the substrate-binding sites adjacent to the catalytic triad (Perona and Craik, 1995; Perona *et al.*, 1995; Czapinska and Otlewski, 1999; Hedstrom, L., 2002).

I.2.4.2. Serine protease inhibitors

Serine protease inhibitors are widely found in most living organisms and have been isolated from different cells and tissues. Their relative abundance exceeds inhibitors that inhibit proteases of other mechanistic classes. It is unclear if this fact is due to their distribution in nature, the preference or convenience of its isolation by researchers, or because of the extensive abundance of serine proteases, such as trypsin, chymotrypsin, subtilisin, among others.

About 20 structurally diverse inhibitors families have been identified, comprising α-helical, β-sheet, and α/β proteins, and different folds of small disulfide-rich proteins (Krowrash *et al.*, 2003). Serine protease inhibitors are classified in three different types, according to their structural interaction established with the target protease: canonical (standard mechanism), non-canonical inhibitors and serpins (Bode and Huber, 1992, 2000; Laskowski and Qasim, 2000; Krowrash *et al.*, 2003).

Canonical inhibitors represent the largest group of serine protease inhibitors. They are proteins of 3-21 kDa per domain, where BPTI (bovine pancreatic trypsin inhibitor) (figure 13A), OMTKY3 (turkey ovomucoid third domain), and eglin are listed as representative proteins in this group (Krowrash *et al.*, 2003). In addition, they are tight-binding inhibitors, with a non-covalent interaction resembling enzyme-substrate Michaelis complex, direct blockage of the active-site, non-conformational, antiparallel β-sheet between enzyme and inhibitor (Bode and Huber, 2000), similar mode of interaction through canonical protease-binding loop despite completely different inhibitor structures (Ardelt and Laskowski, 1991), important role of P1 residue, additive effects on association energy (Lu *et al.*, 1993).

Non-canonical inhibitors display an extremely strong and specific interaction through their N-terminal segment which binds to the protease active site forming a short parallel β-sheet. These inhibitors also form extensive secondary interactions with the target protease outside the active site, which provide additional buried area and contribute significantly to strength, speed, and specificity of recognition. The non-canonical inhibitors are much less abundant than canonical inhibitors or serpins as they only occur in blood-sucking organisms and inhibit proteases involved in clot formation – thrombin or factor Xa. These proteins have molecular masses between 6-8 kDa per domain, and hirudin from the medical leech *Hirudo medicinalis*, TAP (tick anticoagulant peptide), and Ornithodorin from soft ticks, are some examples of proteins belonging to these type of inhibitors (Krowrash *et al.*, 2003).

Unlike canonical inhibitors, serpins are much larger proteins, typically 350–500 amino acids in size, distributed from viruses to mammals (figure 13C) (Gettings, G.P., 2002; Silverman *et al.*, 2001). They are abundant in human plasma and mutations in serpins lead to numerous serious genetic diseases in humans (Stein and Carrell, 1995). In contrast to canonical inhibitors, serpins utilize the kinetic features of a hydrolytic reaction to form a very stable acyl-enzyme intermediate. The enzyme-serpin complex is a covalent acyl-enzyme adduct and upon acylation the protease is translocated by over 70 Å from its initial recognition site (Huntington *et al.*, 2000). Serpins are the only family of serine protease inhibitors capable of interacting with proteases of other mechanistic classes, such as cysteine (Komiyama *et* al., 1994) and aspartic proteases (Mathialagan and Hansen, 1996).

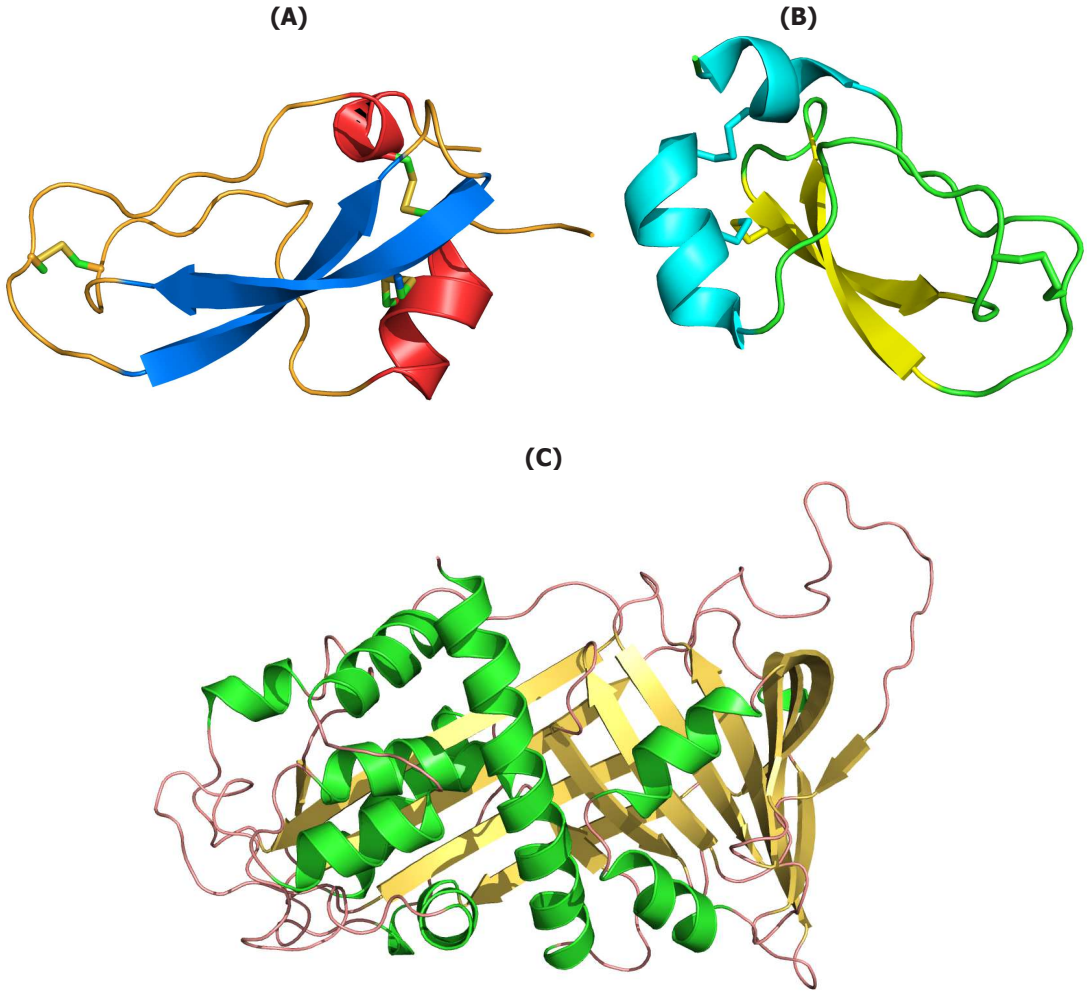


Figure 13. Three-dimensional structure of various proteinaceous serine protease inhibitors. **(A) BPTI.** PDB code: 5PTI (Wlodawer *et al.*, 1984). **(B) ShPI-1.** PDB code: 3OFW (García-Fernández *et al.*, 2011). **(C) Serpin from *Manduca sexta*.** PDB code: 1K9O (Ye *et al.*, 2001).

I.3. Marine invertebrates as a source of protease inhibitors

The importance of marine organisms as a source of new substances is growing. Marine organisms offer an enormous resource for novel compounds and it has been classified as the largest remaining reservoir of natural molecules, which could be used for basic research, industry and biomedicine. There are several reasons to support this fact: great biodiversity (chemical diversity) largely unexplored, organisms submitted to unique host environmental conditions of predation-defense relationships (Mebs y Gebauer, 1980; Shiomi *et al*., 1985; Sencic y Macek, 1990; Castañeda *et al*., 1995). This environment is very different in many aspects from the terrestrial, a situation that demands the production of quite specific and potent active molecules. Marine organisms have great potential as source of protease inhibitors and other bioactive polypeptides (Antuch *et al*., 1993; Castañeda *et al*., 1995; Alvarez *et al*., 1996).

Metalloprotease inhibitors: A metalloendopeptidase inhibitor called Jaspisin has been isolated and characterized from the marine sponge *Jaspis sp.* This molecule displays an IC_{50} value of 8.6 μg/ml against hatching enzyme, a kind of Zn^{2+}-dependent metalloendopeptidases (Ikegami *et al*., 1994).

Another interesting metalloprotease inhibitor is HcPI isolated from the marine annelide *Hermodice carunculata*, which is a non-peptide inhibitor (MW=580 Da). This molecule is able to specifically inhibit the metallop-ectoenzyme pyroglutamyl aminopeptidase II (PPII) involved in the inactivation of the thyrotropin-releasing hormone (TRH). HcPI represents the most potent inhibitor described against PPII (K_i = 51 nM) and may be used in the treatment of brain and spinal injury as well as in CNS disorders (Pascual *et al*., 2004; Chávez-Gutiérrez *et al*., 2005).

As mentioned above, a polyvalent inhibitor has been isolated and characterized from the marine annelide *Sabellastarte magnifica*. SmCI is a N-glycosylated protein of 165 amino acid residues, which is able to strongly inhibit carboxypeptidase A (K_i = 24 nM), trypsin (K_i = 4.5 nM), pancreatic elastase (K_i = 2.0 nM) and more weakly chymotrypsin (0.93 μM). However, SmCI does not display inhibitory activity against carboxypeptidase B, papain, or pepsin. Structure modeling studies have revealed that SmCI consists of three domains, with high similarity to serine proteases of the BPTI/Kunitz family (Alonso del Rivero *et al*., 2012).

Ancorinosides B–D have been isolated from the marine sponge *Penares sollasi* Thiele as inhibitors of MT1-MMP. Their structures were elucidated to be tetramic acid glycosides related to ancorinoside A by spectroscopic and chemical methods. Ancorinosides B–D inhibit MT1-MMP with IC_{50} values of 180–500 μg/ml (Fujita *et al*., 2001). In addition, callysponginol sulfate A has been isolated from the marine sponge *Callyspongia truncata* as a MT1-MMP inhibitor. Its structure was elucidated by a combination of spectroscopic and chemical methods and found to be a new sulfated C24 acetylenic fatty acid. This compound inhibits MT1-MMP with an IC_{50} value of 15.0 μg/ml (Fujita *et al*., 2003).

Aspartic protease inhibitors: As mentioned above, aspartic protease inhibitors are less frequent in nature, compared to inhibitors against proteases belonging to other mechanistic classes. The HIV protease is probably the most target enzyme used in the research of inhibitors from marine invertebrates. The Didemnaketals A and B are alkaloids isolated from the ascidian *Didemmun sp*, which are able to inhibit HIV-1 protease with IC_{50} values of 2 and 10 μM, respectively (Potts *et al*., 1991). The unusual structures of these compounds along with their inhibitory activity have become in leader molecules (Nakao and Fusetani, 2007). On the other hand, brominated polyacetylenic acids have been identified in the marine sponge *Xetospongia muta*, which exhibit inhibitory activity against HIV-1 protease with IC_{50} values about 12 μM (Patil *et al*., 1992). In addition, an alkaloid with inhibitory activity against plasmepsin II has been isolated and structurally characterized from the marine sponge *Smenospongia aurea* (Hu *et al*., 2002). Recently, a polypeptide isolated from the gorgonian *Plexaura homomalla* has been described as a tight-binding inhibitor of plasmepsins II and IV, with inhibition constants of 4.3 and 17.9 nM, respectively (Salas *et al*., 2012, in press).

Cysteine protease inhibitors: Papain has been mainly used as model enzyme in the research of cysteine protease inhibitors. Equistatin is a 22 kDa protein isolated from the sea anemone *Actinia equina*. This molecule displays a strong inhibitory activity against cysteine and aspartic proteases such as cathepsin L (K_i = 0.051 nM), papain (K_i = 0.57 nM), and cathepsin B (K_i = 0.3 nM) in the case of cysteine proteases, as well as the aspartic protease cathepsin D (K_i = 0.3 nM) (Lenarcic *et al.*, 1997; Lenarcic and Turk, 1999). Equistatin consists of three homologue domains, where the domain I is responsible of the papain inhibitory activity, whereas the domain II is involved in the inhibition of cathepsin D (Strukelj *et al.*, 2000)

On the other hand, tokaramide A and miraziridine A are two proteinaceous inhibitors of cathepsin D, which is involved in tumor metastasis, isolated from the marine sponge *Theonella aff. mirabilis* (Fusetani *et al.*, 1999; Nakao *et al.*, 2000). Miraziridine A also displays inhibitory activity against proteases of other mechanistic classes such as trypsin (serine protease) and pepsin (aspartic protease) (Schaschke, N., 2004).

Secobatzelline A, a batzelline natural analogue, and Discorhabdin P, a discorhabdin analogue, have been isolated from a deep-water caribbean sponge of the genus *Batzella*. These molecules exhibit inhibitory activity against the caspase CPP32, which are a group of at least 10 cysteine proteases (also known as interleukin-2 converting enzymes or ICE_2), which play a major role in the programmed cell-death mechanism known as apoptosis (Gunasekera *et al.*, 1999a; Gunasekera *et al.*, 1999b).

Serine protease inhibitors: Most of the proteases and inhibitors that have been identified in these sources belong to the serine class (Fritz *et al.*, 1972; Delfín *et al.*, 1996; Zykova *et al.*, 1998; Chávez *et al.*, 1998; Nagle, et al., 2001; González *et al.*, 2007).

Canonical inhibitors have been isolated from the ascidians *Halocynthia roretzi* and *Ciona intestinalis*. The first trypsin inhibitor isolated from the hemolymph of *H. roretzi*, called ATI, is a strong inhibitor of trypsin, plasmin, and an acrosin purified from sperm of this sea squirt (Yokosawa *et al.*, 1985). ATI consists of a polypeptide chain of 55 amino acid residues, four disulfide bridges, and its reactive site at Lys16-Met17 (Kumazaki *et al.*, 1990).

The serine protease inhibitor identified in *C. intestinalis* has a molecular mass of 6.6 kDa, and a non-classical Kazal-type domain with an Arg residue at is P1 site. It is a tight-binding trypsin inhibitor (K_i=0.05 nM) although no inhibitory activity is detected against chymotrypsin and elastase (Odum *et al.*, 1999). In addition, a 21 kDa protein inhibitor has been purified and characterized from the tropical gorgonian *Melithea cf. stormi*. The protein displays inhibitory activity against pancreatic elastase with a K_i value of 1.5 nM. The sequence of its first 39-residues revealed the presence of a non-classical Kazal-type domain (La Barre *et al.*, 1996).

Marine molluscs constitute another important source explored in the search of protease inhibitors. A proteinaceous inhibitor (MW=2.9 kDa) has been isolated from the sea hare *Aplysia dactylomela*. This is a plasma kallikrein inhibitor (K_i=0.22 nM), which is also able to inhibit human pancreatic trypsin and plasmin in the nanomolar range (González *et al.*, 2004).

In addition, three inhibitors of human neutrophil elastase (CmPI-I, II and III) have been identified from the marine snail *Cenchritis muricatus* (González *et al.*, 2007a). Analysis of the amino acid sequence and topology of the disulfide bridges showed that the major one (CmPI-II) displays homology with the Kazal domain that characterizes the family of inhibitors of the same name. However, the position of the disulfide bridge between cysteines I and V, and the number of amino acids between them suggests that CmPI-II belongs to a new group of inhibitors "non-classical" Kazal family (Gonzalez *et al.*, 2007b). This molecule has the exceptional capability of inhibiting human neutrophyl elastase having an Arg residue at the P1 site.

ShPI-1 is a proteinaceous inhibitor belonging to the Kunitz/BPTI family isolated from the sea anemone *Stichodactyla helianthus* (Delfín *et al.*, 1996). From the functional point of view, ShPI-I represents an exception among the protease inhibitors described, since this molecule has not only demonstrated the ability to inhibit trypsin and other serine proteases such as bovine chymotrypsin, plasmin and kallikrein, but also cysteine proteases such as bromelain and papain

and aspartic proteases as pepsin. This inhibitor has a molecular mass of 6.1 kDa, it consists of a polypeptide chain of 55 amino acids and three disulfide bridges (Antuch *et al.*, 1993). In addition, a second protease inhibitor called ShPI-II has been isolated and characterized from this anemone (Huerta *et al.*, 1998). The 3D structure of this inhibitor was solved by RMN (Antuch *et al*, 1993). Recently, their crystallographic structures (alone and in complexe with bovine trypsin) has been elucidated (García-Fernández *et al.*, 2012, in press). The latter represents the first X-ray structure of a marine invertebrate with a serine protease.

On the other hand, two inhibitors named AXPI-I and AXPI-II have been purified from the sea anemone *Anthopleura aff. Xanthogrammica*. Both molecules belong to the BPTI/Kunitz family and they are strong trypsin inhibitors (Minagawa *et al.*, 1997). Furthermore, an elastase inhibitor has been identified in *Anemonia sulcata*, which unlike trypsin inhibitors it revealed to have a Kazal-type serine protease inhibitor domain (Kolkenbrock and Tschesche, 1987).

Non-peptide inhibitors of serine proteases have also been found in marine invertebrates. The Cyclotheonamides A-D is a family of cyclic peptides isolated from the sponge *Theonella swinhoei* that display inhibitory activity against serine proteases (Fusetani et al., 1990, Nakao et al., 1995). On the other hand, Dynosins A-D have been purified from the Australian sponge *Lamellodysidea chlorine*, which inhibit the factor VIIa of the coagulation cascade (Carroll *et al.*, 2002; Carroll *et al.*, 2004).

I.4. Analytical approaches used for the identification and characterization of protease inhibitors

I.4.1. Kinetics of reversible tight-binding inhibition

The most important kinetic feature of these inhibitors, derived from their biological function is the ability to strongly bind to the target enzyme, which ensures high efficiency as inhibitors, with K_i values lower than $1x10^{-7}$ – $1x10^{-8}$ M. These molecules are generally transition state analogues and act as competitive inhibitors. Consequently, these characteristics led to effective concentrations of tight-binding inhibitors in the same range of enzyme concentrations to be inhibited; therefore the basic principles of Michaelis-Menten kinetics are not fulfilled. Thus, the inhibitory activity assays, which are usually expressed in percentages of inhibition or inhibition fraction, may not adequately reflect the total amount of inhibitor present in a sample, or its effectiveness against the enzyme tested. Moreover, the K_i determination can not be performed according to the principles of classical competitive inhibition, since it requires a special treatment. This kinetic parameter, ultimately, is the unique way to measure the real efficiency of a tight-binding inhibitor against the target enzyme tested.

As mentioned above, the effective concentrations of tight-binding inhibitors are in the same range of enzyme concentrations used in the kinetic assays. For this reason, as in the case of the enzyme, a mass conservation equation for the inhibitor must be included: $[I_0] = [I] + [EI]$ (where $[I_0]$: initial concentration of total inhibitor, [I]: concentration of free inhibitor in the equilibrium, [EI]: concentration of enzyme-inhibitor complex in the equilibrium). These conditions lead to a kinetic equation different from the classic equation for Michaelis-Menten competitive inhibition (Bieth J.G., 1995). The experimental procedure for determining the inhibition constant (K_i) is relatively simple. Initially, the active inhibitor concentration must be known, especially if it is a molecule of proteinaceous nature. This active concentration is determined by titrating the inhibitor with a known active enzyme concentration.

In the same way, the determination of the active enzyme concentration can be done by titrating the molecule with substrates, irreversible inhibitors or tight-binding inhibitors (Knight G.C.; 1995; Chávez and Gutiérrez, 2012). After the active enzyme concentration is measured, appropriate conditions of enzyme and substrate concentration are selected, in which an initial study of the effect of preincubation time of the enzyme-inhibitor on the inhibitory activity is performed, with the aim of establishing the time required to achieve the equilibrium. This study also enables to determine if the inhibitor acts as a slow or fast binding molecule.

Subsequently, the active inhibitor concentration is assessed by titrating the molecule with the standard enzyme and the fraction of not bound enzyme is calculated (**a**), and expressed as percentage or residual enzymatic activity (**a**=v_i/v_0) at different inhibitor concentrations. A linear behavior of residual enzymatic activity (v_i/v_0) according to the relationship $[I_0]/[E_0]$ (where $[E_0]$: initial concentration of active enzyme) means that the interaction is so strong that virtually all the inhibitor is combined in the EI complex, in other words, the inhibitor is able to titrate the enzyme.

This behavior occurs experimentally in the presence of reversible tight-binding inhibitors or pseudo-irreversible inhibitors, when the following conditions are used: high $[E_0]/K_i$ ratio ≥ 100 and $[S_0] \leq K_M$ where $[S_0]$: initial substrate concentration and K_M: Michaelis-Menten constant to prevent the induction of EI dissociation caused by high substrate concentrations. Moreover, the time of enzyme-inhibitor preincubation must be sufficient to ensure maximal binding.

In the procedure to determine the K_i value, experimental conditions must be adjusted in order to obtain a concave behavior of **a**=v_i/v_0 as a function of varying inhibitor concentrations at a fixed concentration of enzyme and substrate. This behavior occurs in the presence of a reversible tight-binding inhibitor and it is achieved by using the following experimental condictions: $[E_0]/K_i \leq 10$ y $[S_0] \approx K_M$. In addition, the time required to reach the equilibrium previously evaluated must be corroborated in these new experimental (Bieth J.G., 1995; Copeland R.A., 2000; Chávez and Gutiérrez, 2012).

This behavior is described by the general equation for tight-binding inhibition proposed by Morrison J.F., 1982, which clearly shows that both K_i as $[E_0]$, are important factors in the enzyme-inhibitor interaction.

$$a = 1 - \frac{\left\{\left([E_0] + [I_0] + K_{iapp}\right) - \left([E_0] + [I_0] + K_{iapp}\right)^2 - 4[E_0][I_0]\right\}^{1/2}}{2[E_0]}$$

The value **a** as a function of $[I_0]/[E_0]$ (concave curve) demonstrates the reversibility of inhibition and the possibility of determining the K_i value by this equation. However, the K_i value assessed is apparent and the real value can only be calculated, taking into account the effect of substrate concentration on the inhibitory activity, according to the equation described by Bieth J.G., 1995.

$$K_{iapp} = K_i\left(1 + \frac{[S_0]}{K_M}\right)$$

This equation is used in order to take into account the possible induction of enzyme-inhibitor complex dissociation caused by the substrate, bearing in mind the reversibility of the system indicated by obtaining a concave behavior.

I.4.2. Enzyme immobilization

The rapid development of biotechnology in the last two decades increasingly demands the use of enzymes for industrial, analytical and biomedical applications. The practical use of enzymes in many areas has been limited by factors related to its high cost, inability to re-use and low stability. Immobilization technology, therefore, represents not only an alternative to remove some of the limitations associated with the use of enzymes and other proteins in practice, but it is also a key tool for the successful application of biotechnology in new products.

Enzyme immobilization can be defined as the retention of the enzyme molecules in a phase (microphase enzyme phase), which maintains a constant exchange with the rest of the system or macrophase, where effectors are involved. Another definition of enzyme immobilization that includes requirements and the most important advantages of this process is the following: the location of the enzyme in one phase, with retention of its functional activity and its ability to be used continuously and repetitively in a process.

The benefits derived from the enzyme immobilization are:

- The easy recovery of the enzyme from the reaction mixture. It is possible to stop the reaction at any time, absence of contaminant products derived from the enzyme as well as the reuse of the enzyme preparation.
- The use of immobilized enzyme in continuous reactors and, therefore, the possibility of regulating the enzymatic reaction rate and product yield, by varying the flow rate in the reactor.
- The feasibility, in many cases in a targeted way, to change the enzyme properties, such as its catalytic activity and stability.

The possibility of using the enzymes repeatedly and continuously in processes, as already mentioned above, contributes to the reduction of process costs. In addition, enzyme immobilization can improve the properties of therapeutic enzymes by reducing, in some cases, their antigenicity, increased half-life in blood and the possibility of ensuring a better distribution of the molecule in the body.

Immobilization supports: The fundamental properties of the immobilization supports are listed as follows: large surface area, permeability and porosity, hydrophilic character, insolubility, rigidity, particle size and shape, physical, chemical and biological stability, presence of reactive groups or susceptible to activation, regenerability, and economy.

Immobilization supports can be classified according to their chemical composition in organic and inorganic supports. However, it is also important to take into account their morphological features allowing classify them into non-porous and porous media. This property determines the surface area and pore size of the support, which influences the capability of the immobilization support as well as the efficiency of the enzyme operating. The non-porous materials are commonly used in nanotechnology; particularly as inorganic matrices (Chávez *et al.*, 2011) although they have the disadvantage of the presence of small surface areas, which leads to low degrees of immobilization. The increased surface area through the use of very fine particles results in high pressure drops and low flow rates in reactors with immobilized enzymes, as well as difficulties in the separation of the immobilized enzyme derivative that limits its potential use in a continuous system.

I.4.2.1. Enzyme immobilization methods

The enzyme immobilization methods most frequently used, shown in figures 14 and 15, are based on combining the nature of the interaction responsible for the immobilization between enzyme and support (Kennedy and Cabral, 1983). This classification, therefore, ranges from more traditional forms of immobilized enzymes by the use of insoluble supports to retention of enzymes in systems where the support is not directly involved. Figure 14 outlines the different immobilization methods as follows:

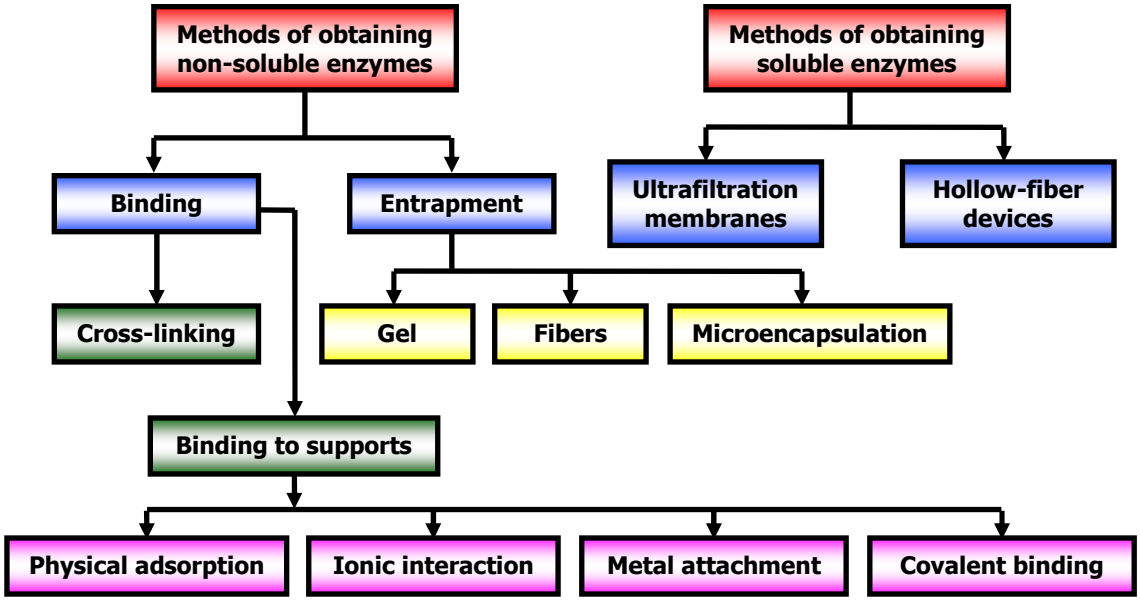


Figure 14. Classification of enzyme immobilization methods

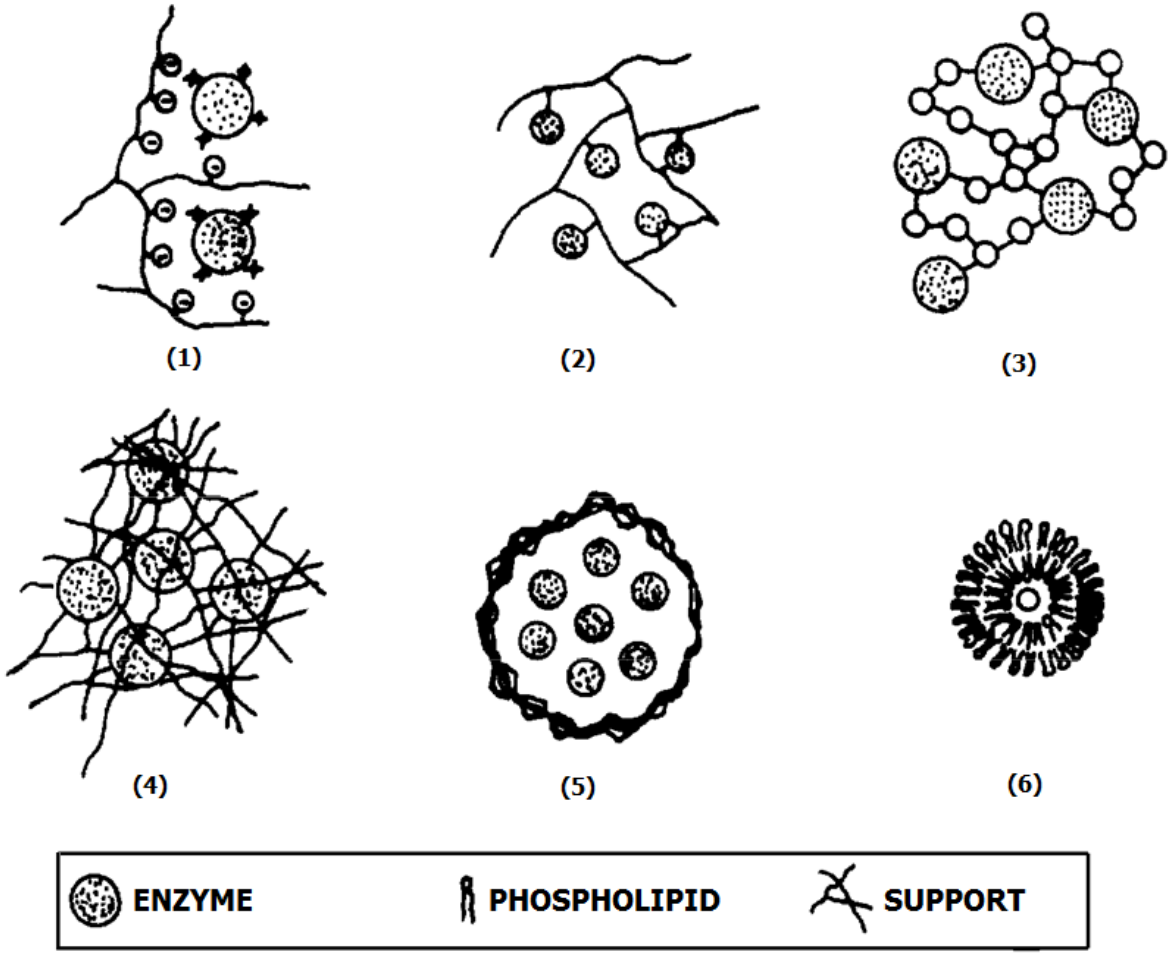


Figure 15. Example of enzyme immobilization methods (figure from Chávez and Gutiérrez, 2012)
(1) Binding by electrostatic interactions. **(2)** Covalent immobilization. **(3)** Intramolecular cross-linking. **(4)** Gel entrapment. **(5)** Microencapsulation in membranes. **(6)** Microencapsulation in liposomes

I.4.2.1.1. Physical adsorption methods

Adsorption is based on weak attractive interactions between the enzyme and the external or internal surface in solid supports, such as hydrogen bonds, hydrophobic associations and Van der Waals interactions. The degree of immobilization depends on several factors including the concentration of enzyme and adsorbent, solvent polarity, pH, ionic strength, temperature and contact time. It requires proper control of all these factors in order to achieve not only optimum enzyme adsorption, but also a good retention of functional activity.

Adsorption methods have several advantages. They are simple and economical procedures that produce few or no conformational change in the enzyme during the immobilization, since no reactive species are involved. However, the main disadvantage of these methods is the low stability obtained in the immobilized enzyme derivatives. The causes are related to the low strength of the interactions that bind the enzyme onto support, which results in enzyme desorption during the process. This can easily occurs when some conditions such as pH, ionic strength, solvent, among others, are changed. This limits the use of these methods in practice. The most frequently used supports are alumina, activated carbon, collagen, glass, diatomaceous earth and hydroxylapatite, among others.

I.4.2.1.2. Ionic interaction methods

The fundamental difference between the physical adsorption methods and ionic interaction methods for enzyme immobilization lies in the strength of the interaction between enzyme and support. Ionic interaction methods are based on electrostatic interactions between the enzyme and solid supports, in which have been introduced ion exchange groups. These interactions are stronger than those occurring during immobilization by physical adsorption.

These methods have advantages similar to those described for the physical adsorption methods, so often lead to the obtaining of immobilized derivatives with a high retention of catalytic activity. Its stability is higher, although variations of pH and ionic strength can cause loss of the enzyme by desorption. The common supports used in ionic interaction methods are polysaccharides such as cellulose and dextran, as well as synthetic polymers such as polystyrene, to which groups that can act as anionic and cationic exchangers are covalently bound.

I.4.2.1.3. Metal attachment methods

These methods are based on the chelating properties of transition metals which can be exploited for the enzyme immobilization. There have been immobilized enzymes in this way, using organic supports activated with transition metal salts. The salts used in this activation process have been chlorides and sulfates of titanium, iron, zirconium, vanadium and tin. This method has been used not only to activate organic supports such as cellulose, chitin, alginic acid, etc., but also inorganic and silica-based supports, celite, glass, wool, among others. Nucleophilic groups (hydroxyl, amino, thiol, etc., are effective ligands for transition metal ions that can complex both the enzyme as support. Enzymes have groups that can act as ligands, for example, free carboxyl groups at the C-terminal tail of the polypeptide chain and the side chains of glutamic and aspartic amino acids, the hydroxyl groups of the phenolic side chains of tyrosines, hydroxyl groups of serine and threonine residues, thiol group of cysteines, and free amino groups of the N-terminus of the chain and lysine side-chain (ε-amine).

The results obtained in the enzyme immobilization by this method depend mainly on the accessibility of groups in the enzyme that can act as ligands, steric factors that allow the interaction of such groups with transition metal atoms, the non-participation of residues of the active-site in the enzyme and the proximity of the molecules linked into the support. These methods allow obtaining immobilized enzyme derivatives with a high retention of their catalytic activities, but with varying operational stabilities.

I.4.2.1.4. Covalent binding methods

These methods are based on the formation of a covalent bond between enzyme and support. The main advantage of the covalent biding methods is the strength of enzyme-support interaction, which allows obtaining very stable immobilized enzyme derivatives that rarely cause loss of enzyme under normal conditions of use of these preparations. However, the selection of the experimental conditions for immobilization is more difficult than other methods, and conditions are usually severe and they can lead to immobilized derivatives with a low retention of functional activity. An additional problem is the guarantee that the covalent binding of the enzyme to the support groups do not involve the active-site of the enzyme, which must not be modified during the immobilization process.

The covalent immobilization of enzymes to a support requires the analysis of three key factors involved in the process: the functional groups of proteins suitable for covalent binding under mild conditions; the chemical reaction that allows the covalent binding of proteins to the support; and an activated support suitably modified for the protein immobilization.

Functional groups of proteins: Most of the chemical reactions described for covalent immobilization of enzymes are classified as carbonyl-like reactions with the nucleophilic groups of proteins $-HN_2$, -SH and -OH. The anion sulfur has a greater nucleophilic reactivity than nitrogen and oxygen compounds of similar basicity. However, the thioesters formed are much less stable than the esters, and these are, in turn, less stable than the substituted amines formed during the reaction with amine groups. Given these reasons and the frequency of occurrence of different amino acids in the protein composition, the more suitable residues for protein immobilization are: L-lysine, followed by L-cysteine, L-tyrosine, L-histidine, L-aspartic acid, L-glutamic acid, L-arginine, L-tryptophan, L-serine, L-threonine and L-methionine.

Most of the covalent binding reactions of proteins are based on reactions with amine, thiol, and carboxylic groups as well as with aromatic rings of L-tyrosine and L-histidine. These reactions are: diazotization, amide bond formation (peptide bond), alkylation and arylation, Schiff base formation, Ugi reaction, amidination reactions, reactions of thiol-disulfide exchange, mercury-enzyme interactions, and radiation-induced reactions.

Covalent immobilization represents one of the most commonly used enzyme immobilization method used and therefore the most important for different applications such as enzyme bioconversion, biosensors and nanotechnology. Besides, this method has been traditionally used for the purification of different molecules based on the molecular recognition principle. Another example is Intensity Fading MALDI-TOF MS for the identification of biomolecules such as protease inhibitors or proteases.

Covalent immobilization requires previous activation of the support, followed by the covalent attachment of the biomolecule (for example, an enzyme) to the support. In some cases a spacer arm is need to guarantee an effective immobilization as when the immobilized molecule is very small in comparison with its counterpart (interaction of an immobilized inhibitor with a target enzyme).

The support activation with cyanogen bromide (CNBr) is one of the methods most frequently used although it has some disadvantages related to the high toxicity of the activation process and instability at pH below 5 and above 10 producing releasing of the immobilized ligand, among others.

Glyoxal-agarose immobilizes proteins via reversible Schiff's bases at alkaline pH values, where the density and reactivity of exposed lysine residues on the protein surface enable its immobilization on the support (Mateo *et al.*, 2005; Grazu *et al.*, 2006). The stabilization of immobilized enzyme is achieved through a final reduction reaction, in which the weak Schiff's bases are transformed in very stable secondary amino bonds and all remaining aldehyde on the solid support into inert hydroxyl groups (Blanco and Guisán, 1988)

Glyoxal-Sepharose® supports display suitable properties for immobilization-stabilization of proteins by covalent attachment. These include: very high reactivity of the glyoxal groups with non-ionized amino groups; very low steric hindrances in the reactive amino groups in the protein and glyoxal groups in the support; good geometrical congruence between the protein and the support surface; easy control of the activation degree, which leads to very high values; very high stability of the glyoxal groups, allowing very long enzyme–support multi-interaction and long storage times; very short spacer arms that permit to fix the relative positions of the groups implied in the immobilization, increasing the rigidity of the protein (Mateo *et al.*, 2006). The following graph summarizes the process:

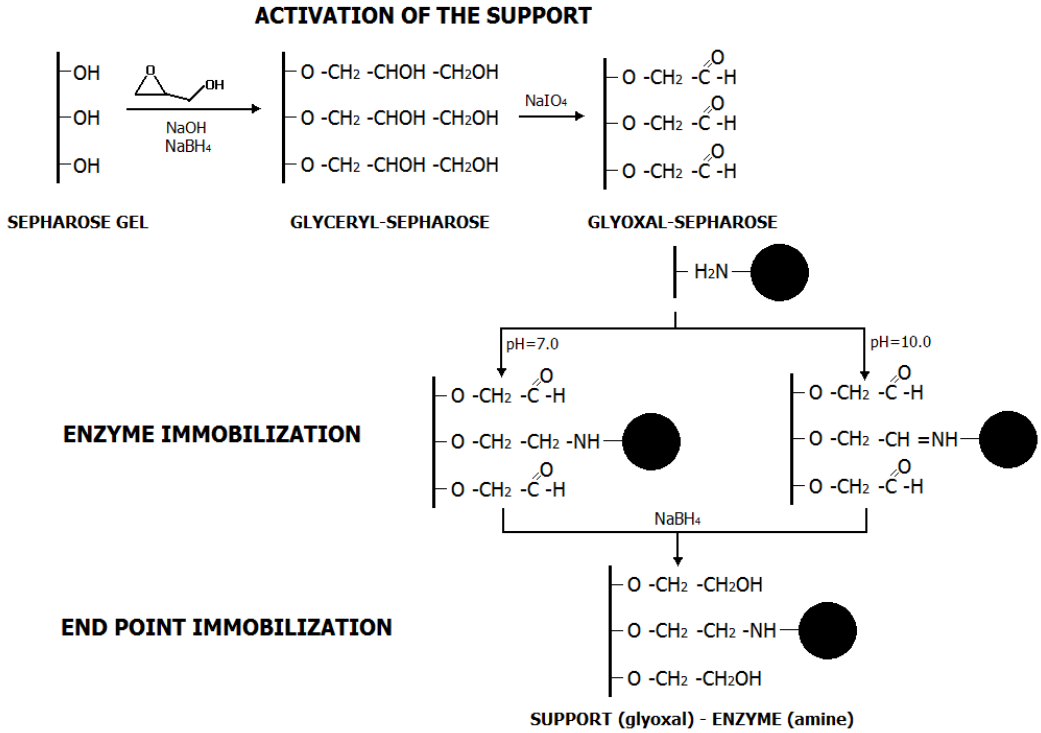


Figure 16. Sepharose activation with glyoxal groups and chemistry of the immobilization process (figure from Guisan *et al.*, 1996)

I.4.2.1.5. Cross-linking methods

The method is based on inter-and intramolecular cross-linking of enzyme molecules in the absence of a support. By this method there are three-dimensionally cross-linked enzyme aggregates that are insoluble in water. Cross-linking is achieved by using bi-or multifunctional reagents that bind covalently to the enzyme. These reagents are carbonyl compounds that react with lysine residues of the enzyme by forming Schiff bases, diazo compounds that bind to lysine residues, histidine, tyrosine, arginine or cysteine, isocyanates react with primary amine groups by peptide bond formation; alkyl iodides react with nucleophilic groups and iodoacetamide with cysteine residues, both by alkylation. Glutaraldehyde has been the most extensively used as cross-linking agent. The bond formed between the aldehyde and the enzyme is irreversible and resistant to extreme pH values and temperature, suggesting that the Schiff base is stabilized in the process.

The main disadvantages of this method are the difficulty of controlling the cross-linking reaction to obtain a high retention of catalytic activity, and the poor mechanical properties of these preparations, which do not allow good flow characteristics. For these reasons this method is usually employed in combination with others. However, an important advantage of this method is that immobilized preparations obtained are mainly composed of only the pure protein.

I.4.2.1.6. Entrapment methods

Entrapment methods are based on the inclusion of the enzyme within the three-dimensional network of a polymer or semipermeable membranes, so that the enzyme molecules can not be released but they allow the diffusion of substrates and products. These methods differ from those described above in that the enzyme does not bind to the gel matrix or membrane, which are considered as universal methods of protein immobilization, as they can be applied to different biomolecules, organelles and whole cells of different sizes and properties. This feature of the lack of interaction between the support and the enzyme minimizes the inactivation of the enzyme molecule and leads to obtain preparations with a high retention of functional activity.

However, the main disadvantage of this type of immobilization is that diffusional limitations are accentuated for substrates and products. It requires a rigorous control of the support pore size, so that no loss of enzyme occurs during its use, but ensures a good diffusion of the enzyme effectors. Use of this immobilization method is limited when the effectors of the enzyme are macromolecules. An alternative for these cases is to increase the size of the enzyme molecule by cross-linking with bifunctional reagents or binding (covalent or otherwise) to supports.

I.4.2.1.7. Microencapsulation methods

Enzymes can be immobilized inside microcapsules, which are prepared from organic polymers. Membranes that surround the enzyme must be semipermeable membranes in order to prevent the release of the enzyme molecules, but does allow the free diffusion of substrates and products.

The main advantages of this immobilization method are the presence of a large surface area for contact between the enzyme and the substrate in a relatively small volume and the ability to immobilize, in one step, several different enzymes, from soluble or immobilized enzymes previously obtained by another method. The disadvantages are related to the possible inactivation of the enzyme during the immobilization procedure, the addition of the enzyme, in some cases, into the walls of the membrane and the inability to immobilize enzymes acting on substrates of high molecular weight.

I.4.2.1.8. Soluble immobilized enzymes

The enzyme immobilization methods described above involve the modification of the enzyme molecule and/or its microenvironment with subsequent alterations in its functional activity. As a result of these procedures, very low enzyme activities are usually recovered. In order to use the native enzyme in its original state in continuous and repeated processes, it have designed systems that physically confine the enzyme in devices or systems, ultrafiltration membranes or hollow fibers.

The main advantages of this procedure are that it does not require chemical enzyme modification leading to a high functional activity. This procedure is also very easy to perform, operate, clean and regenerate, compared to other immobilization methods.

However, some disadvantages inherent of the method are the presence of diffusional mass transfer limitations due to the resistance imposed by the membrane towards the transport of species in solution, adsorption, the possibility of inactivation of the enzyme due to the vigorous agitation and the need for fine control of the residence time of low molecular weight substrates to ensure high conversions. However, this method can be combined with other immobilization methods (e.g. adsorption) that may reduce some of the above limitations.

I.4.3. Proteomic methods used for the identification and characterization of protease inhibitors

I.4.3.1. Intensity Fading MALDI-TOF MS

One of the most complex problems in the process of isolating molecules from natural sources is the stage of screening and identification. In the identification process of protease inhibitors, kinetic assays have been traditionally used and although they remain an obligatory tool based on functionality. The high-throughput techniques have as advantage their fast analysis and ability in many cases, to evaluate the direct interaction between molecules.

In the identification process of protease inhibitors, kinetic assays are an essential tool in the initial stage of identification, since they are based on the functionality (functional properties) of molecules to be detected. This method can be used in a high-throughput manner such as in immunoenzymatic tests, enzymatic microwell plate assays, among others. Moreover, the best strategy is the combination of these assays along with other identification methods based on the molecular interaction between the target enzyme and ligand such as the Biacore, IF MALDI-TOF MS, among others. These methods can confirm, based on a different criterion, the presence of the ligand (inhibitor) in those most promising positive extracts previously identified based on functional activity.

MALDI-TOF MS (Matrix-Assisted Laser Desorption/Ionization- Time-Of-Flight Mass Spectrometry) is a suitable analytical technique for ligand screening owing to its extremely high resolution and sensitivity (and thus low sample requirement), its fast analysis and easy automation, which facilitates the use of high-throughput protocols (Mann *et al.*, 2001).

Thus, a MALDI-TOF MS-based approach, called Intensity-Fading MALDI-TOF MS (IF MALDI-TOF MS), has been designed for just such a purpose. This methodology is based on the use of the MALDI ion intensities to detect quickly the formation of complexes between biomolecules in solution, where a target protein is one of the partners (protein-protein, protein-peptide, protein-organic molecule, and protein-nucleic acid complexes) (Villanueva *et al.*, 2003). In this strategy of direct detection, the target-ligand interaction is detected through the decrease (fading) of the molecular ion intensities of the partners as directly compared to the MALDI mass spectrum of the initial mixture following the addition of the target molecule (figure 17) (Villanueva *et al.*, 2003; Yanes *et al.*, 2007a).

Furthermore, IF MALDI-TOF MS technique that included an affinity step (with the target protein bound to a solid support, e.g., agarose/sepharose beads) has been established in order to resolve the target-ligand complex from unreacted analytes, followed by analysis with MALDI-TOF MS (Yanes *et al.*, 2005) (figure 17). In this methodology the complex is detected through the decrease of the molecular ion intensities of the ligands as directly compared with the MALDI mass spectrum of the mixture after the addition of the target protein. Because the protein is immobilized on beads, an additional step can be performed to assess the specificity of the assay; the protein-conjugated beads are washed after the binding reaction to discard unbound molecules and, by lowering the pH, the species bound to the protein can be eluted for further MS analysis. Breaking the interactions of the protein with the putative binding molecules of the mixture causes the faded signals to reappear in a subsequent MS analysis. Therefore, these signals can be attributed to molecules that specifically bind to the target protein (Yanes *et al.*, 2007b). This approach of indirect detection can be used with an immobilized protease as target enzyme for the screening of protease inhibitors and vice versa.

In some experiments, detection, fading or disappearance of ion signal intensities is not clearly visualized in the control mass spectrum, probably due to signal suppression effects (Yanes *et al.*, 2005), excessive amount of sample applied, complexity of the extract, among others. However, it is observed the reappearance of ion signals in the elution fraction, which indicates that some molecules were specifically bound to the target protease, although they were not displayed in the control mass spectrum. In these cases, the presence of ion signals in the elution fraction is also considered as a positive result.

The potential of the approach has been examined in several examples of model interactions, mainly involving small non-protein and protainaceous protease inhibitors, at both the qualitative and semiquantitative levels. Using this method, different protein ligands of proteolytic enzymes in total extracts of invertebrate organisms have been identified in a simple way (Villanueva *et al.*, 2003).

IF MALDI-TOF MS has proved to be a successful strategy in the analysis of non-covalent interactions between proteases and protease inhibitors. Thus, it has been demonstrated the selective interaction of CPA and protainaceous inhibitors such as LCI and PCI (Villanueva *et* al., 2003; Yanes *et al.*, 2006); the molecular recognition between the non-peptide inhibitor E-64 and the stefin A with papain used as target protease (Villanueva *et al.*, 2003). Regarding serine protease inhibitors, it has been studied the interaction of BPTI and trypsin, as well as the identification of more than 75 molecules present in the medical leech *H. medicinalis* that selectively interact with trypsin (Yanes *et al.*, 2005). Furthermore, it has been corroborated the interaction of ShPI in the crude extract of *S. helianthus* with trypsin (Villanueva *et al.*, 2003; Yanes *et al.*, 2004; Yanes *et al.*, 2007).

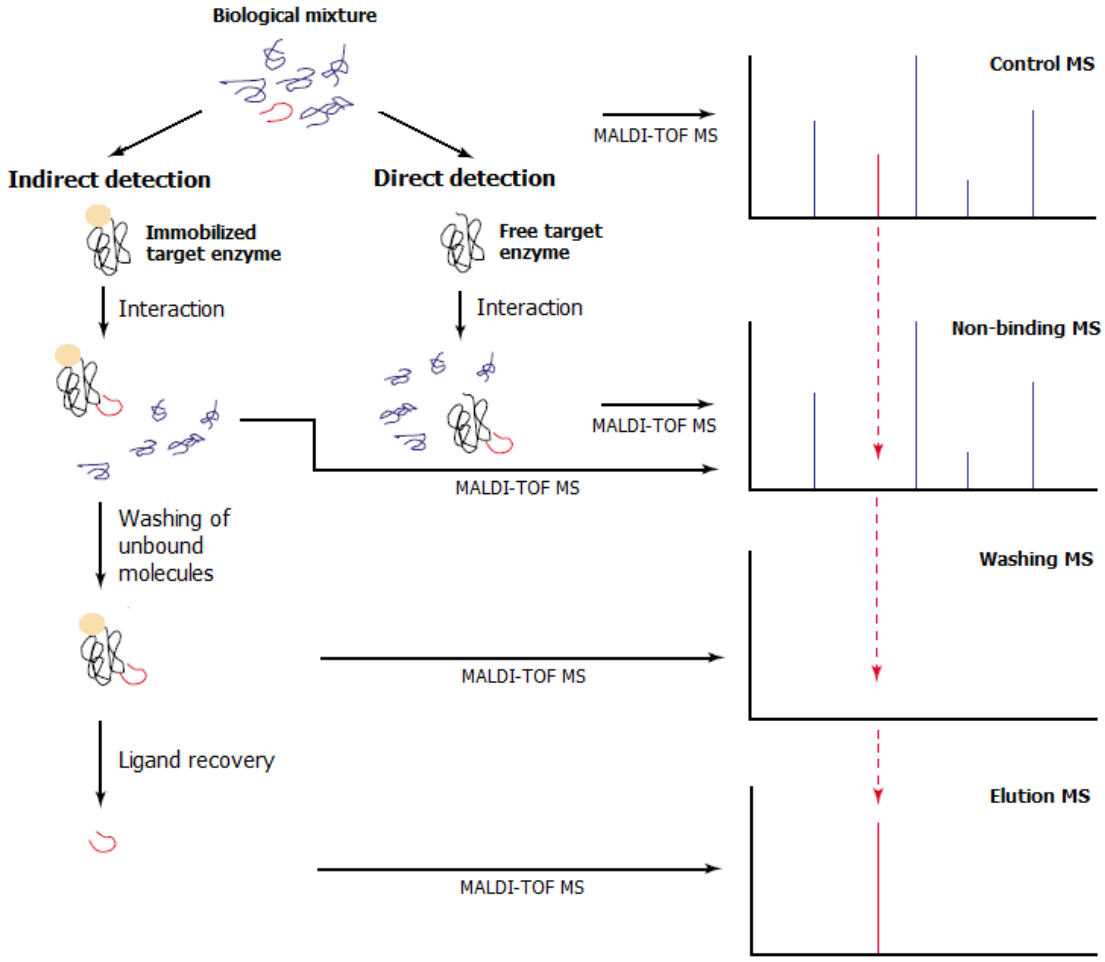


Figure 17. Strategy of IF MALDI-TOF MS in applications to PI discovery (figure adapted from Yanes *et al.*, 2004)

I.4.3.2. Tandem mass spectrometry and fragmentation approaches

Tandem mass spectrometry (MS/MS) experiments are major tools used in protein identification. Mass spectrometers measure the mass/ charge ratio of analytes; for protein studies, this can include intact proteins and protein complexes (Sobota and Robinson, 2002), fragment ions produced by gas-phase activation of protein ions (top-down sequencing) (Loo *et al.*, 1990; Senko *et al.*, 1994; Little *et al.*, 1994; Breuker *et al.*, 2002; Kelleher N.L., 2004), peptides produced by enzymatic or chemical digestion of proteins (mass mapping) (Kleno *et al.*, 2004; Pan *et al.*, 2003), and fragment ions produced by gas-phase activation of mass-selected peptide ions (bottom-up sequencing) (Yates J.R., 1998).

The application of mass spectrometry and MS/MS to proteomics takes advantage of the vast and growing array of genome and protein data stored in databases. The information produced by the mass spectrometer, lists of peak intensities and mass-to-charge (m/z) values, can be manipulated and compared with lists generated from "theoretical" digestion of a protein or "theoretical" fragmentation of a peptide.

Applications to analyze even smaller quantities of sample are driving the development of more sensitive mass spectrometers, as well as low flow, high resolution separation technologies, to provide structural information on individual components in complex mixtures of thousands of proteins derived from biological samples. Protein identification by mass spectrometry requires an interplay between mass spectrometry instrumentation (how molecules are ionized, activated, and detected) and gas-phase peptide chemistry (which bonds are broken, at what rate, and how cleavage depends on factors such as peptide/protein charge state, size, composition, and sequence) (Wysocki *et al.*, 2005).

I.4.3.2.1. Dissociation techniques in mass spectrometry

Peptide or protein sequence identification by mass spectrometry involves fragmentation of a peptide or protein to produce smaller m/z fragments; ideally, measured m/z values of these pieces can be assembled to produce the original sequence. Cleavage is commonly accepted to occur predominantly through charge-directed pathways, i.e., cleavage is initiated by a charge that is transferred to the vicinity of the cleavage site.

Fragment ion types: A nomenclature is used to describe the fragment ion types that are produced by cleavage of different bonds along the peptide backbone and/or side chain (Roepstorff P., 1984; Johnson *et al.*, 1987; Johnson *et al.*, 1988) (figure 18A). Typical ion structures will be illustrated below, although actual structures of a particular fragment ion are often only inferred from model studies. Cleavage of the backbone typically occurs at the peptide amide bond to produce ***b*** ions, if the amino terminal fragment retains the charge, or ***y*** ions, if the carboxy-terminal fragment retains the charge (figures 18B and 18C).

In the case of multiply charged ions, a charge separation can occur to produce complementary ion pairs (e.g., a doubly charged ion can fragment to produce a $\boldsymbol{b_n/y_m}$ ion pair where $n + m$ = total residues in peptide). Both partners of the complementary pair are not always detected in equal abundance, because they are not equally stable against further fragmentation or because instrument discrimination may enhance or diminish one partner of the pair. Although ***b*** and ***y*** ions are considered to be the most useful sequence ion types, because they correspond to cleavage of the amide bond, other ion types are observed and used in spectral interpretation or database searches. These include ***a*** ions (figure 18D), which correspond formally to loss of CO from a ***b*** ion; a m/z difference of 28 between two peaks suggests an ***a–b*** ion pair and is useful in identifying the ion series to which the peaks belong.

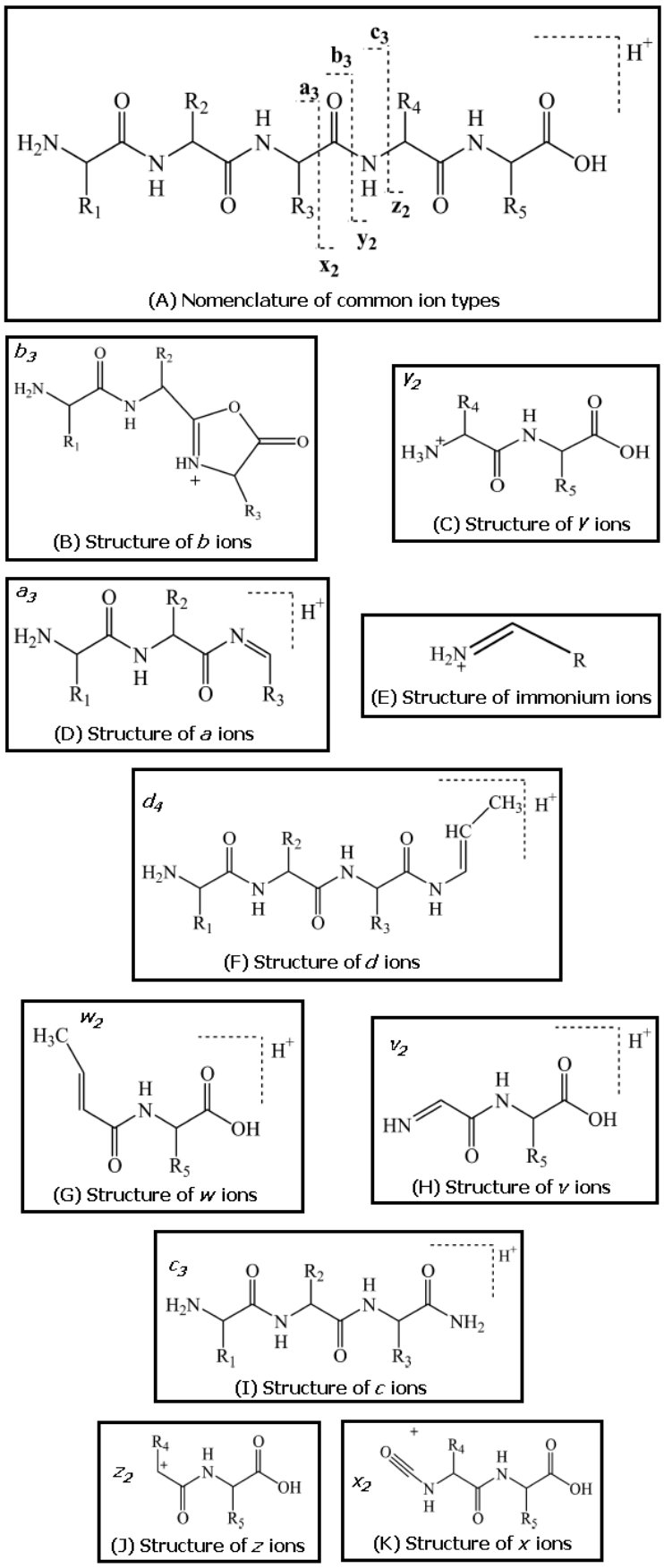


Figure 18. Nomenclature and representative structures of fragment ion types (figure adapted from Wysocki *et al.*, 2005)

The ***y*** series is sometimes accompanied by peaks formally corresponding to loss of NH3 from the ***y*** ions, allowing designation of the higher m/z ion of each delta 17 pair as belonging to the ***y*** ion series. Ions that correspond to immonium ions (figure 18E), or fragments of immonium ions, of individual amino acid residues in a peptide are often detected, even for residues from the internal portion of the sequence (Falick *et al.*, 1993).

Ions that correspond to cleavage of a side-chain bond in addition to backbone cleavage(s) are referred to as side-chain cleavage ions. These are designated as ***d***, ***v***, and ***w*** ions, and allow distinction between isomeric or isobaric ions (e.g., Ile from Leu) (figures 18F, 18G, and 18H).

Homolytic cleavage at the N-Cα bond produces ***c*** ions when charges are present in the amino-terminal fragment, and ***z*** · ions when charges are present at the carboxyl-terminal fragment (figures 18I, 18J). The ***z*** · ions are unique and different from the generic ***z*** ion structure shown in figure 18J in that they are odd electron radical cations. Ions of type ***x*** have recently been reported for photodissociation experiments (Thompson *et al.*, 2003) (figure 18K).

I.4.3.2.1.1. MALDI Post-Source Decay

With reflectron time-of-flight mass spectrometers, it is in theory possible to obtain structural information on a selected quasimolecular ion by mass analysis of daughter ions issued from in-flight fragmentation of the parent ion. Intact molecular ions leaving the ion source and having acquired sufficient internal energy during the desorption process (photoactivation, low energy collisions, etc.) can release this energy by undergoing fragmentation while traveling the first field-free drift path of the instrument (called post-source decay, or PSD (Spengler *et al.*, 1992; Kaufmann *et al.*, 1993; Nhuknan *et al.*, 1996; Stimson *et al.*, 1997).

The fragment ions have the same velocity as their precursor ions but have different energy as a function of their mass. Fragment ions are then discriminated as a function of their kinetic energy (thus their mass) by the time dispersions induced by the electrostatic reflector. Larger fragment ions (with higher kinetic energy) will penetrate deeper into the reflectron than smaller fragment ions and will appear at a later time on the resulting reflectron time-of-flight spectrum (Chaurand *et al.*, 1999).

The principle of MALDI post-source decay TOF mass spectrometry is illustrated in figure 19. With typical reflectors (single stage, two stage, gridded/gridless (Stahl-Zeng and Hillenkamp, 1996), it is necessary to progressively decrease the potential(s) applied to the reflector to acquire a complete product-ion spectrum by concatenation of several sections. Different from that, a so-called curved-field reflector is able to disperse all PSD ions over the acquired time-of-flight spectrum with the same set of electrostatic potentials applied to the reflector. A complete PSD ion spectrum can thus be acquired within a single step (Cornish *et al.*, 1994). The performance of this approach, however, has not been shown so far to be superior to stepped-mode PSD instruments.

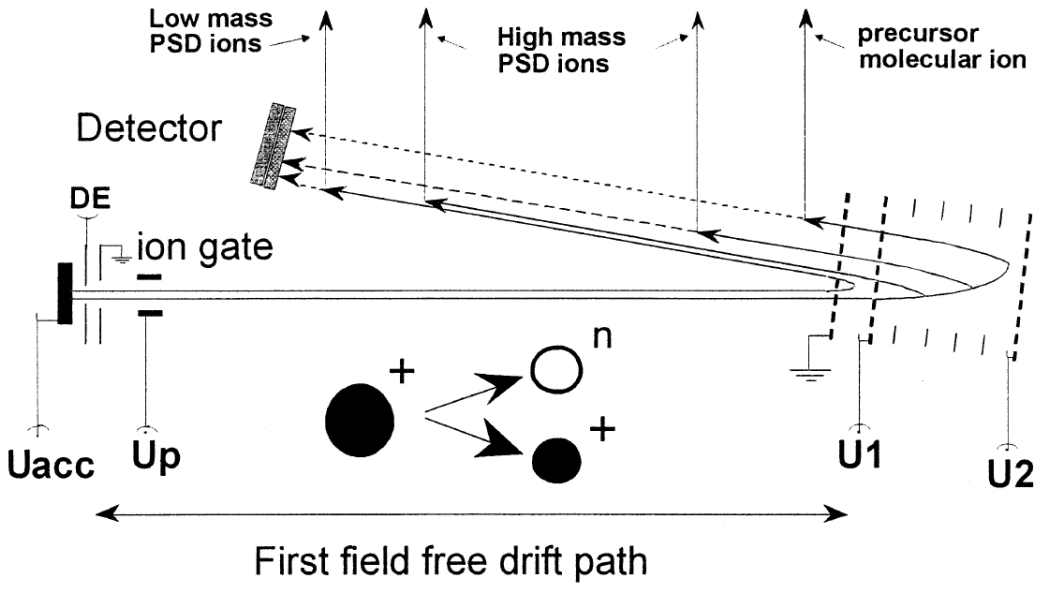


Figure 19. Principle of MALDI-PSD time-of-flight mass spectrometry (figure from Chaurand *et* al., 1999)

Peptide sequencing has become a major field of application of the MALDI-PSD technique, driven by the growing analytical demand. Generally a PSD spectrum recorded from a singly charged peptidic molecule contains mass signals coming from immonium ions (Ambihapathy *et al.*, 1997), N-terminal fragment ions (***a***-, ***b***-, ***c***-, and ***d***- type ions) (Kaufmann *et al.*, 1994; Yalcin *et al.*, 1995; Yalcin *et al.*, 1996), C-terminal fragment ions (***x***-, ***y***-, and ***z***-type ions) (Biemann K., 1990) and internal (double chain cleavage) fragment ions (***AY***-, ***BY***-, and ***CY***-type ions). Furthermore, all these ions potentially display satellite peaks at minus 17 mass units due to loss of ammonia (NH_3) from lysine or arginine, or at minus 18 mass units due to loss of H_2O from serine or threonine.

I.4.3.2.1.2. In-source decay

In-source decay (ISD) fragmentation appears at higher laser fluence than those conventionally used in MALDI studies. Higher laser fluence gives internal energy to the analyte that favors the backbone fragmentation (Pfeifer *et al.*, 1999; Purcell and Gorman, 2001). In ISD, proteins are principally cleaved at the N–Cα bond of the peptide backbone, giving $\boldsymbol{c_n}^-$ and ($\boldsymbol{z_n+2}$)-fragment types (Brown *et al.*, 1996; Brown *et al.*, 1997; Takayama M., 2001a).

A representation of the ISD fragmentation is shown in figure 20, where four steps are described: first, an electronic excitation of the MALDI matrix by the photon absorption of the laser; second, an intermolecular hydrogen transfer from the matrix to the peptide backbone carbonyloxygen; third, the formation of a peptide radical; fourth, the cleavage of the NH-CH bonds to form *c*- and (*z+2*)-ions. By increasing the laser power, formation of $\boldsymbol{b_n}^-$, $\boldsymbol{y_n}^-$, $\boldsymbol{a_n}^-$, and $\boldsymbol{x_n}^-$ions may occur (Takayama and Tsugita, 1998; Takayama, 2001b). Other fragments like $\boldsymbol{d_n}$, $\boldsymbol{v_n}$, and $\boldsymbol{w_n}$ are rarely present in the mass spectrum (Marzilli *et al.*, 2000; Takayama M., 2001b). The distinction between isobaric amino acids, like isoleucine/leucine (Ile/Leu) and lysine/glutamine (Lys/Gln), is thus impossible.

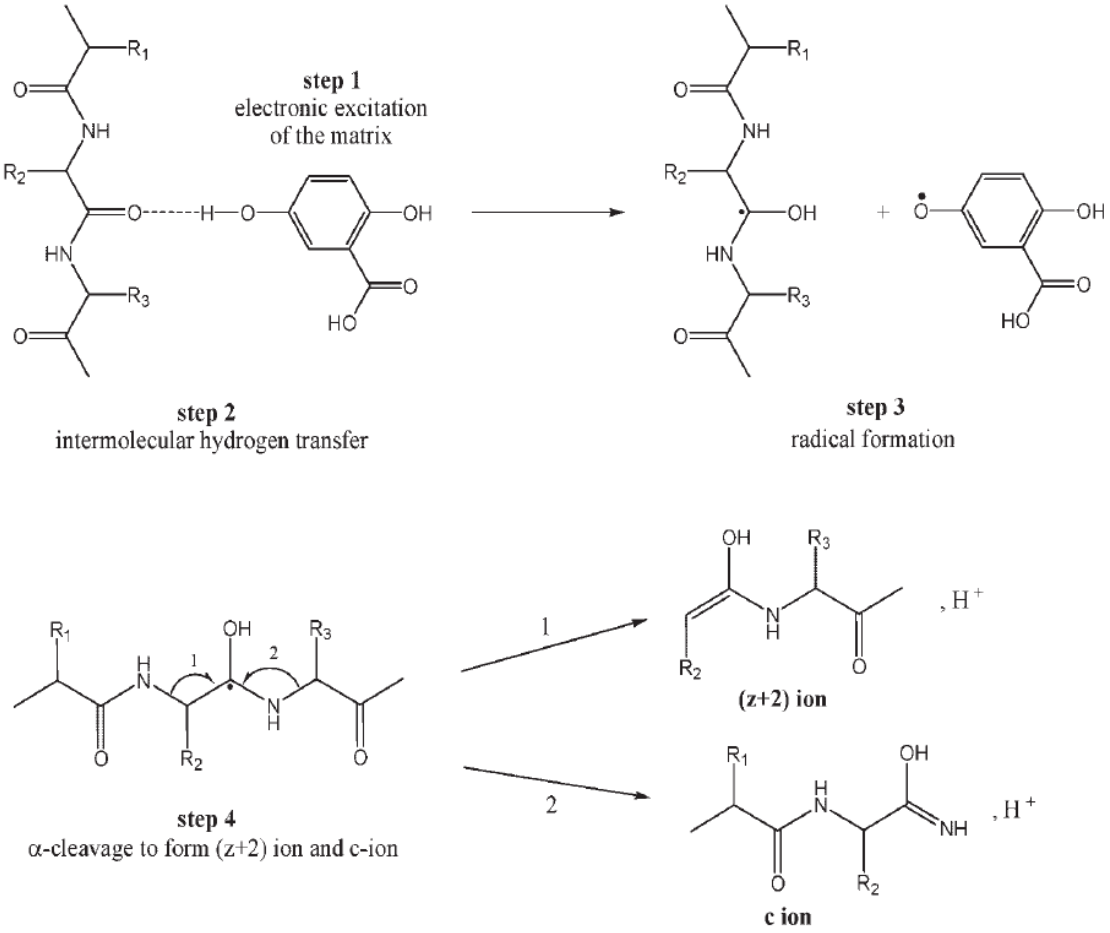


Figure 20. Illustration of the formation of *c*- and (*z+2*)-ions by ISD (figure from Hardouin J., 2007)

The choice of the MALDI matrix can dramatically affect the mass spectrum quality and, in the case of ISD mode, can increase the extent of in-source fragmentation (Brown *et al.*, 1997). Different matrices have been tested to observe their ISD fragmentation efficiency (Brown *et al.*,, 1996; Brown *et al.*, 1997; Takayama M., 2001b). Hydroxybenzoic acid matrix types seem to be the best candidates and 2,5-DHB matrix favours the c_n- and y_n-series ions and was far superior to other di- or trihydroxybenzoic acid derivates (Takayama, 2001b).

Although ISD is not efficient in the case of protein mixture, it is a great approach to obtain sequence information of intact purified peptides or proteins that needs no enzymatic digestion (Reiber *et al.*, 1998). Even if the N-terminal of the protein is modified, the ISD fragmentation occurs, contrary what occurs during Edman sequencing. ISD is not limited to small molecular mass peptide fragments issued from the digestion of protein in contrast to MS/MS or PSD that fragment only peptides containing around 10–20 amino acids (Hardouin J., 2007).

ISD fragmentation data are proven to be very complementary to PSD MALDI in determining unknown sequences because of the significant differences in the observed fragmentions. Recent advances in instrumentation (new available configurations of tandem mass spectrometers) coupled with the MALDI technique have led to powerful new instruments that are especially well suited for study of ISD fragments. This approach allows promises for protein sequencing and characterizing Post-translational modifications (PTMs) (Hardouin J., 2007).

I.4.3.2.1.3. Collision-induced dissociation

Collision-induced dissociation (CID), also referred to as collisionally activated dissociation (CAD), was first described by Jennings K.R., 1968 and McLafferty and Bryce, 1967. In CID, precursor ions are accelerated to higher kinetic energy and allowed to collide with neutral gas atoms or molecules (typically helium, nitrogen or argon). Following this inelastic collision, some of the ion's kinetic energy is converted into internal vibrational energy resulting in bond cleavage. CID is classed as a slow activation method in which multiple discrete activation events (collisions) occur throughout the activation period. The activation period is as long, or longer, than the timescale for unimolecular fragmentation of the ion. If the activation events are in competition with deactivation events, the process is considered a 'very slow' activation method, also known as a 'slow-heating' method (McLuckey and Goeringer, 1997).

For peptides and proteins, the lowest energy pathway tends to be cleavage of the amide N–CO bond to produce *b* and *y* fragment ions (Roepstorff and Fohlman, 1984) (figure 21). The fragmentation observed following peptide CID can be explained by the 'mobile proton' model (Summerfield *et al.*, 1997; Dongre *et al.*, 1996; Wysocki *et al.*, 2000). Collisional activation of the peptide ion mobilises a proton from any basic site (with varying efficiencies), i.e. N-terminus or amino acid residue, e.g., arginine or lysine, to a backbone heteroatom (either a carbonyl oxygen or amide nitrogen) instigating charge-directed fragmentation. CID is the most widely used MS/MS technique in bottom-up proteomics. It is a fast and efficient way for sequencing of linear and cyclic peptides (Lippstreufisher and Gross, 1985; Eckart *et al.*, 1985).

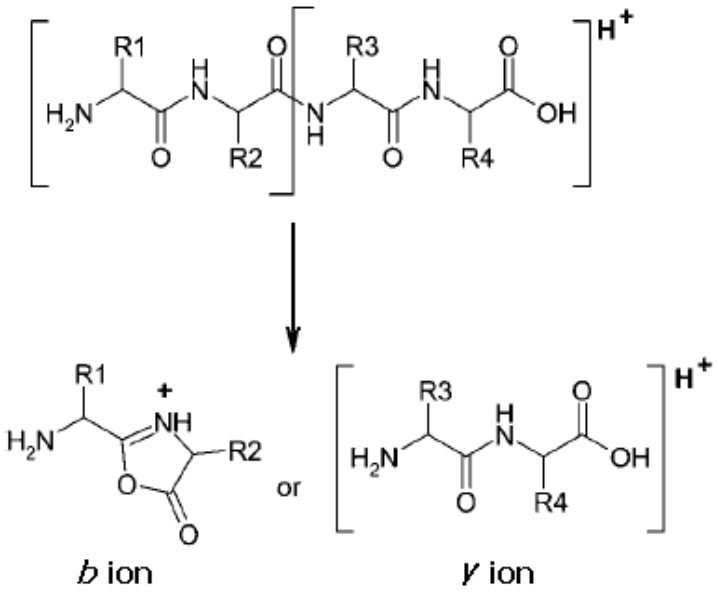


Figure 21. Fragmentation of the peptide amide N–CO bond produced by CID (figure from Jones and Cooper, 2011)

I.4.3.2.1.4. Electron capture dissociation

Electron capture dissociation (ECD) was first described by Zubarev *et al.* in 1998 and has since proved an invaluable MS/MS technique for biomolecular analysis (Cooper *et al.*, 2005), particularly for peptides and proteins. The advantages of ECD for peptide/protein analysis are firstly that cleavage is random (the only exception being N-terminal to proline, (Cooper *et al.*, 2003)), therefore sequence coverage tends to be higher for ECD than for slow-heating techniques (Axelsson *et al.*, 1999; Zubarev, 2000), and secondly, that labile post-translational modifications (PTMs) are retained on peptide/protein backbone fragments (Kelleher *et al.*, 1999).

In ECD, multiply charged ions of interest are irradiated with low energy electrons (<0.2 eV) producing charge-reduced species, which dissociate along radical-driven pathways. The dominant peptide fragmentation pathways proceed via cleavage of the backbone N–Cα bond to give ***c*** and ***z*** · fragment ions (figure 22A) (which may be accompanied by hydrogen atom transfer to give and, more commonly, ***z*** fragment ions) and disulfide bonds (figure 22B) (Zubarev *et al.*, 2002; Savitski *et al.*, 2007; Zubarev *et al.*, 1999). However, ECD is essentially limited to Fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS), due to the requirement to trap electrons, with the attendant cost implications (Jones and Cooper, 2011).

ECD has been proposed as a "non-ergodic" fragmentation process, in which the intra-molecular energy randomization is slower than the ECD cleavages (Zubarev R.A., 2003; Cooper *et al.*, 2005). This unique non-ergodic feature has been exemplified by the ability of ECD in preserving the labile side chain modifications groups on the peptide backbone fragments. This is in contrast to CID and other conventional fragmentation techniques, which typically eject the labile modifications prior to the peptide backbone dissociation (Zubarev R.A., 2003).

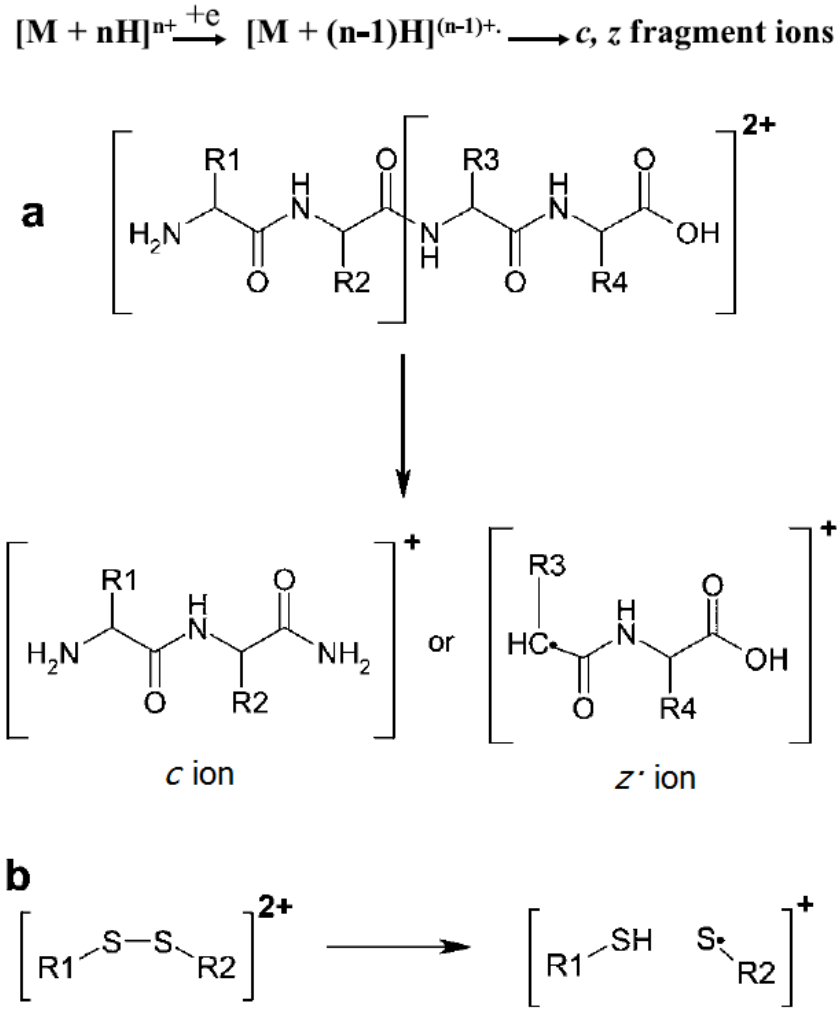


Figure 22. ECD fragmentation pathway. (a) Representation of the ECD process and the nomenclature used for assignment of product ions in the ECD spectra of peptides and proteins. **(b)** Fragmentation of a disulfide bond (figure from Jones and Cooper, 2011).

I.4.3.2.1.5. Electron transfer dissociation

Electron transfer dissociation (ETD) is a new method to fragment peptides that utilizes ion/ion chemistry (Syka *et al.*, 2004; Coon *et al.*, 2004; Coon *et al.*, 2005; Pitteri *et al.*, 2005). ETD fragments peptides by transferring an electron from a radical anion to a protonated peptide. This induces fragmentation of the peptide backbone, causing cleavage of the N-Cα bond (figure 23) just as ECD does. This creates complementary ***c*** and ***z***-type ions instead of the typical ***b*** and ***y***-type ions observed in CID. Although the exact mechanism of ECD and ETD are debated, ETD preserves PTMs that are labile by CID and sequence information on the peptide can be obtained. ETD uses a radio frequency (RF) quadrupole ion trapping device instead of an FT-ICRMS for ion trapping and detection (Syka *et al.*, 2004). RF ion trap mass spectrometers are low-cost, low-maintenance, and widely accessible as compared to the FT-ICR-MS.

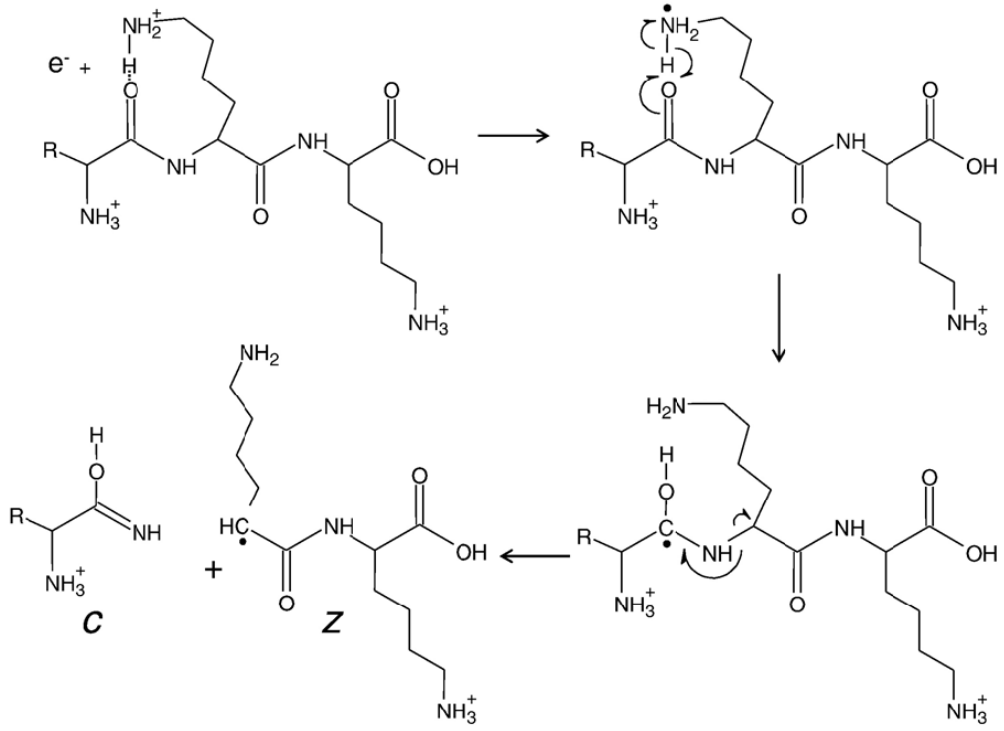


Figure 23. ETD fragmentation pathway (figure from Syka *et al.*, 2004)

I.4.3.2.1.6. Electron-detachment dissociation

A recent method for inducing backbone fragmentation, termed electron-detachment dissociation (EDD) mass spectrometry, has shown substantial promise as a means of generating CR-C bond cleavages (Kjeldsen *et al.*, 2005). This experiment is carried out in negative-ion mode and involves using electrons with kinetic energies above 10 eV to detach electrons from the negative sites in the parent peptide or protein ion to convert the negative sites to radical centers. The locations of the negative charges in such samples include amide nitrogen atoms along the backbone (as well, of course, as C-terminus and side-chain carboxylate groups) (Kjeldsen *et al.*, 2005). When an electron is ejected from an amide nitrogen anion, nitrogen-centered radicals such as those shown in figure 24 are formed.

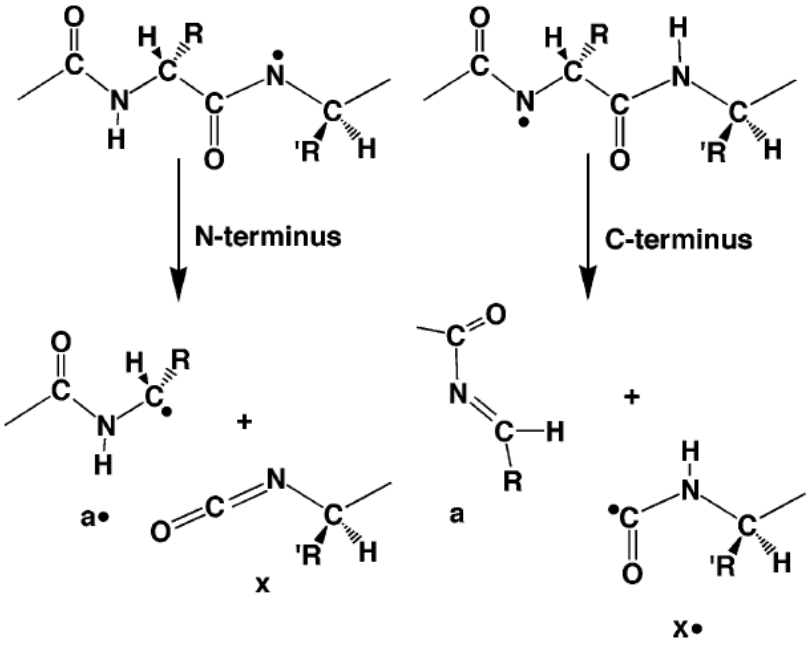


Figure 24. EDD fragmentation pathway (figure from Anusiewicz *et al.*, 2005)

Upon forming such a nitrogen-centered radical, two Cα-C bond cleavage paths might be expected to occur. The path shown in figure 24 involves cleavage of a Cα-C bond in which the Cα carbon atom is adjacent to the nitrogen radical, while that shown on the left involves cleavage in which the Cα carbon atom is adjacent to an amide -NH- group rather than the nitrogen radical. It is found that Cα-C bond cleavage is favoured over side-chain loss, although loss of a tyrosine side chain may compete with Cα-C cleavage because the tyrosine radical formed can delocalize its unpaired electron over its aromatic ring. In addition, it is found that fragmentation of the nitrogen-centered radicals formed in EDD results in cleavage to produce so-called ***a*** **˙/*x*** fragments rather than ***a/x*** ˙ fragments both because producing the former involves a significantly smaller barrier and is nearly thermoneutral, while cleavage to yield ***a/x*** ˙ is significantly endothermic (Anusiewicz *et al.*, 2005).

I.4.3.2.1.7. Photodissociation

It has been demonstrated that photodissociation (PD) could be a useful technique for tandem mass spectrometry of protonated peptides generated by MALDI (Barbacci *et al.*, 1999; Hettick *et al.*, 2001; Thompson *et al.*, 2003; Thompson *et al.*, 2004; Oh *et al.*, 2004a; Oh *et al.*, 2004b). The performance of a tandem time-of-flight (TOF) mass spectrometer built for PD studies, which records the PD and PSD spectra simultaneously, was reported (Oh *et al.*, 2004a; Oh *et al.*, 2004b).

The PD technique has some advantages over the usual CID tandem TOF mass spectrometry (Cornish *et al.*, 1993; Morris *et al.*, 1996; Medzihradszhy *et al.*, 2000). The fact that collision gas is not needed, that the ion trajectory is not disturbed by the activation step, and that monoisotopic selection of the precursor ion is possible through the ion pulse-laser pulse synchronization, are the obvious advantages of PD (Barbacci *et al.*, 1999). Also, the fact that the initial energy deposition site (the chromophore site) can be selected by choosing the PD wavelength is a potential advantage which may be found useful in the future (Hettick *et al.*, 2001). It was observed that the PD yield could be as good as that of CID and depended on the intensity of the PD laser. As mentioned above, photodissociation produces ***x***-type ions (Thompson *et al.*, 2003).

I.4.3.2.1.8. Infrared multiphoton dissociation

Infrared multiphoton dissociation (IRMPD) can be used to dissociate such thermally stable ions (Little *et al.*, 1994). In this technique, ions undergo stepwise vibrational excitation. Low-power cw CO_2 radiation gives fragments similar to those obtained by BIRD and collision-induced dissociation (Little *et al.*, 1994; Tonner *et al.*, 1997). The utility of IRMPD for efficient fragmentation of multiply charged ions from proteins and oligonucleotides has been demonstrated (Little *et al.*, 1994). Since then low-power cw IRMPD has been widely used for structural characterization of large biologically relevant molecules in FTICR MS (Laskin and Futrell, 2005).

IRMPD dissociation efficiency, defined as the relative abundance of the precursor ion in the IRMPD mass spectrum, depends on the number of ions irradiated by the laser, the laser intensity and irradiation time. An important intrinsic property of IRMPD is that it is not mass selective: all ions that are in the optical path of the laser beam are excited simultaneously. This can be advantageous for sequencing large ions produced by electrospray with a distribution of charge states. Simultaneous irradiation of different charge states produces rich MS/MS spectra that contain both intact precursor ions (lower charge states) and sequence specific fragment ions originating from higher charge state precursors (Laskin and Futrell, 2005).

I.4.3.2.1.9. Blackbody infrared radiative dissociation

Blackbody infrared radiative dissociation (BIRD) describes the observation of ion-dissociation reactions at essentially zero pressure by the ambient blackbody radiation field, which is usually studied in the ion-trapping ion cyclotron resonance (ICR) mass spectrometer. Focussing on the quantitative observation of the temperature dependence of BIRD rates, methods are developed for connecting BIRD observations with activation parameters and dissociation thermochemistry. Three regimes are differentiated and described, comprising large molecules, small molecules, and intermediate-sized molecules (Dunbar, R.C., 2004).

BIRD has spread over a wide variety of applications, such as the characterization of solvent– molecule detachment from solvated ions; dissociation reactions of biomolecules (polypeptides, oligonucleotides, complexes involving polysaccharides) and the structural information to be deduced from them; and dissociations of proton-bound and metal–ion-containing complexes. Polypeptide and protein ions were among the first species to be electrosprayed and studied by BIRD (Price *et al.*, 1996; Price *et al.*, 1997; Schnier *et al.*, 1996; Schnier *et al.*, 1997; Jockusch *et al.*, 1997). Several techniques related to BIRD are noted, including collisional dissociation in the FTICR ion trap; high-pressure thermal dissociation in quadrupole ion traps and in heated inlet capillary regions; hot-filament assisted dissociation; as well as the above mentioned IRMPD (Dunbar, R.C., 2004).

I.4.3.2.1.10. Surface-induced dissociation

Surface-induced dissociation (SID) is analogous to CID, except that a surface replaces the neutral gas as the collision target. A typical ion–surface collision event is illustrated in figure 25, where surface collision event deposits energy into the precursor ion. Collisions can generate fragment ions, neutralized precursor molecules, sputtered surface atoms, and ion–surface reaction products. The incorporation of a surface into a mass spectrometer for ion activation was pioneered in the laboratory of R. Graham Cooks in the mid-1970s and early 1980s (Cooks *et al.*, 1975a; Cooks *et al.*, 1975b; Mabud *et al.*, 1985). Since that time, collisions of low-energy (eV) organic ions with surfaces within the tandem mass spectrometer have been valuable for analyzing surface composition, characterizing reactions between organic projectile ions and surface adsorbates, chemically modifying surfaces, and determining projectile ion structure. A major motivation for development of SID is that energy transfer to ionic projectiles can be improved by increasing the mass of the collision target (Wysocki *et al.*, 2008).

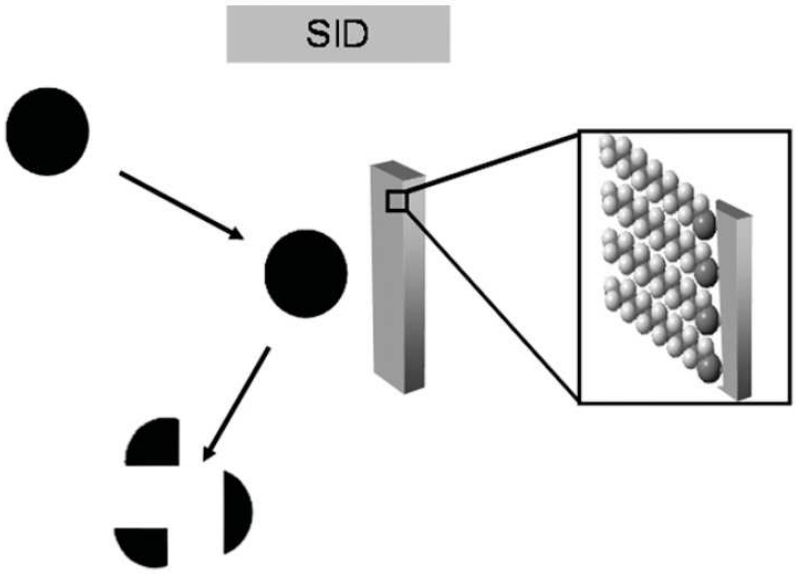


Figure 25. Diagram of surface-induced dissociation (figure from Wysocki *et al.*, 2008)

Surface-induced dissociation has been implemented in many different instrument platforms including magnetic sector-electric sector (BE)– surface–electric sector quadrupole (EQ), reflectron time of- flight mass spectrometry (TOF MS), Q–surface–Q, TOF–surface–TOF, Q–surface–TOF, Fourier transform ion cyclotron resonance (FT-ICR), and matrix-assisted laser desorption ionization–ion mobility–surface-induced dissociation–time-of-flight (MALDI-IM-SID-TOF) (Wysocki *et al.*, 2008).

Surface-induced dissociation was influential in helping to develop a general framework for peptide fragmentation through the energy-resolved study of systematically varied model peptides. SID results from the Wysocki group contributed to the mobile proton model, a description of peptide dissociation in which fragmentation is initiated by rapid intra-molecular proton transfer among backbone protonation sites (Wysocki *et al.*, 2008).

I.4.3.2.1.11. Charge-remote fragmentation

Charge-remote fragmentation is a class of gas-phase decompositions that occur physically remote from the charge site (Adams J., 1990; Gross M.L., 1992; Adams and Songer, 1993). The reactions, analogous to gas-phase thermolysis (Adams and Gross, 1989), are independent of the charge status because there appears to be no important interaction between reaction sites and the charge. High-energy collisional activation (CA) always favours charge-remote fragmentation, although sometimes these reactions do occur under low-energy CA or even as a metastable-ion process (Cheng and Gross, 2000).

Charge-remote fragmentation has been proven useful in the structural determination of long-chain or poly-ring molecules. Since 1983 (Tomer *et al.*, 1983), structures of a variety of synthetic and natural molecules, including fatty acids and their derivatives, glycerophospholipids, glycolipids, ceramides, carbohydrates, antibiotics, steroids, and peptides, have been determined by charge-remote fragmentations (Adams J., 1990; Gross M.L., 1992).

I.4.3.3. Peptide and protein *de novo* sequencing by mass spectrometry

So far the genome sequences of most organisms are still unknown. Even for those that are known, modifications such as post-translational events may hinder the identification of all or part of the protein sequence, or at least the definition of the modifications. Thus, complete characterization of the protein primary structure often requires determination of the protein sequence with minimal assistance from genomic data — *de novo* protein sequencing. Early *de novo* protein sequencing measurements relied on Edman degradation of the protein, but mass spectrometry (MS) has reduced the need for this technique because it is more sensitive and provides higher sample throughput. It can also cope better with protein mixtures and with modifications to the protein N terminus (Standing K., 2003).

Considering all amino acid sequence combinations that are theoretically possible, only a very minor portion of protein sequences occurs in nature, and therefore a short peptide sequence is already highly protein-specific. This situation effects that a database-supported, probability-based annotation of peptide MS/MS spectra leads to protein identification at a high level of confidence from fragmentary sequence information. Thus, database-supported protein identification is very effective, but it precludes the recognition of all peptides not present in the reference database.

In spite of the continuously growing sequence databases, *de novo* sequencing of peptides, i.e. sequencing without assistance of a linear sequence database, is still essential in several analytical situations. For example, analyses of protein sequence variants or their splice isoforms require *de novo* sequencing, as well as protein analysis from organisms with unsequenced genomes. In addition, *de novo* sequencing is essential for analysis of peptides containing non-proteinic or modified amino acids, as typically present, e.g. in bioactive peptides of bacteria or fungi (Degenkolb *et al.*, 2008). The performance of both the MS/MS and of the LC part influences the utility of an LC-MS system for *de novo* sequencing. This is because the significance of peptide MS/MS data is connected with the purity of the peptide ions selected for fragmentation (Seidler *et al.*, 2010).

The majority of peptide *de novo* sequencing has been performed using CID, although *de novo* sequencing using fragmentation data from a combination of CID and ECD or ETD spectra gives significantly improved results when compared with *de novo* sequencing based on CID data only. The better *de novo* sequencing performance is due to the complementary fragmentation information contained in the two types of spectra, which results in more complete ion series covering the peptide sequence (Bertsch *et* al., 2009; Seidler *et al.*, 2010; Jones and Cooper, 2011).

I.4.3.4. Bottom-up protein identification

Protein identification by MS can be performed using sequence-specific peptide fragmentation or peptide mass fingerprinting (PMF), also known as peptide mass mapping (Aebersold and Goodlett, 2001). The standard approach to identify proteins includes separation of proteins by gel electrophoresis or liquid chromatography. Subsequently, the proteins are cleaved with sequence-specific endoproteases, most notably trypsin. Following digestion, the generated peptides are investigated by determination of molecular masses or generation of peptide fragments. For protein identification, the experimentally obtained masses are compared with the theoretical peptide masses of proteins stored in databases by means of mass search programs (figure 26) (Aebersold and Mann, 2003; Thiede *et al.*, 2005).

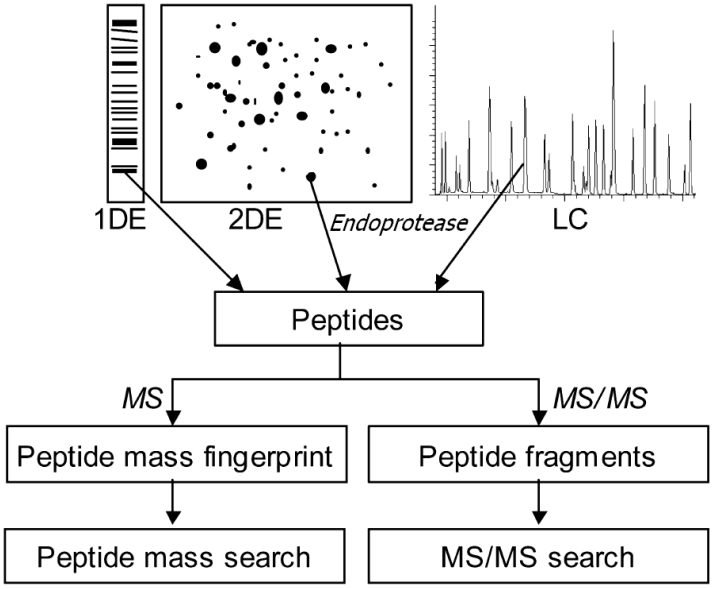


Figure 26. Common procedures to identify proteins by MS (figure from Thiede *et al.*, 2005)

Within the context of the workflows presented in figure 27, the bottom-up approach is represented by the direct mass analysis of the 'entity' subjected to ionization. The ability to extract information about the protein from the peptide masses derives from the specificity of the proteolysis approach and the quality of the mass measurement. The fingerprinting approach is most useful when it is applied to relatively pure proteins (Reid and McLuckey, 2002). Another powerful approach for protein identification is called shotgun proteomics, which is based in the initial digestion of the protein or protein mixtures (bottom-up-based approach) followed by liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis (figure 27). It is important to be noted that shotgun proteomics do not requires the use of pure proteins, compared as bottom-up approach.

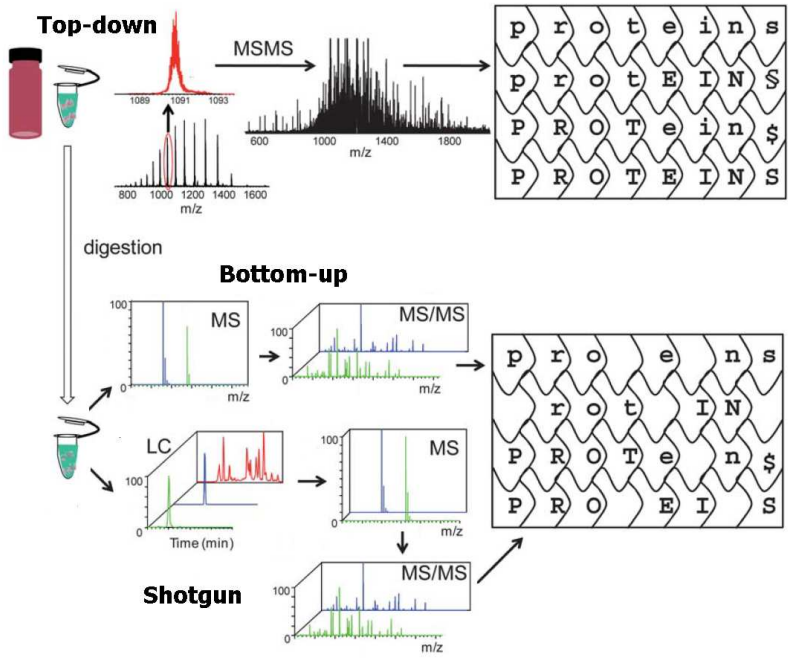


Figure 27. Comparison of top-down, bottom-up and shotgun proteomics (figure adapted from Cui *et al.*, 2011)

I.4.3.5. Top-down protein identification and characterization

Top down approaches for protein identification and characterization have recently been developed that enable primary structural information to be obtained directly from the gas-phase dissociation of intact protein ions without prior recourse to extensive separation or digestion (Kelleher *et* al., 1999). In this scenario, ions derived from the intact protein (the 'entity', figure 27) are fragmented in the mass spectrometer and identification is made by the 'sequence tag' strategy (Mortz *et al.*, 1996; Cargile *et al.*, 2001; Demirev *et al.*, 2001), via database searching of the uninterpreted product ion spectrum (Meng *et al.*, 2001; Reid *et* al., 2002) or through '*de novo*' determination of the complete amino acid sequence (Horn *et al.*, 2000a; Horn *et al.*, 2000b).

A major advantage of top-down methods is that performing an MS/MS experiment on an intact protein ion in principle makes the entire sequence available for examination, better enabling complete characterization of the protein and any associated post-translational modifications (Reid *et* al., 2001; Kelleher *et al.*, 1999; Fridriksson *et al.*, 2000; Shi *et al.*, 2001). Additionally, the large number of redundant identifications that are usually required with the bottom-up peptide sequencing strategies can be avoided. Furthermore, while electrospray ionization produces multiply charged ions with a relatively narrow *m/z* range, the masses of intact proteins are spread over a wider range compared with those of enzymatically derived peptides, potentially simplifying the task of analyzing proteins in complex mixtures (Reid and McLuckey, 2002).

I.5. Expression of heterologous proteins in *Pichia pastoris* system

Pichia pastoris is a methylotrophic yeast that can be genetically engineered to express proteins for both basic research and industrial use (Higgins and Cregg, 1998). The production of a functional protein is intimately related to the cellular machinery of the organism producing the protein. The yeast *P. pastoris* is a useful system for the expression of milligram-to-gram quantities of proteins for both basic laboratory research and industrial manufacture. The fermentation can be readily scaled up to meet greater demands, and parameters influencing protein productivity and activity, such as pH, aeration and carbon source feed rate, can be controlled (Higgins and Cregg, 1998). Compared with mammalian cells, *P. pastoris* does not require a complex growth medium or culture conditions, is genetically relatively easy to manipulate, and has a eukaryotic protein synthesis pathway (Macauley *et al.*, 2005).

There are many reasons for the popularity of the *P. pastoris* expression system, but two are most compelling. The first is an unusually efficient and tightly regulated promoter from the alcohol oxidase I gene (*AOX1*) that is used to drive the expression of the foreign gene (Cregg and Madden, 1988). The *AOX1* promoter is strongly repressed in cells grown on glucose and most other carbon sources, but is induced over 1000-fold when cells are shifted to a medium containing methanol as a sole carbon source. The ability to repress expression of the foreign protein is advantageous if the protein is toxic to the cell (as many recombinant proteins are when synthesized at high levels). The second reason why this is such a popular expression system is that, physiologically, *P. pastoris* prefers a respiratory rather than a fermentative mode of growth. Fermentation products include ethanol and acetic acid, which quickly reach toxic levels in the high cell density environment of a fermenter with strongly fermentative organisms (Cereghino *et al.*, 2002).

The expression of any foreign gene in *P. pastoris* comprises three principal steps: (a) insertion of the gene into an expression vector; (b) introduction of the expression vector into the *P. pastoris* host; and (c) examination of potential strains for the expression of the foreign gene (Macauley *et al.*, 2005).

The *AOX* promoters have therefore been the most widely employed; however, other promoter options are available for the production of foreign proteins in *P. pastoris* (Cereghino and Cregg, 2000). There are two genes that encode alcohol oxidase in *P. pastoris*: *AOX1* and *AOX2*; *AOX1* is responsible for a vast majority of alcohol oxidase activity in the cell (Ellis *et al.*, 1985; Tschopp *et al.*, 1987; Cregg *et al.*, 1989). The *AOX1* promoter has been the most widely reported and utilized of all the available promoters for *P. pastoris* (Cereghino *et al.*, 2001).

In fact, the abundance of the *AOX* enzyme can reach 30% of total cellular protein (TCP) content when grown on methanol as a sole carbon source (Gellissen G., 2000; Cregg *et al.*, 1993). This strong *AOX1* promoter can therefore be used to drive the expression of recombinant proteins to high levels even with a single integrated copy of the expression cassette (Cregg *et al.*, 1993; Clare *et al.*, 1991). Other benefits of this promoter are that it can be switched off, as non-limiting amounts of carbon sources such as glycerol and glucose repress the AOX1 promoter at the transcriptional level and minimize the possibility of selecting non-expressing mutants/contaminants during biomass generation (Cregg *et al.*, 1993; Clare *et al.*, 1991; Cereghino *et al.*, 1999; Katakura *et al.*, 1998). The *AOX2* gene also produces alcohol oxidase, although this gene yields 10–20 times less *AOX* activity than the *AOX1* gene (Cregg and Madden, 1988).

Most *P. pastoris* host strains grow on methanol at the wild-type rate (Mut^+, methanol utilization plus phenotype). However, two other types of host strains are available which vary with regard to their ability to utilize methanol because of deletions in one or both *AOX* genes. Strains with AOX mutations are sometimes better producers of foreign proteins than wild-type strains (Tschopp *et al.*, 1987; Cregg *et al.*, 1987; Chiruvolu *et al.*, 1997). Additionally, these strains do not require the large amounts of methanol routinely used for large-scale fermentations of Mut^+ strains. KM71 (*his4 arg4 aox1v: : SARG4*) is a strain where *AOX1* has been partially deleted and replaced with the *S. cerevisiae ARG4* gene (Cregg and Madden, 1989). Since the strain must rely on the weaker *AOX2* for methanol metabolism, it grows slowly on this carbon source (Mut^s, methanol utilization slow phenotype). Another strain, MC100-3 (*his4 arg4 aox1v: :SARG4 aox2v: : Phis4*), is deleted for both *AOX* genes and is totally unable to grow on methanol (Mut^-, methanol utilization minus phenotype) (Cregg and Madden, 1989). All of these strains, even the Mut^- strain, retain the ability to induce expression at high levels from the *AOX1* promoter (Chiruvolu *et al.*, 1997).

Different expression vectors for *P. pastoris* contain signal sequences to direct the heterologous protein towards the secretion pathway. This feature, along with the low concentration of native proteins of *P. pastoris* in the culture supernatant, significantly simplifies the purification process. In order to obtain a functional system is necessary to generate in frame fusion of the coding sequence of the protein of interest with the secretion signal. The sequences used in *P. pastoris* are α and SUC2 factor, both of *S. cerevisiae* (Cereghino and Cregg, 2000; Beldarrain *et al.*, 2000).

A number of selectable marker genes are known for the molecular genetic manipulation of *P. pastoris*, such as *HIS4* (histidinol dehydrogenase gene), *ARG4* (argininosuccinate lyase gene), *ZeoR* (zeocin resistance gene), Blasticidin S deaminase gene (Cereghino *et al.*, 2001), *ADE1*–PR amidoimidazolesuccinocarboximide Synthase (Cereghino *et al.*, 2001), *URA3*–orotidine 5_-phosphate decarboxylase and *SorR* –acetyl-CoA carboxylase (Wan *et al.*, 2004; Cereghino and Cregg, 2000). The genetic manipulation of *P. pastoris* for the production of various heterologous proteins is simplified by the use of a wide range of selectable markers and promoters. Choosing the correct markers and promoters is essential for obtaining a high productivity from this system and, as has been described here, is different for every heterologous protein. This demonstrates that a great deal of optimization at the molecular level may be required, as well as optimization of the protein production process (Macauley *et al.*, 2005).

It is often desirable to select for transformants containing multiple integration events as such clones potentially express significantly higher levels of the recombinant protein. A further advantage of selecting for multi-copy transformants is that if there is a mutation in one particular copy of the expression cassette, arising from the integration process, then the protein that results from this mutant copy may not contribute as significantly to the total amount of protein expressed. Multiple integration events occur relatively infrequently at a rate of 1–10%. The number of integrated copies of the expression cassette can affect the amount of protein expressed (Cereghino and Cregg, 2001).

Importantly, even at high gene dosages there is no evidence for saturation of the secretory pathway with each integration event found to contribute equally to the levels of protein expressed. For example, the selection effects for hyper-resistance to 100, 500, 1000 and 2000 mg/mL of zeocin were examined with the transformants analysed for copy number. It was found that transformants resistant to 100 mg/mL of zeocin generally contained one copy, those resistant to 500 mg/mL had two copies, those resistance to 1000 mg/mL had three copies and transformants resistant to 2000 mg/mL had four copies. In contrast, other studies have found that transformants resistant to 1000 mg/mL zeocin had copy numbers of 15–25 (Cereghino and Cregg, 2001).

Despite all the advantages mentioned above, the *P. pastoris* system has a number of technological limitations. One of the most important is the high proteolytic activity in the extracellular medium during the fermentation process, which affects the yield and quality of produced proteins. However, different strategies have been used to decrease the proteolytic activity observed in cultures of *P. pastoris*, which include: a) control of operational conditions, such as the pH of the fermentation (Clare et al. 1991a; Werte et al., 1999), reducing the culture temperature to values of 15-25°C (Li *et al*., 2001; Jahic *et al*., 2003), the addition of organic nitrogen sources (Clare *et al*. 1991; Brankamp *et al*., 1995), b) obtaining protease deficient strains such as the development of mutant protease a (encoded by the *PEP4* gene), which in turn activates other vacuolar proteases (Gleeson and Howard, 1994). The use of protease deficient strains, combined with culture strategies mentioned above, has allowed the expression of different proteins (Brierley et al, 1994; Brankamp *et al*., 1995, Jonsson *et al*., 1997). Another limitation that hinders the expression processes of heterologous proteins is the low oxygen transfer capacity in high density cultures, which often reaches a high cell concentration. Finally, the low growth rate during the stage of expression using methanol as a carbon source hinders the establishment of continuous culture processes, usually more productive than batch or fed-batch processes (Cereghino and Cregg, 2000).

Fermentation protocols for *P. pastoris* generally include three separate phases. First is the glycerol batch phase (GBP), in which cells are initially grown on glycerol in a batch mode. In the second phase, the glycerol fed-batch phase (GFP), a limited glycerol feed is initiated following exhaustion of the glycerol, and cell mass is increased to a desired level prior to induction. Furthermore, the *AOX1* promoter is derepressed during this phase due to the absence of excess glycerol. The third phase is the methanol fed-batch phase (MFP), in which methanol is fed at a limited feed rate or maintained at some level to induce the *AOX1* promoter for protein expression. A limited glycerol feed can be simultaneously performed for promoting production when necessary (Zhang *et al*., 2000).

Over 400 proteins, from human endostatin to spider dragline silk protein, have been produced in *P. pastoris* (Cereghino and Cregg, 2001). This expression system is uniquely suited for foreign protein expression for three key reasons: it can be easily manipulated at the molecular genetic level (e.g. gene targeting, high-frequency DNA transformation, cloning by functional complementation); it can express proteins at high levels, intracellularly or extracellularly; and it can perform many 'higher eukaryotic' protein modifications, such as glycosylation, disulfide-bond formation, and proteolytic processing (Cregg *et al*., 2000).

II. OBJECTIVES

II.1. General objective

- Identification of metallocarboxypeptidase, aspartic, cysteine and serine protease inhibitors in marine invertebrates, using enzyme kinetic and proteomic-like techniques. Purification and characterization of a carboxypeptidase inhibitor from the marine snail *Nerita versicolor*.

II.2. Specific objectives

- Design and optimization of a strategy that combines enzyme kinetic and proteomic methods for identifying CPA, CPB, pepsin, papain, trypsin and subtilisin inhibitory activities in extracts of marine invertebrates.
- Design and implementation of purification procedures for a metallocarboxypeptidase inhibitor from the marine mollusc *Nerita versicolor* (NvCI).
- Production of NvCI by recombinant methods in *Pichia pastoris* system.
- Achievement the functional and structural characterization of natural and recombinant NvCI.

III. MATERIALS AND METHODS

III.1. Biological extracts

A set of thirty marine invertebrates were collected at the North Coast of Cuba. The selection and capture control of the species were monitored in collaboration with the Institute of Oceanology of Havana, to ensure the identification, preservation, rational exploitation of marine biodiversity and prevent damage to biological species.

Among the marine species, were prioritized those belonging to the following Phyla: *Annelida, Bryozoa, Chordata, Cnidaria, Echinodermata, Mollusca,* and *Porifera*.

III.2. Extract preparation

Species were transported alive on ice, dried on paper filter, weighted, cut in small pieces with scissors (without any organ dissection) and homogenized in distilled water (2 ml/g) at 4 ºC in a blender (3× 10 s); maximum time between collection and homogenization: 2 h. Homogenates were centrifuged at 10 000 × *g* for 30 min at 4 ºC. The supernatants were filtered on glass fibers (crude extract).

Extracts were lyophilized and kept at −20 ºC until assayed for CPA, CPB, pepsin, papain, trypsin and subtilisin inhibitory activity by resuspending 200 mg (dry weight) in 5ml distilled water (Pascual *et al.*, 2004; González *et al.*, 2007a, Alonso del Rivero *et al.*, 2012.).

III.3. Determination of total protein concentration

The samples were analyzed for total protein in triplicate using a colorimetric detection with bicinchoninic acid (BCA) (Smith *et al.*, 1985). Protocol was followed according to manufacturer's instructions, using 96-well plates. Bovine serum albumin was used as the standard and the respective buffer as the blank. The absorbance was determined in an automatic ELISA reader (Labsystems; EMS Reader MF) at a wavelength of 540 nm.

III.4. Measurement of inhibitory activities

All assays were performed in triplicate at 25 ºC. For 96-well assays an iEMS reader/dispenser FM (Labsystems, Finland) was used in a final reaction volume of 250 µl. For cuvette assays a Cary 400 Bio (Varian Inc, USA) spectrophotometer was used in a final reaction volume of 1 ml. The reactions were followed at 5 sec interval for 5 min, and measured in terms of initial velocities. Inhibitory activity was measured by following the change in absorbance due to the catalytic hydrolysis of the substrate.

• **Extract assay:** Mixtures of activity buffer, biological extract and enzyme were placed in a cuvette or in a 96-well microplate and preincubated at 25ºC for 10 min. A fixed volume of substrate was then added to initiate the reaction.

• **Control assay:** Mixtures of activity buffer and enzyme were placed in cuvette or in a 96-well microplate and preincubated at 25ºC for 10 min. A fixed volume of substrate was then added to initiate the reaction.

III.4.1. CPA inhibitory activity

- **Wavelength (λ):** 340 nm
- **Classification:** 96-well assay

- **Activity buffer:** 20 mM Tris-HCl, pH 7.5, containing 500 mM NaCl, 1% v/v DMSO and 0.05% w/v Brij®-35

- **Enzyme:**
  - CPA from bovine pancreas (E.C.3.4.17.1) supplied by Sigma-Aldrich (USA)
  - **MW:** 38400 Da
  - **ε_{278}:** 65.8 $(\mu mol/ml)^{-1}$ cm^{-1}
  - **Concentration in the assay (M):** $7.0x10^{-9}$
  - **Stock solution:** 22 mg/ml
  - **Working solution:** 1/20000 dilution from stock solution in activity buffer
  - **Volume of enzyme in the assay:** 55 µl

- **Substrate:**
  - N-(4-Methoxyphenylazoformyl)-Phenylalanine-OH potassium salt (AAFP) [CAS Nº 396717-86-5] supplied by Bachem (Switzerland)
  - **MW:** 365.43 Da
  - **ε_{350}:** 19 $(\mu mol/ml)^{-1}$ cm^{-1}
  - **K_M (M):** $1.1x10^{-4}$ (Mock *et al.*, 1996)
  - **Concentration in the assay (M):** $1.0x10^{-4}$
  - **[Substrate] / K_M:** 0.9
  - **Stock solution:** 100 mM in DMSO
  - **Working solution:** 1 mM in 20 mM Tris-HCl, pH 7.5, 500 mM NaCl
  - **Volume of substrate in the assay:** 25 µl

III.4.2. CPB inhibitory activity

- **Wavelength (λ):** 340 nm
- **Classification:** 96-well assay

- **Activity buffer:** 20 mM Tris-HCl, pH 7.5, containing 100 mM NaCl, 1% v/v DMSO and 0.05% w/v Brij®-35

- **Enzyme:**
  - CPB from porcine pancreas (E.C. 3.4.17.2) supplied by Merck & Co., Inc. (USA)
  - **MW:** 34000 Da
  - **ε_{280}:** 77.1 $(\mu mol/ml)^{-1}$ cm^{-1}
  - **Concentration in the assay (M):** $3.0x10^{-9}$
  - **Stock solution:** 5 mg/ml
  - **Working solution:** 1/5000 dilution from stock solution in activity buffer
  - **Volume of enzyme in the assay:** 26 µl

- **Substrate:**
  - N-(4-Methoxyphenylazoformyl)-Arginine-OH potassium salt (AAFA) [CAS Nº 442158-31-8] supplied by Bachem (Switzerland)
  - **MW:** 372.81 Da
  - **ε_{350}:** 19 $(\mu mol/ml)^{-1}$ cm^{-1}
  - **K_M (M):** $4.4x10^{-5}$ (Mock and Stanford, 2002)
  - **Concentration in the assay (M):** $1.0x10^{-4}$
  - **[Substrate] / K_M:** 2.3
  - **Stock solution:** 100 mM in DMSO
  - **Working solution:** 1 mM in 20 mM Tris-HCl, pH 7.5, 100 mM NaCl
  - **Volume of substrate in the assay:** 25 µl

III.4.3. Pepsin inhibitory activity

- **Wavelength (λ):** 310 nm
- **Classification:** cuvette assay

- **Activity buffer:** 100 mM acetate buffer, pH 4.4

- **Enzyme:**
  - Pepsin from porcine gastric mucosa (E.C.3.4.23.1) supplied by Sigma-Aldrich (USA)
  - **MW:** 35000 Da
  - **ε_{280}:** 56.7 $(\mu mol/ml)^{-1}$ cm^{-1}
  - **Concentration in the assay (M):** $1.0x10^{-8}$
  - **Stock solution:** 10 mg/ml in water
  - **Working solution:** 1/1000 dilution from stock solution in activity buffer
  - **Volume of enzyme in the assay:** 35 µl

- **Substrate:**
  - H-Leu-Ser-p-nitro-Phe-Nle-Ala-Leu-OMe.TFA (LSPNPLAL) [CAS Nº 99764-63-3]
  - **MW:** 835.88 Da
  - **ε_{310}:** 10 $(\mu mol/ml)^{-1}$ cm^{-1}
  - **K_M (M):** $2.3x10^{-4}$ (Martin *et al.*, 1980)
  - **Concentration in the assay (M):** $2.0x10^{-4}$
  - **[Substrate] / K_M:** 0.9
  - **Stock solution:** 100 mM in DMSO
  - **Working solution:** 20 mM in DMSO
  - **Volume of substrate in the assay:** 10 µl

III.4.4. Papain inhibitory activity

- **Wavelength (λ):** 414 nm
- **Classification:** 96-well assay

- **Activity buffer:** 100 mM phosphate buffer, pH 6.5 containing 100 mM KCl, 0,1 mM EDTA, 3 mM Dithioerythritol (DTT) and 0.05% w/v Brij®-35

- **Enzyme:**
  - Papain from *Carica papaya* (E.C.3.4.22.2) supplied by Roche (Switzerland)
  - **MW:** 23000 Da
  - **ε_{280}:** 76.6 $(\mu mol/ml)^{-1}$ cm^{-1}
  - **Concentration in the assay (M):** $4.5x10^{-8}$
  - **Stock solution:** 10 mg/ml
  - **Working solution:** 1/2000 dilution from stock solution in activity buffer
  - **Volume of enzyme in the assay:** 52 µl

- **Substrate:**
  - L-Pyroglutamyl-L-phenylalanyl-L-leucine-*p*-nitroanilide (PFLNA) [CAS Nº 85901-57-1] supplied by Bachem (Switzerland)
  - **MW:** 509.56 Da
  - **ε_{410}:** 8.8 $(\mu mol/ml)^{-1}$ cm^{-1}
  - **K_M (M):** $3.4x10^{-4}$ (Filippova *et al.*, 1984)
  - **Concentration in the assay (M):** $4.0x10^{-4}$
  - **[Substrate] / K_M:** 1,2
  - **Stock solution:** 100 mM in DMSO
  - **Working solution:** 4 mM in DMSO
  - **Volume of substrate in the assay:** 25 µl

III.4.5. Trypsin inhibitory activity

- **Wavelength (λ):** 414 nm
- **Classification:** 96-well assay

- **Activity buffer:** 20 mM Tris-HCl, pH 8.0 containing 20 mM $CaCl_2$, 150 mM NaCl and 0.05% v/v Triton X-100

- **Enzyme:**
  - Trypsin from bovine pancreas (E.C.3.4.21.4) supplied by Sigma-Aldrich (USA)
  - **MW:** 24000 Da
  - **ε_{280}:** 34.6 $(\mu mol/ml)^{-1}$ cm^{-1}
  - **Concentration in the assay (M):** $2.8x10^{-7}$
  - **Stock solution:** 10 mg/ml in 1 mM HCl
  - **Working solution:** 1/300 dilution from stock solution in activity buffer
  - **Volume of enzyme in the assay:** 50 µl

- **Substrate:**
  - Nα-Benzoyl-L-arginine 4-nitroanilide hydrochloride (BAPNA) [CAS Nº 21653-40-7] supplied by Bachem (Switzerland)
  - **MW:** 434.88 Da
  - **ε_{410}:** 8.8 $(\mu mol/ml)^{-1}$ cm^{-1}
  - **K_M (M):** $9.4x10^{-4}$ (Erlanger *et al.*, 1961)
  - **Concentration in the assay (M):** $1.0x10^{-3}$
  - **[Substrate] / K_M:** 1.1
  - **Stock solution:** 100 mM in DMSO
  - **Working solution:** 10 mM in 20 mM Tris-HCl, pH 8.0, 20 mM $CaCl_2$, 150 mM NaCl
  - **Volume of substrate in the assay:** 25 µl

III.4.6. Subtilisin inhibitory activity

- **Wavelength (λ):** 414 nm
- **Classification:** 96-well assay

- **Activity buffer:** 50 mM Tris-HCl, pH 8.5 containing 10% v/v DMSO

- **Enzyme:**
  - Subtilisin from Bacillus licheniformis (E.C.3.4.21.14) supplied by Sigma-Aldrich (USA)
  - **MW:** 27290 Da
  - **ε_{280}:** 23.5 $(\mu mol/ml)^{-1}$ cm^{-1}
  - **Concentration in the assay (M):** $2.0x10^{-7}$
  - **Stock solution:** 10 mg/ml in water
  - **Working solution:** 1/400 dilution from stock solution in activity buffer
  - **Volume of enzyme in the assay:** 55 µl

- **Substrate:**
  - Benzylocarbonyl-glycyl-glycyl-L-Leucine 4-nitroanilide (GGLPNA) [CAS Nº 53046-98-3] supplied by Bachem (Switzerland)
  - **MW:** 499.52 Da
  - **ε_{410}:** 8.8 $(\mu mol/ml)^{-1}$ cm^{-1}
  - **K_M (M):** $8.0x10^{-4}$ (Lyublinskaya *et al.*, 1974)
  - **Concentration in the assay (M):** $4.0x10^{-4}$
  - **[Substrate] / K_M:** 0.5
  - **Stock solution:** 100 mM in DMSO
  - **Working solution:** 4 mM in DMSO
  - **Volume of substrate in the assay:** 25 µl

III.4.7. Determination of specific inhibitory activity

One unit of inhibitory activity (equivalent to v_i) was defined as the amount of inhibitor able to reduce one unit of enzymatic activity (equivalent to v_0), as follow:

$$IA(U/ml) = \left(\left(\frac{\Delta Abs}{\Delta t} \right)_{control} - \left(\frac{\Delta Abs}{\Delta t} \right)_{extract} \right) \frac{1}{\xi} \frac{V_{total}}{V_{extract}} \frac{F}{L} \quad \textbf{(1)}$$

Where ***IA*** is the inhibitory activity (U/ml), ***(ΔAbs/Δt)*** $_{control}$ is the absorbance variation per min during the control assay, ***(ΔAbs/Δt)*** $_{extract}$ is the absorbance variation per min during the assay in the presence of extract, ξ is the extinction coefficient for the substrate ($(\mu mol/ml)^{-1}.cm^{-1}$), ***L*** is the path length (cm), V_{total} is the total assay volume (ml), ***V*** $_{extract}$ is the volume of extract in the assay (ml) and ***F*** is the dilution factor for the extract.

The term residual enzymatic activity, also known as v_i/v_o=**a**, was referred to the enzymatic activity obtained after preincubation of the enzyme with the extract containing the inhibitor.

The specific activities were obtained dividing the inhibitory activity by the protein concentration, both obtained on the same sample.

III.4.8. Dose-response relationship and determination of IC_{50} value

Dose-response relationship was obtained by measuring the inhibitory activity (determined in triplicate) at different extract concentrations.

The IC_{50} value was calculated by fitting the experimental data to the following equation using the GraphPad prisma 5 software (GraphPad Software, Inc., USA):

$$\frac{v_i}{v_0} = \frac{1}{1 + \left(\frac{[I]}{IC_{50}} \right)} \quad \textbf{(2)}$$

Where V_i and V_0 are the velocities of the enzyme catalyzed reaction in the presence and absence of inhibitor, respectively, [I] is the concentration of inhibitor in the assay, and the IC_{50} is that concentration of inhibitor required to reduce the reaction velocity to one half that observed in the absence of inhibitor (Copeland R.A., 1995)

III.5. Electrophoretic methods

III.5.1. Tris-glycine SDS-PAGE

This technique described by Laemmli, U.K., 1970 is the most commonly used in SDS-PAGE (sodium dodecyl sulphate polyacrylamide gel electrophoresis) analyses. Electrophoresis was performed with a discontinuous buffer system (4.0% polyacrylamide stacking gel and 12.5% resolving gel) in a Mini-PROTEAN Tetra Cell (Bio-Rad Laboratories, USA). Tris-Glycine-SDS buffer (0.025 M Tris base, 1.92 M Glycine and 1% w/v SDS) was used as anode and cathode buffer. The BenchMark™ Protein Ladder was used as a molecular weight standard. It consists of 15 recombinant proteins ranging in molecular weight from 10 to 220 kDa (Invitrogen Corp., USA). Gels were run with constant current at 15 mA during stacking and then increased to 30 mA in resolving stage for each gel using a PowerPac HC power supply (Bio-Rad Laboratories, USA). Gels were then stained with Instant*Blue*™ (Expedeon, UK).

After destaining, gels were stored in distilled water until they were scanned. Stain intensities of individual protein bands were quantified using a Bio-Rad **620** scanning densitometer (Bio-Rad Laboratories). Data from the densitometer were interpreted using Quantity One® software (version **4.6.3;** Bio-Rad Laboratories, USA)

III.5.2. Tris-tricine-SDS-PAGE

Tris-tricine-SDS-PAGE is commonly used to separate proteins in the mass range 1-100 kDa. It is the preferred electrophoretic system for the resolution of protein smaller than 30 kDa (Schägger, H., 2006). Gels were composed of three sections: a stacking gel (4%T, 3%C), a spacer gel (10%T, 3%C) and a separating gel (16.5%T, 6%C), where %T and %C are defined as follows:

$$\%T = \frac{(acrylamide + bisacrylamide)(g)}{totalvolume\,(ml)} x100 \qquad \textbf{(3)}$$

$$\%C = \frac{bisacrylamide\,(g)}{(acrylamide + bisacrylamide)(g)} x100 \qquad \textbf{(4)}$$

1 M Tris-HCl pH 8.45, 0.1% w/v SDS was used as gel buffer; 0.1 M Tris-HCl pH 8.9 was employed as anode tampon and 0.1 M Tris base, 0.1 M Tricine, pH ~8.25, 0.3% w/v SDS was used as cathode buffer. Furthermore, 3% w/v SDS, 6% v/v b-mercaptoethanol, 30% w/v glycerol, 0.05% w/v Coomassie blue G-250, 150 mM Tris-HCl pH 7.0 was used as reducing sample buffer 4X. Samples were mixed with SDS-containing sample buffer and incubated at 100ºC for 5 min.

Electrophoresis started with an initial voltage of 30 V and it was maintained at this value until the sample entered completely into the stacking gel. At this point, voltage was progressively increased every 10 min, until proteins entered into the spacer gel at 100 V. This value is kept constant until the end of electrophoresis. Gels may warm up, but the temperature should not exceed 35-40ºC (Schägger, H., 2006). Gels were then stained with Instant*Blue*™ (Expedeon, UK). Gel scanning was performed according to the method described above.

III.5.3. Agarose gel electrophoresis

Agarose gels were employed for analysis of samples containing DNA. Two horizontal electrophoresis units were used: Bluemarine™ 100 for gel formar 7x10 cm and Bluemarine™ 200 for gel formats 15x15 cm or 15x20 cm (SERVA electrophoresis, Germany). A power supply EPS 200 (GE healthcare, USA) was used in both cases.
In order to load the DNA sample into the gel a one-tenth volume of 10 X sample buffer [50X TAE buffer (40 mM Tris-acetate buffer containing 1 mM EDTA), glycerol and Orange G] was added to each. Analysis of double restriction enzyme digestion of pPICZαA and pMA-T was performed using gels with 1% w/v and 2% w/v agarose, respectively.

Agarose gels were prepared in TAE buffer containing a final concentration of ethidium bromide at 0.2 μg / ml. Electrophoresis was carried out at constant voltage to maintain a ratio of 10 mV per cm of gel using TAE running buffer. The DNA bands were visualized in an UV light transilluminator (UVP, USA). Estimation of DNA molecular weights was performed by comparison of known molecular weight standards: DNA MW Marker X and III (Roche, Switzerland)

III.6. Immobilization of target proteases

III.6.1. Enzyme immobilization on glyoxal-Sepharose® supports

The immobilization process of proteases on glyoxal-sepharose® support was carried out as described by Guisan *et al.*, 1998. For this purpose, a first stage of support preparation (activation and oxidation) was performed, followed by an enzyme immobilization stage that included a final step of enzyme stabilization by multipoint covalent attachment to the support. The detailed procedure for enzyme immobilization was followed according to Montes *et al.*, 2006:

III.6.1.1. Preparation of glyoxal-Sepharose® support

All reagents used were purchased from Sigma-Aldrich (USA).

III.6.1.1.1. Activation of sepharose gel to glyceryl-Sepharose®

105 g (150 ml) of Sepharose® CL-4B were washed thoroughly with distilled water. The gel was resuspended in distilled water up to a total volume of 180 ml. Then, 50 ml of 1.7 N NaOH solution containing 3.4 g of sodium borohydride ($NaBH_4$) were added to this suspension. The vessel was placed in ice bath; while gently stirring, 36 ml of glycidol were added dropwise. The suspension was gently stirred for 18 h at 4°C. Finally, the suspension was filtered and the support was thoroughly washed with distilled water.

III.6.1.1.2. Oxidation of glyceryl-Sepharose® to Glyoxal-Sepharose®

105 g of glyceryl-Sepharose® (previously obtained) were resuspended in 1500 ml of distilled water. Subsequently, 60 ml of support oxidation solution (0.1 M $NaIO_4$ in water) were slowly added to this suspension while stirring. This amount of oxidation solution enables an activation degree of 40 μmol aldehyde / ml gel. The suspension was gently stirred in the dark for 2 h. Finally, the support was washed with an excess of distilled water and filtered to dryness.

III.6.1.1.3. Determination of activation degree of the support

Oxidation of glycols with sodium periodate is a stoichiometric reaction. Therefore, the activation degree of the support can be easily controlled through the periodate concentration used.

Dilutions of a stock solution of 100 mM $NaIO_4$ were prepared as follows: (1/500, 1/2000, 1/4000, 1/5000, 1/10000). 1 ml of each dilution was placed in separate test tubes. Then, 1 ml of 10% w/v KI and 1 ml of saturated $NaHCO_3$ solution were added in each tube. Subsequently, the absorbance at 390 nm (OD_{390}) of the reaction mixture was measured against a blank (1 ml of distilled water instead of the $NaIO_4$ dilution). A plot of OD_{390} *vs.* [$NaIO_4$] was carried out and a simple linear regression was performed.

$$OD_{390} = m\left[NaIO_4\right] + b \qquad \textbf{(5)}$$

Where **m:** slope and **b:** Y-intercept

Sample Analysis: 1 ml of supernatant (filtered and washed) of the activation mixture was added instead of $NaIO_4$ dilution. The absorbance at 390 nm was measured and its corresponding $NaIO_4$ content was determined by the above equation.

The activation degree of the support (AD) was calculated by the following equations:

$$\mu mol\ NaIO_{4\ consumed} = \mu mol\ of\ carbonyl\ groups\ formed = \mu mol\ NaIO_{4\ initial} - \mu mol\ NaIO_{4\ final} \quad \textbf{(6)}$$

$$AD = \frac{\mu mol\ NaIO_{4\ consumed}}{V\ \mathrm{sup}\ port\left(ml\right)} \quad \textbf{(7)}$$

Where,
μmol $NaIO_{4\ initial}$, periodate measured in the solution at t=0 h
μmol $NaIO_{4\ final}$, periodate measured in the supernatant after filtrating and washing steps

III.6.1.2. Enzyme Immobilization

10 g of glyoxal-Sepharose® were incubated with 100 ml of enzyme solution prepared in immobilization buffer (50 mM sodium bicarbonate-NaOH, pH 10.5). Here, it is necessary not to use Tris or other chemical reagents with primary amino groups $–NH_2$, (this is true for protein immobilization on glyoxal-Sepharose® as well as on N- hydroxysuccinimide activated Sepharose™ supports). At once, 1 ml of distilled water was added to 9 ml of enzyme solution to be used as reference solution. If the enzyme activity decreases during the course of immobilization resulting from enzyme inactivation, this effect must be distinguished from loss of the supernatant resulting from immobilization. The mixture was gently stirred by end-over-end rotation at 4°C. Aliquots of supernatant and suspension were withdrawn at regular time intervals to assay enzyme activity. Supernatant was achieved by centrifugation of the suspension. An assay of enzymatic activity was also performed in the reference solution using the same time intervals and aliquot volumes as in the earlier step. The immobilization process finishes when activity of the supernatant is zero or remains constant. Then, the immobilized preparation was washed five times with three volumes of immobilization buffer.

III.6.1.3. Structural Stabilization by Multipoint Covalent Attachment

The enzyme derivative (previously obtained) was resuspended in 50 ml of immobilization buffer. Then, 50 mg of solid $NaBH_4$ were added to the suspension and stirred for 30 min at 4°C. Finally, the enzyme derivative was washed with distilled water, while vacuum filtering to eliminate the $NaBH_4$. Enzyme derivatives were stored at 4°C in 20% v/v ethanol in water.

III.6.2. Immobilization of pepsin on NHS activated Sepharose™ 4 Fast Flow support

In order to immobilize the porcine pepsin on the activated sepharose support, the following solutions were prepared:

- Matrix of N- hydroxysuccinimide (NHS) activated Sepharose™ 4 Fast Flow (GE Healthcare, USA)
- Solution of pepsin in 20 mM Hepes, pH 7.4 or in sodium phosphate buffer at pH 6. The enzyme concentration depends on the amount of immobilized derivative to be prepared.
- 0.2 M ethanolamine (blocking agent free NH2 groups) in the same buffer in which the protein was dissolved.
- 100 mM acetate buffer at pH 4.4
- Sodium phosphate buffer at pH 6.0
- 1 mM HCl

For the immobilization of pepsin on the activated Sepharose support, the protocol described by Ramírez *et al.*, 2009 was followed:

A fixed amount of matrix (0.8 g equivalent to approximately 1 ml) is washed several times with 1 mM HCl. The pepsin solution was mixed with the matrix and the suspension was allowed to stir 6 hours at 4ºC. Then, the suspension was filtered and the matrix was collected. One volume of ethanolamine solution was added to the matrix (equivalent to the matrix volume). The pH was adjusted and the suspension was stirred during 2 hours. Then, the suspension was filtered and the matrix was collected. Subsequently, the matrix was washed alternately with the solution of pH 4.4 and pH 6.0. The first washing filtrate was collected and stored. A final washing was carried out with a solution at a pH 6.0 and the immobilized derivative was stored at 4ºC in 20% v/v aqueous ethanol.

The filtrates are used to assess the immobilization yield. Since the volume and protein concentration are known, the total amount of protein applied was calculated. The final quantity of immobilized enzymes on the matrices were calculated by differences between the total amounts of protein applied and the total protein measured in all filtrates. This is an indirect method, but it is useful. The direct method measures the enzymatic activity into the matrix.

The enzymatic activities in the immobilized derivatives were determined using a fixed substrate concentration as described in the paragraphs III.4.1 to III.4.6, varying the final concentration of immobilized enzyme in a final reaction volume of 3 ml and 25ºC. Other experimental conditions are summarized in the next table:

Table 2. Experimental conditions for the determination of enzymatic activities in the immobilized derivatives of proteases

Enzyme	Immobilized derivative	Buffer	$[S_0]$ (M)	$[E_0]$ range (M)
Bovine CPA	CPA-1	20 mM Tris-HCl pH 7.5, 500 mM NaCl	$1.0x10^{-4}$	$3.6x10^{-8} – 2.2x10^{-7}$
	CPA-2			$3.7x10^{-8} – 2.2x10^{-7}$
	CPA-3			$3.8x10^{-8} – 2.3x10^{-7}$
Porcine CPB	CPB-1	20 mM Tris-HCl pH 7.5, 100 mM NaCl	$1.0x10^{-4}$	$1.0x10^{-8} – 1.0x10^{-7}$
	CPB-2			$1.6x10^{-8} – 9.7x10^{-8}$
	CPB-3			$1.6x10^{-8} – 9.9x10^{-8}$
Papain	PAP-1	100 mM phosphate buffer pH 6.5, 100 mM KCl, 0.1 mM EDTA, 3 mM DTT	$4.0x10^{-4}$	$1.8x10^{-7} – 9.1x10^{-7}$
	PAP-2			$2.0x10^{-7} – 9.8x10^{-7}$
	PAP-3			$2.3x10^{-7} – 1.1x10^{-6}$
Bovine trypsin	TRYP-1	20 mM Tris-HCl pH 8.0, 20 mM $CaCl_2$, 150 mM NaCl	$1.0x10^{-3}$	$9.3x10^{-8} – 4.6x10^{-7}$
	TRYP-2			$1.4x10^{-7} – 6.9x10^{-7}$
	TRYP-3			$1.4x10^{-7} – 7.0x10^{-7}$
Subtilisin	SUBT-1	50 mM Tris-HCl pH 8.5, 10 %v/v DMSO	$4.0x10^{-4}$	$4.4x10^{-8} – 2.2x10^{-7}$
	SUBT-2			$6.4x10^{-8} – 3.2x10^{-7}$
	SUBT-3			$6.2x10^{-8} – 3.1x10^{-7}$
Porcine pepsin	PEPS-1	100 mM acetate buffer pH 4.4	$2.0x10^{-4}$	$1.0x10^{-7}$
	PEPS-2			$2.2x10^{-7}$
	PEPS-3			$2.0x10^{-7}$

The three concentration ranges used for each enzyme, representing different loads are labeled with 1, 2 and 3

III.6.3. Analysis of the immobilization process by tris-glycine SDS-PAGE

Samples of supernatant and immobilized enzyme were withdrawn at different times. Supernatant samples were submitted to Tris-glycine SDS-PAGE as described above. Non-reduced glyoxal protein preparations were washed with 50 mM sodium bicarbonate buffer at pH 10.5. This derivative was resuspended in 1 M ethanolamine and 2% w/v SDS / 1 M 2-mercaptoethanol. This treatment releases any protein bound to the support (via weak Schiff's bases) (Guisán *et al.*, 2006). The supernatants obtained were also analyzed by Tris-glycine SDS-PAGE.

III.6.4. Determination of immobilization parameters

The immobilization parameters were calculated by the following equations:

$$DI_{prot} = \frac{mg\ prot_{difference}}{V_{\sup port}} \quad \textbf{(8)} \qquad \%I\ prot = \frac{mg\ prot_{difference}}{mg\ prot_{initial}} * 100 \quad \textbf{(9)}$$

$$DI_{EA} = \frac{TU_{difference}}{V_{\sup port}} \quad \textbf{(10)} \qquad \%I\ EA = \frac{TU_{difference}}{TU_{initial}} * 100 \quad \textbf{(11)}$$

$$\%I\ RFA = \frac{TU_{direct}}{TU_{initial}} * 100 \quad \textbf{(12)}$$

Where,
DI prot, degree of enzyme immobilization in terms of protein
DI EA, degree of enzyme immobilization in terms of enzymatic activity
%I prot, percentage of enzyme immobilization in terms of protein
%I EA, percentage of enzyme immobilization in terms of enzymatic activity
%RFA, percentage of retention of functional activity
$TU_{initial}$, Total units of enzymatic activity of the initial solution (t=0) in the immobilization process
$TU_{difference}$, difference between the initial total units and total units of the filtrate and subsequent washings, after the immobilization has been completed
TU_{direct}, Total units of enzymatic activity determined in the immobilized derivative
mg $prot_{initial}$, milligrams of protein of the initial solution (t=0) in the immobilization process
mg $prot_{difference}$, difference between the initial milligrams of protein and milligrams of protein in the filtrate and subsequent washings, after the immobilization has been completed
$V_{support}$, volume of the support used in the immobilization

III.6.5. Experimental determination of external mass transfer limitations in protease-Sepharose® derivatives

In order to evaluate the effect of external mass transfer limitations on the immobilized enzyme derivatives, enzymatic activity was determined using four stirring conditions: 0 rpm (no stirring), 500 rpm, 750 rpm and 1000 rpm (maximum stirring). The rest of the experimental conditions for the enzymatic assay are described in table 2. If the observed reaction rate does not change as the stirring is increased, external mass transfer is not rate-limiting. Furthermore, in stirred-reaction systems, as in this case, external mass transfer limitations can be eliminated using a good stirring system (Guin and Xiaofeng, 1997).

III.6.6. Estimation of internal mass transfer limitations in protease-Sepharose® derivatives

The strategy described by Suau *et al.*, 2009 has been employed in this work for the evaluation of internal mass transfer limitations in the immobilized derivatives on activated Sepharose® supports. In order to evaluate the effect of internal mass transfer limitations, the effectiveness factor for each derivative was determined. For this purpose, the Weisz-Prater modulus (ϕ_{WP}) (Weisz and Prater, 1954), also known as observable Thiele modulus (Φ) was calculated according to the following equation:

$$\phi_{WP} = \left(\frac{R}{3}\right)^2 \frac{\gamma_{obs} \cdot \rho_c}{D_{eff} \cdot C_S} \qquad \textbf{(13)}$$

Where,

γ_{obs}, observed reaction rate (mmol/h.g biocatalyst)
ρ_c, biocatalyst density (g/m^3 biocatalyst)
D_{eff}, effective diffusivity (m^3 solution/h.m biocatalyst) (h=hour)
C_S, substrate concentration (mmol/m^3 solution)
R, paticle radius (m biocatalyst). For the support employed, the mean radius was 45 μm.

The dimensionaless Weisz-Prater modulus is convenient to use because observed reaction rate is measured direcly, C_S can be assumed to be equal to the bulk substrate concentration if all external mass transfer limitations are removed by proper mixing (Vannice, M., 2005). Furthermore, as particle radius is known (45 μM, manufacturer's data) and the catalyst density can be determined gravimetrically, the effective diffusivity can be calculated.

The effective diffusivity describes the steady state diffusion through the pore space in natural porous media (Grathwohl, P., 1958). This coefficient is defined as:

$$D_{eff} = \frac{D_{AB} \cdot \varepsilon_p \cdot \sigma}{\tau} \qquad \textbf{(14)}$$

Where,

D_{AB}, diffusion coefficient of solute A in solvent B for a dilute solution (cm^2 /s)
ε_p, porosity
τ, tortuosity (a measure of the pore structure, Missen *et al.*, 1999)
σ, constriction factor (parameter which accounts for the variation in the cross-sectional area that is normal to diffusion, Satterfield and Sherwood, 1963)

Tortuosity is considered as the inverse of the porosity ($\tau \sim 1/\varepsilon$). Thus, for sepharose ε=0.8 (from the above equation 1/0.8=1.25) and constriction factor $\sigma \sim 1$ described for Sepharose by Suau *et al.*, 2009.

The diffusivity of substrates in water was calculated from the Wilke and Chang equation (Wilke and Chang, 1955) for non-electrolytes in dilute aqueous solutions:

$$D_{AB} = 7.4x10^{-8} \frac{(\phi_B M_B)^{1/2} T}{\mu_B V_A^{0.6}} \qquad \textbf{(15)}$$

ϕ_B, association parameter for solvent (ϕ_{water}=2.6)
M_B, molecular weight of the solvent (M_{Bwater}=18.05 g/gmol)
μ_B, viscosity of the solvent, cp (μ_{water}=0.89 cp at 298 K)
T, system temperature, K (298 K)
V_M, molar volume of the solute A (substrate) at normal boiling point (cm^3/gmol)

The molar volume value was obtained by applying LeBas additive method, where the atomic contributions are added. The following table summarizes the atomic and molecular volumes according to Le Bas, G. 1915:

Table 3. Atomic and molecular volumes

Atom	Atomic volume (cm^3/gmol)	Molecule	Molecular volume (cm^3/gmol)
Carbon	14.8	H_2	14.3
Hydrogen (in bonds)	3.7	O_2	25.6
Chlorine (R-CHCl-R)	24.6	N_2	31.2
Bromine	27.0	Air	29.9
Iodine	37.0	CO	30.7
Sulfur	25.6	CO_2	34.0
Nitrogen (double bond)	15.6	SO_2	44.8
Nitrogen in primary amines	10.5	NO	23.6
Nitrogen in secondary amines	12.0	N_2O	36.4
Oxygen	7.4	NH_3	25.8
Oxygen in methyl esters, aldehydes, ketones	9.1	H_2O	18.9
Oxygen in acids	11.0	H_2S	32.9
Oxygen in methyl ethers	12.0	COS	51.5
Oxygen in higher ethers	9.9	Cl_2	48.4
Oxygen in higher esters	11.0	Br_2	53.2
Benzene ring	15.0	I_2	71.5
Naphthalene ring	30.0		

Criteria for finding internal mass transfer limitations. Once known the observable Thiele modulus, the criteria of the Weisz-Prater modulus was applied in order to identify whether the internal mass transfer limitations are significant and the effectiveness factor was interpolated into the plot ϕ_{WP} vs. η presented by Asenjo and Merchuck, 1995, using the figure representing the first-order reaction for a spherical geometry (figure 28):

Table 4. Criteria for finding internal mass transfer limitations

ϕ_{WP}	η	Internal mass transfer
<0.3	≅1.0	Insignificant
>0.3	<1.0	Significant

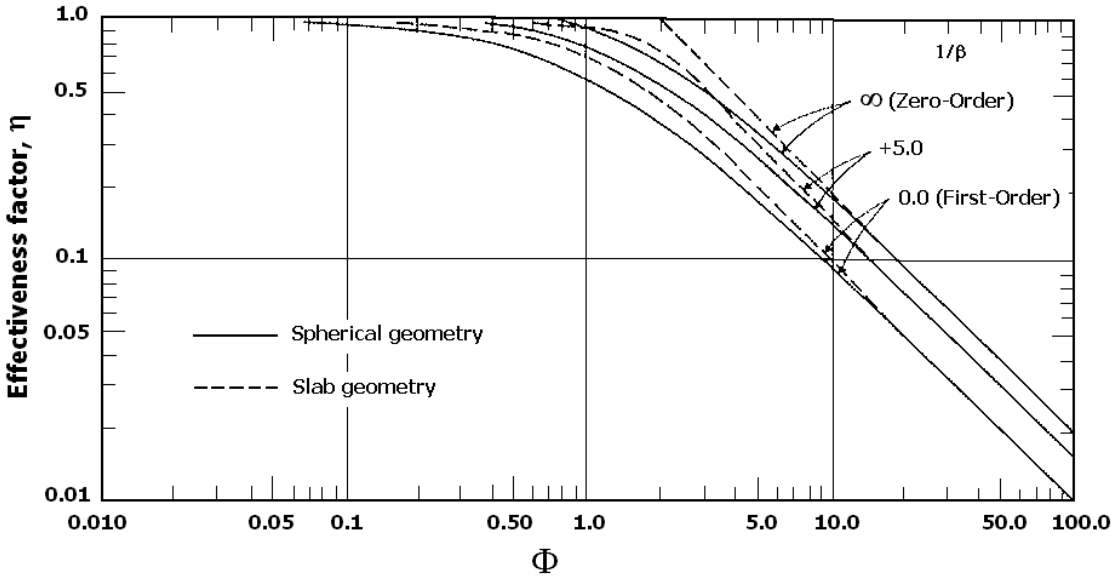


Figure 28. Effectiveness factors for substrate utilization with Michaelis-Menten kinetics. Φ is the observable Thiele modulus

III.7. Identification and analysis of protease inhibitors by proteomic methods

III.7.1. Desalting of extracts using ZipTip®C_4 tips

Due to the sample complexity, in terms of high contents of salts and other interfering components for MALDI analysis, biological extracts were desalted and concentrated using ZipTip®C_4 pipette tips (Millipore, USA) according to the manufacturer's protocol:

- **Resin activation:** Depress pippetor plunger to a dead stop. Using the maximum volume setting of 10 µl, aspirate pure acetonitrile into the tip. Dispense to waste.
- **Equilibration:** Aspirate 0.1% v/v aqueous trifluoroacetic acid (TFA). Dispense to waste and repeat 2 more times.
- **Interaction:** Bind peptides and/or proteins to ZipTip®C_4 pipette tip by fully depressing the pipette plunger to a dead stop. Aspirate and dispense the sample (dissolved in 0.5% aqueous TFA) 15-20 cycles for maximum binding of complex mixtures.
- **Washing:** Aspirate 0.1% aqueous TFA into tip and dispense to waste. Repeat 9 more times.
- **Elution:** Dispense 5 µl of 0.1 % v/V TFA in ACN into a clean vial using a standard pipette tip. Carefully, aspirate and dispense eluant through ZipTip®C_4 pipette tip at least three times without introducing air into the sample.

III.7.2. Intensity Fading MALDI-TOF MS

The most promising extracts in terms of specific inhibitory activity (the highest values), dose-response relationships and IC_{50} (the lowest values) were analyzed by IF MALDI-TOF MS using an indirect protocol (Yanes *et al.*, 2005), that includes an affinity step using the following immobilized proteases:

- **CPA-Glyoxal Sepharose® CL-4B:** 3.5 mg CPA / ml drained gel
- **CPB-Glyoxal Sepharose® CL-4B:** 2.5 mg CPB / ml drained gel
- **Pepsin-NHS-activated Sepharose™ 4 fast flow:** 3.9 mg pepsin / ml drained gel
- **Papain-Glyoxal Sepharose® CL-4B:** 2.7 mg papain / ml drained gel
- **Trypsin-Glyoxal Sepharose® CL-4B:** 5.1 mg trypsin / ml drained gel
- **Subtilisin-Glyoxal Sepharose® CL-4B:** 2.5 mg subtilisin / ml drained gel

Immobilized proteases were prepared according to the immobilization methods described in section III.6. All IF MALDI-TOF MS assays were carried out in duplicate at room temperature.

Firstly, the crude extract was submitted to MALDI-TOF MS analysis in order to optimize its MALDI mass spectrum (TOTAL). Then, the immobilized protease was mixed with the crude extract and the flow-through fraction (NB: Non-biding molecules) was analyzed by MALDI-TOF MS in order to check which molecules of the biological extract were bound to the protease. After several washing steps with the interaction buffer (WS-1 to WS-5), complexes between the immobilized protease and binding molecules from the extract were dissociated by acid treatment of the Sheparose support in the elution step (EL). Those ion signals presenting a fading or disappearance of its relative signal intensity after interaction with the immobilized enzyme and final reappearance in the elution fraction were considered as a positive result (Yanes *et al.*, 2007).

In order to confirm the specificity of interaction between putative binding molecules of the extract and the target protease, an IF MALDI-TOF MS assay using Sepharose CL-4B instead of the immobilized enzyme was performed. Then, the elution fraction of this experiment was compared with the elution fraction obtained in the IF MALDI-TOF MS assay with the immobilized protease and those common ion signals in both mass spectra are attributed to non-specific interactions of proteins in the extract and the Sepharose support.

An additional IF MALDI-TOF MS assay called control of immobilized enzyme was carried out. In this experiment, activity buffer rather than extract was placed in contact with the immobilized enzyme. This experiment was carried out in order to detect peptides from protease hydrolysis resulting in the acidification step of the resin, which may lead to false positives.

The elution fractions of IF MALDI-TOF MS experiments were tested for their respective protease inhibitory activity in order to corroborate the presence of proteinaceous inhibitors in the extract, through a technique of molecular recognition.

A detailed protocol is described as follows:

Column preparation:
- Remove top and bottom cap of an empty spin column supplied by ABT (Spain). Place column into a receiver tube.
- Load 200 μl of immobilized enzyme or Sepharose matrix (dissolved 1:1 v:v in 20% v/v ethanol) into the column.
- Centrifuge at 1500 rpm for 2 min. Discard flow-through.
- Add 200 μl of water LC-MS Chromasolv® (Sigma-Aldrich, USA).
- Centrifuge at 1500 rpm for 2 min. Discard flow-through and repeat the previous step one more time.

Equilibration:
- Add 200 μl of activity buffer 1X into the column.
- Centrifuge at 1500 rpm for 2 min. Discard flow-through and repeat 2 more times.

The buffers employed in the IF MALDI-TOF MS assays are described in the next table, as follows:

Table 5. Buffers used in the IF MALDI-TOF MS assays

Immobilized enzyme	Buffer 1X
CPA	20 mM Tris-HCl pH 7.5, 500 mM NaCl
CPB	20 mM Tris-HCl pH 7.5, 100 mM NaCl
Pepsin	100 mM acetate buffer pH 4.4
Papain	20 mM Tris-HCl pH 7.0, 100 mM KCl, 0.1 mM EDTA, 3 mM DTT
Trypsin	20 mM Tris-HCl pH 8.0, 20 mM $CaCl_2$, 150 mM NaCl
Subtilisin	50 mM Tris-HCl pH 8.5

Interaction:
- Secure the bottom cap.
- Load 200 μl of sample (dissolved 1:1 v:v in activity buffer 2X) and secure the top cap.
- Shake at 400 rpm for 10 min at room temperature.
- Remove top and bottom cap. Place column into a new receiver tube.
- Centrifuge at 2000 rpm for 1 min. Pass the sample through the column 2 more times.
- Recover flow-through (non-binding fraction)

Washing:
- Place column into a new receiver tube. Add 200 μl of activity buffer 1X.
- Centrifuge at 1500 rpm for 2 min. Recover flow-through (washing fraction). Repeat the process 4 more times.

Elution:
- Place column into a new receiver tube. Add 50-100 μl of 0.5% v/v aqueous TFA.
- Centrifuge at 1500 rpm for 2 min. Recover the eluate and pass through the column 2 more times.
- Recover flow-through (elution fraction)

- **Optional step:** adjust the pH of the eluate with 1 M Tris-HCl buffer, pH 7.5 (approximately 10 μl of buffer per 200 μl sample). This step is required in order to measure protease inhibitory activity in the elution fraction

Note: In the control of immobilized enzyme, 200 μl of activity buffer 1X instead of sample is placed into the column.

III.7.2.1. MALDI matrices

All the solid matrices used in these analyses were purchased from Sigma-Aldrich (USA). Matrix solutions were prepared as follows:

- **α-Cyano-4-hydroxycinnamic acid (α-CHCA):** 10 mg/ml in 0.1% v/v aqueous TFA, 30% v/v acetonitrile (ACN)
- **Sinapic acid (SA):** 10 mg/ml in 0.1% v/v aqueous TFA, 30 % v/v ACN
- **2,6-Dihydroxyacetophenone (DHAP):** 5 mg/ml in 20 mM aqueous ammonium citrate dibasic, 30% v/v ACN
- **2,5-dihydroxy-benzoic acid (DHB):** 20 mg/ml in 0.1% v/v aqueous TFA, 30 % v/v ACN
- **Mixture of 2,5-dihydroxy-benzoic acid and 2-hydroxy-5-methoxy-benzoic acid (super-DHB):** 20 mg/ml in 0.1% v/v aqueous TFA, 30 % v/v ACN
- **1,5-Diaminonaphthalene (1,5-DAN):** 20 mg/ml in 0.1% v/v aqueous TFA, 50 % v/v ACN

III.7.2.2. Sample preparation for mass spectrometry

All fractions from IF MALDI-TOF MS were desalted using ZipTip®C_4 pipette tips (millipore). Then, 1 μl of five replicates of each desalted sample was deposited on a MTP 384 target plate polished steel T F (Bruker Daltonics), followed by deposition of 1 μl of DHAP as matrix. The mixture was allowed to dry at room temperature (Karas *et al.*, 1990).

III.7.2.3. MALDI-TOF mass spectrometry analysis

MALDI-TOF mass spectra were acquired on an Ultraflextreme™ MALDI-TOF mass spectrometer (Bruker Daltonics, Germany). The instrument is equipped with a nitrogen laser emitting at 337 nm. Spectra were automatically acquired using the AutoExecute™ acquisition control program, setting the following parameter: initial and maximal power laser of 60 and 100%, respectively; range from 1000-20000 m/z in linear mode geometry; extraction delay time set at 250 ns; accelerating voltage operating in positive ion mode of 25 kV; 0.5 Gs/s sampling rate digitizer; time ion selector deflection of mass ions 1000 m/z. To improve the signal-to-noise ratio, 500 laser shot steps were acquired for each mass spectrum until sum up 3000. An external calibration was performed using a standard protein calibration mixture (Protein Calibration Standard I, Bruker Daltonics). All data were reprocessed using the Flex Analysis software version 3.3 (Bruker Daltonics, Germany).

III.7.3. MALDI-Top-Down sequencing

Taking into account that top-down sequencing (TDS) with mass spectrometry is an approach that offers the possibility to sequence intact proteins, the elution fractions obtained from IF MALDI-TOF MS were subsequently analyzed using this strategy according to the characteristics of their MALDI spectra. Thus, fractions containing peaks below 3000 m/z were fragmented by collision-induced dissociation (CID), while those elution fractions containing a single MALDI signal or few components with a major peak were analyzed by in-source decay (ISD) fragmentation. Thus, the elution fraction of IF MALDI-TOF MS with the crude extract of *H. carunculata* on pepsin-NHS activated Sepharose™ 4 fast flow was used for CID fragmentation while the elution fraction of IF MALDI-TOF MS with the crude extract of *S. helianthus* on trypsin-glyoxal Sepharose® CL 4B was subjected to SID fragmentation.

III.7.3.1. MALDI-Top-Down sequencing using CID fragmentation

III.7.3.1.1. Sample Preparation for mass spectrometry

The selected elution fraction was desalted using ZipTip®C_4 pipette tips (millipore). Then, 1 μl of five replicates of each desalted sample was deposited on a MTP 384 target plate polished steel T F (Bruker Daltonics), followed by deposition of 1 μl of α-CHCA as matrix. The mixture was allowed to dry at room temperature (Karas *et al.*, 1990).

III.7.3.1.2. MALDI TOF/TOF analysis

Spectra were obtained in a reflectron mode and selected peptides were subsequently fragmented using the LIFT method (Suckau *et al.*, 2003a). Spectra were automatically acquired using the AutoExecute™ acquisition control program, setting the following parameter:

- ***Spectra acquisition*** – initial and maximal power laser of 50 and 100%, respectively; range from 500-5000 m/z in reflectron mode; extraction delay time set at 90 ns; accelerating voltage operating in positive ion mode of 25 kV; 4 Gs/s sampling rate digitizer; time ion selector deflection of mass ions 1000 m/z. To improve the signal-to-noise ratio, 500 laser shot steps were acquired for each mass spectrum until sum up 3000. An external calibration was applied using a standard peptide calibration mixture (Peptide Calibration Standard, Bruker Daltonics).

- ***CID fragmentation*** – 40% of initial power laser; precursor mass range from 0-3500 m/z; extraction delay time set at 90 ns; accelerating voltage operating in positive ion mode of 19 kV; 2 Gs/s sampling rate digitizer; parent accumulation of 6000 shots; fragment accumulation of 16000 shots. Maximally four MS/MS precursors were selected per MS run.

III.7.3.2. MALDI-Top-Down sequencing using ISD fragmentation

III.7.3.2.1. Sample preparation for mass spectrometry

The selected elution fraction was desalted using ZipTip®C_4 pipette tips (millipore). Here, three different matrices were used in order to analyze the intact protein of interest by ISD fragmentation. For this purpose, 1 μl of five replicates of each desalted sample was premixed with 1 μl of DHB, 1 μl of super-DHB and 2 μl of 1,5-DAN, in each case. The mixture was deposited on a MTP 384 target plate polished steel T F (Bruker Daltonics) and allowed to dry at room temperature (Karas *et al.*, 1990).

III.7.3.2.2. ISD fragmentation

Proteins were fragmented using a reflector ISD method (Suckau and Cornett, 1998). Spectra were automatically acquired using the AutoExecute™ acquisition control program, setting the following parameter: initial and maximal power laser of 60 and 100%, respectively; range from 1000-5000 m/z in reflectron mode; extraction delay time set at 100 ns; accelerating voltage operating in positive ion mode of 25 kV; 4 Gs/s sampling rate digitizer; time ion selector deflection of mass ions 1000 m/z. To improve the signal-to-noise ratio, 500 laser shot steps were acquired for each mass spectrum until sum up 6000. An external calibration was applied using a standard peptide calibration mixture (Peptide Calibration Standard, Bruker Daltonics).

III.7.3.3. Sequencing analysis

Monoisotopic masses were determined using FlexAnalysis 3.3 (Bruker Daltonics) with the SNAP peak picking algorithm (Koester and Holle, 1999). All fragmentation spectra were analyzed using the Biotools software (version 3.2, Bruker Daltonics). Automated *de novo* sequence analysis was carried out using RapidDeNovo™ module for MALDI-LIFT spectra.

In the case of ISD spectra, a simple *de novo* sequencing strategy was employed, just assigning contiguous sequence tags using the 20 standard amino acid residue masses with a mass tolerance of 0.5 Da (Suckau and Resemann, 2003b).

Homology searches were performed for both the resulting sequences from RapiDeNovo analysis and tag sequences obtained from ISD spectra using MS-BLAST (http://dove.embl-heidelberg.de/Blast2/msblast.html) and databases were searched with Mascot (Matrix Science, London) on a local intranet server.

III.8. Purification procedure of a carboxypeptidase inhibitor from the marine extract of *Nerita versicolor* (NvCI)

Taking into account the strength of inhibition detected in the crude extract of *N. versicolor* against CPA (high specific inhibitory activity and low IC_{50} value), the results obtained in the IF MALDI-TOF MS experiment, in which was detected a major signal around 6 kDa (indicating the presence of putative proteinaceous CPA inhibitors) as well as the bioavailability of this species in nature, the marine extract of *N. versicolor* was selected in order to purify and characterize the molecule(s) responsible for CPA inhibition.

III.8.1. Clarification of the crude extract by heat treatment

Considering the thermal stability displayed by certain protease inhibitors isolated from molluscs (Tschesche and Dietl, 1976, González *et al.*, 2007a) and that heat precipitation is a fast and simple clarification method for removing labile protein contaminants (Chavez *et al.* 1990), heat treatment was performed with the aim of clarifying the crude extract. The crude extract was heated at 60ºC for 30 min, then cooled to room temperature and centrifuged at 10000 g at 4ºC for 20 min. The supernatant (heat-treated extract) was analyzed for total protein using the BCA method (section III.3). In addition, the heat-treated extract was tested for CPA inhibitory activity as described in section III.4.1, in order to determine its specific inhibitory activity, establish the dose-response curve and calculate the IC_{50} value (sections III.4.7. and III.4.8.). The heat-treated extract was stored at -20°C for later use.

III.8.2. Affinity chromatography

All procedures were carried out at room temperature. The heat-treated extract was applied to a CPA-glyoxal Sepharose® CL-4B column (20 ml bed volume at 3.5 mg CPA / ml drained gel). The column was prepared by coupling bovine CPA to Glyoxal Sepharose® CL-4B according to the method earlier described (section III.6.1.). The column was equilibrated at 15 cm/h with 20 mM Tris-HCl buffer, pH 7.5, containing 500 mM NaCl. Then, 20 ml of extract (dissolved 1/1 v/v in equilibration buffer) were applied to the chromatographic column (1.6 cm x 9.9 cm) at 15 cm/h.

Ten-ml fractions of non-retained proteins were collected. The low affinity proteins were removed by washing the column at 30 cm/h with 20 mM Tris-HCl buffer, pH 7.5, containing 500 mM NaCl (10 column volumes) and 10 ml fractions were collected. Proteins were eluted with 6 column volumes using 10 mM NaOH, pH 12 at 30 cm/h (Alonso del Rivero *et al.*, 2012) and 3 ml fractions were collected. The separation was followed by absorbance at 280 nm and by measuring total protein by BCA method and CPA inhibitory activity in each fraction as well as using MALDI-TOF MS analysis.

Fractions containing the peak of inhibitory activity were pooled, dialyzed against 1 L of water using CelluSep H1 tubular dialysis membranes (molecular weight cut-off, MWCO: 1 kDa) (Membrane filtration products Inc., USA) and finally lyophilized in a Telstar lyophilizer (Telstar, USA). This lyophilized protein pool was stored at -20°C for later use.

III.8.3. Reversed-phase HPLC

The peak of CPA inhibitory activity from affinity chromatography was applied to RP-HPLC, using a C_4 (3,9 mm x 150 mm) column (Waters Corp., USA). Solution A was 0.1% v/v TFA, while solution B was acetonitrile (ACN) in 0.1% v/v TFA. Elution was performed under the following experimental conditions: At the start, 10% of solution B during 15 min was used, followed by a linear gradient from 10 to 40% over 100 min and a linear gradient from 40 to 98% over 1 min. The flow rate was 0.5 ml/min and room temperature.

All collected fractions were analyzed by MALDI-TOF MS and those containing the molecules of interest were dried in vacuum at 35 ºC (Savant SPD111V SpeedVac concentrator, USA). Then, these RP-HPLC peaks were resuspended in water LC-MS Chromasolv® and in each of them were measured absorbance at 280 nm and its CPA inhibitory activity.

Specific inhibitory activity, yield percentage and fold purification were determined from the values of CPA inhibitory activity and protein concentration in each purification step (Janson and Ryden, 1998). The percent yield was determined by the ratio of total inhibitory activity value obtained in each purification step and that obtained in the initial step (100%). The determination of fold purification was performed by calculating the ratio of specific inhibitory activity in each purification step with respect to the initial step, which was considered to have fold purification equal to one.

III.9. Molecular characterization of natural NvCI

III.9.1. Molecular mass determination of NvCI by MALDI-TOF MS

MALDI-TOF mass spectrum of purified NvCI was acquired according to the method described in section III.7.2.3 using DHAP as a matrix.

III.9.2. Primary structure determination of NvCI

III.9.2.1. Determination of the N-terminus amino acid sequence

The N-terminus sequence of NvCI was determined by automated Edman degradation (Edman P., 1949; Edman and Begg, 1967) using a Procise® 492 protein sequencing system (Applied biosystems, USA) in the Proteomics and Bioinformatics facility from UAB, a member of Proteored network.

III.9.2.2. Determination of the number of cysteine residues

600 µl of NvCI (A_{280nm}=0.344) were dissolved in 20 mM ammonium bicarbonate pH 8, containing 8 M urea as denaturant, and 30 mM DTT as reductant. Denaturation and reduction were carried out at 58°C for 1 h. S-carbamidomethylation of cysteine residues was performed with 250 mM iodoacetamide pH 8, at room temperature in the dark for 1 h. In each of these stages, 10 µl of sample were taken, desalted using ZipTip®C_4 pipette tips and analyzed by MALDI-TOF MS. The number of cysteine residues was determined, taking into account the molecular mass difference between native and S-carbamidomethylated inhibitor.

III.9.2.3. Amino acid sequencing of the enzymatic digestion products of NvCI

200 μl of reduced and S-carbamidomethylated inhibitor from the above procedure were diluted 5 times in activity buffer in order to achieve 1 M urea. To obtain internal peptides of NvCI, the following enzymatic digestions were performed:

- **Lysyl endopeptidase from *Lysobacter enzymogenes* (Lys-C, E.C. 3.4.21.50):** activity buffer 5x: 125 mM Tris-HCl pH 8.5, 5 mM EDTA; 14 μl of Lys-C (Roche, Switzerland) at 0.1 μg/μl dissolved in water LC-MS Chromasolv®; incubation time of 16 h at 37ºC.

- **Glutamyl endopeptidase from *Staphylococcus aureus* (Glu-C, E.C. 3.4.21.19):** activity buffer 5X: 500 mM ammonium bicarbonate pH 7.8; 1 μl of Glu-C (Roche, Switzerland) at 1 μg/μl dissolved in water LC-MS Chromasolv®; incubation time of 16 h at room temperature.

- **Modified trypsin from bovine pancreas (E.C. 3.4.21.4):** activity buffer 5X: 500 mM Tris-HCl pH 8.5; 7 μl of trypsin (Roche, Switzerland) at 0.2 μg/μl dissolved in 1 mM HCl; incubation time of 16 h at 37ºC

As the enzymatic digestion was completed, the mixture was acidified with 0.1% v/v aqueous TFA and 50 μl of sample were taken, concentrated in vacuum at 35ºC (Savant SPD111V SpeedVac concentrator, USA), desalted using ZipTip®C_4 pipette tips and finally analyzed by MALDI-TOF MS.

The digestion products were separated by RP-HPLC, using a C_4 column (3,9 mm x 150 mm). Solution A was 0.1% v/v TFA in water, while solution B was 0.1% v/v TFA in acetonitrile. Elution was performed under the following experimental conditions: At the start, 5% of solution B during 10 min was used, followed by a linear gradient from 5 to 55% over 120 min and a linear gradient from 55 to 98% over 5 min. The flow rate was 0.75 ml/min and room temperature.

RP-HPLC fractions containing the purified digestion product were dried under vacuum and resuspended in water LC-MS Chromasolv®. In order to sequence these peptides the following approaches were used:

- ***De novo* peptide sequencing by MALDI MS/MS:** 1 μl of sample was deposited on a MTP 384 target plate polished steel T F (Bruker Daltonics), followed by deposition of 1 μl of α-cyano-4-hydroxycinnamic acid (HCCA) as a matrix. The mixture was allowed to dry at room temperature.

Peptide fragmentation by low energy collision-induced dissociation (CID) was used for tandem mass spectrometry. *De novo* peptide sequencing analysis was carried out using the Biotools software (version 3.2, Burker Daltonics) with the RapiDeNovo™ module.

- **Automated Edman degradation:** it was performed according to the method described above in the determination of the N-terminal amino acid sequence of the entire molecule.

III.9.2.4. Synthesis and cloning of cDNA encoding NvCI

Along with the strategies previously used for NVCI amino acid sequencing, the synthesis of the complementary DNA (cDNA) encoding for this molecule was carried out using the RT (reverse transcription)-PCR (polymerase chain reaction) techniques. Subsequently, this cDNA was cloned and sequenced.

III.9.2.4.1. Total RNA isolation

Initially, the extraction of total RNA was carried out. For this reason, one specimen of the marine snail *Nerita versicolor* was collected and quickly divided into two sections: muscle and body. RNA extraction of each section was performed using a total RNA extraction kit, NucleoSpin® RNA L (Macherey-Nagel, Germany). The tissue was grinded in a mortar to a fine powder in the presence of liquid nitrogen, followed by homogenization with a syringe needle 25GX5/8 inch in the lysis buffer. The purification of RNA was performed following the instructions in the kit of reagents used.

III.9.2.4.2. Primers design

Design of semi-degenerate primers in the synthesis of cDNA encoding NvCI was performed from the N-terminal amino acid sequence, obtained by Edman degradation and *de novo* sequencing by MALDI-TOF MS. Four primers were designed from the amino acid sequence of NvCI: P1_NvCI$_{1\text{-}8}$: 5'-TTYCAYGTSCCNGAYGAYCGNCC-3', P2_NvCI$_{1\text{-}8}$: 5'-TTYCAYGTSCCNGAYGAYAGRCC-3', P3_NvCI$_{1\text{-}8}$: 5'-TTYCAYGTWCCNGAYGAYCGNCC-3' and P4_NvCI$_{1\text{-}8}$: 5'- TTYCAYGTWCCNGAYGAYAGRCC -3'. Where, Y=C/T, R=A/G, S=C/G, W=A/T, N=A/C/G/T.

III.9.2.4.3. 3'-RACE-PCR

The RACE (Rapid amplification of cDNA Ends)-PCR (Frohman *et al.*, 1988) is a technique designed to determine the sequence of the ends of an RNA. This approach generates cDNAs (obtained by PCR) including a copy of the RNA sequence at its 3 'or 5'-end (3'-RACE or 5'-RACE, respectively). In this strategy, a target sequence (nonspecific sequence) is added at the end of the cDNA to be obtained and subsequently, the use of a specific primer allows the amplification a desired gene sequence.

3'-RACE-PCR consists of a first stage of reverse transcription (RT), in which the target sequence attached to a homopolymer poly(dT) is added, followed by a second stage of cDNA amplification by PCR using two primers, the first one corresponding to the gene sequence (specific primer) and another to the target sequence.

First strand cDNA synthesis. Reverse transcription reaction (RT): In order to obtain the first strand cDNA synthesis, the commercial kit "First Strand cDNA Synthesis Kit for RT-PCR, AMV" (Roche, Switzerland) was used. For this stage, a primer containing a target sequence in the 5'-end was designed and attached at the 3'-end by a poly(dT) sequence (R_0R_1poly(dT):5´CCGGAATTCACTGCAGGGTACCCAATACGACTCACTATAGGGCTTTTTTTTTTTTT TTT-3´).

As a previous step to the RT-PCR, RNA sample and R_0R_1poly(dT) primer were subjected to heat treatment (15 min at 65 °C and 5 min in an ice bath) to cause the rupture of the secondary structure and facilitate the correct hybridization of the DNA primer-RNA pair. The following table describes the composition of the reaction mixture:

Table 6. Reaction mixture for RT-PCR

Reagent	Volume (µl)
10X reaction buffer	2
25 mM $MgCl_2$	4
10 mM dNTPs	2
5 µM R_0R_1polydT primer	2
50 U/µl RNase inhibitor	1
AMV Reverse transcriptase*	1
Denatured RNA sample	1.5
Water, PCR grade	6.5
Total volume	20

*Reverse transcriptase from the avian myeloblastosis virus

Reaction was carried out in a thermal cycler XP Cycler, (BIOER TECHNOLOGY CO., China) under the following experimental conditions: 10 min at 25 °C, 90 min at 42 °C, 5 min at 99 °C and 5 min at 10 °C. As a positive control of the RT-PCR, a control Neo pa RNA was employed as a template RNA.

PCR amplification: In order to isolate the specific gene product of the RT-PCR, a PCR step was subsequently performed using the specific primers for NvCI and the R_0 primer: 5´-CCGGAATTCACTGCAGGGT-3´. The above reaction products were used as a template cDNA in the PCR. The following table describes the composition of the reaction mixture:

Table 7. Reaction mixture for PCR

Reagent	Volume (μl)
10X reaction buffer	5
25 mM $MgCl_2$	2
10 mM dNTPs	2
5 μM R_0 primer	2.5
5 μM NvCI specific primer	2.5
DMSO	3
1 U/μl Pfu DNA polymerase*	1.25
Template DNA	5
Water, PCR grade	26.75
Total volume	50

*DNA polymerase from the hyperthermophilic archaeon *Pyrococcus furiosus*

Reaction was carried out using the thermal cycler above mentioned using the following program: 1x (5 min at 95 °C), 35x (00:30 min at 95 °C, 00:30 min at 60 °C; 01:30 min at 72 °C), 1x (15 min at 72 °C) and 1x (16 h at 4 °C). As a negative control of the RT-PCR, an aliquot of RNA sample from *N. versicolor* was used as template in the PCR. In the positive control for RT-PCR, the reaction products above obtained with the control Neo pa RNA were used as a template DNA using the following upstream primer: 5´- GCTCTGATGCCGCCGTGT-3´ and downstream primer: 5´-CCCTGATGCTCTTCGTCC-3´) and the program above detailed for PCR amplification. PCR products were analyzed by 2% w/v agarose gel electrophoresis. DNA bands from the PCR reaction were cut from the agarose gel under UV light and purified using the DNA extraction kit QIAEX II® (QIAGEN, Germany).

III.9.2.4.4. Ligation of DNA fragments

Selected and purified DNA fragments were ligated to the pBE vector. The ligation was carried out according to the following experimental conditions: 7 μl DNA fragment, 1 μl T4 DNA ligase (New England BioLabs Inc., USA), 1 μl 10X T4 DNA ligase reaction buffer (New England BioLabs Inc., USA) and 1 ·l sterile water, at room temperature for 1 h.

III.9.2.4.5. DNA transformation and clone selection

DNA ligation products were transformed in competent *E.coli* XL1-Blue cells by heat shock according to the protocol described by Sambrook and Russell, 2001: 30 min on ice, 90 s at 42ºC and finally 5 min on ice.

The cell suspension from the transformation mixture was grown in 300 ml of LB medium for 15 min at 37ºC, followed by centrifugation at 2000g for 2 min. Cells were carefully resuspended and added on a Petri dish with LBA medium (LB medium containing 50 μg/ml ampicillin), IPTG (0.1 gmol/l) and X-Gal (20 mg/ml). The cell suspension was homogeneously deposited on the agar plate using a Drigalsky glass spatula. Then, the agar plate was incubated at 37ºC, overnight. White colonies were selected and these were subsequently grown in LBA medium at 37ºC, 200 rpm, overnight.

III.9.2.4.6. Storage and purification of plasmid DNA clones

Prior to DNA purification, an aliquot of 92 μl of each clone was mixed with 400 μl of 80% v/v sterile glycerol and stored at -80ºC. This treatment allowed the preservation of clones by long periods of time. Finally, plasmid DNA from *E. coli* transformation was purified using the extraction kit GenElute™ Plasmid Miniprep (Sigma-Aldrich, USA).

III.9.2.4.7. DNA sequencing and sequence analyses

DNA samples were sequenced in the Servei de Genòmica – Universitat Autònoma de Barcelona (Barcelona, Spain) using the following primers: Fup-17 (M13 forward): 5'-GTAAAACGACGGCCAGT-3' and Rup-17 (reverse primer): 5'-GGAAACAGCTATGACCATG-3'. The chromatograms were analyzed with the DNA baser software (Heracle BioSoft SRL, Romania). DNA sequences obtained from the different clones were submitted to multiple sequence alignment using the ClustalX version 2.0 software (Larkin *et al.*, 2007).

Consensus sequence was translated in six reading frames with the *Translate Tool* software available on the ExPASy server (Swiss Institute of Bioinformatics) to identify the nucleotide sequence encoding NvCI.

III.9.3. Computed parameters of NvCI

All the computed parameters were calculated by using ExPaSy Proteomics Server (Swiss Institut of Bioinformatics) with the ProtParam tool. A detailed description about every parameter and its calculation algorithm was made by Gatteiger *et al.*, 2005 and presented as follows:

Extinction coefficients: The extinction coefficient indicates how much light a protein absorbs at a certain wavelength and a predetermined (constant) concentration and path length. It is a useful parameter to quantify the amount of a pure protein.

It is possible to estimate the molar extinction coefficient of a protein from the knowledge of its amino acid composition (Gill and von Hippel, 1989). From the molar extinction coefficient of tyrosine, tryptophan, and cystine (cysteine does not absorb appreciably at wavelengths >260 nm, while cystine does) at a given wavelength, the extinction coefficient of a native and denatured protein can be computed.

The extinction coefficient was calculated using the equation:

$$\varepsilon_{prot} = Numb(Tyr)x\varepsilon_{Tyr} + Numb(Trp)x\varepsilon_{Trp} + Numb(Cystine)x\varepsilon_{Cystine} \qquad \textbf{(16)}$$

The absorbance of a protein concentration at 0.1% ($\varepsilon_{0.1\%}$), equal to 1 g/l or 1 mg/ml solution was calculated using the following formula:

$$\varepsilon_{0.1\%} = \frac{\varepsilon_{prot}}{Molecular\ weight} \qquad \textbf{(17)}$$

***In vivo* Half-life:** The half-life is a prediction of the time taking for disappearing half of the amount of protein after its synthesis in the cell. The prediction is given for three organisms (human, yeast, and *Escherichia coli*), but it is possible to extrapolate the result to similar organisms.

It has been shown that the identity of the N-terminal residue of a protein plays an important role in determining its stability *in vivo* (Bachmair *et al.*, 1986; Gonda *et al.*, 1989; Tobias *et al.*, 1991). Importance of the N-terminal residues is generally known as the 'N-end rule'. Thus, the N-terminal amino acid sequence determines the half-life of proteins.

Instability index (II): The instability index provides an estimate of the stability of the protein in a test tube. A protein whose instability index is smaller than 40 is predicted as stable, a value above 40 predicts that the protein may be unstable.

Statistical analysis of 12 unstable and 32 stable proteins has revealed that there are certain dipeptides, which are significantly different in the unstable proteins compared with those in the stable ones (Guruprasad *et al.*, 1990). The authors of this method have assigned a weight value of instability to each of the 400 different dipeptides (DIWV). Using these weight values, it is possible to compute an instability index (II), which is defined as:

$$\begin{aligned} & i = L - 1 \\ & II = (10/L) * Sum\, DIWV\left(X[i]X[i+1]\right) \\ & i = 1 \end{aligned} \qquad \textbf{(18)}$$

Where,
L is the length of sequence
DIWV (X[i]X[i+1]) is the instability weight value for the dipeptide starting in position i.

Aliphatic index: The aliphatic index of a protein is defined as the relative volume occupied by aliphatic side chains (alanine, valine, isoleucine, and leucine). It may be considered as a positive factor for the increase of thermostability of globular proteins.

The aliphatic index of a protein is calculated according to the following formula (Ikai, A.J., 1980):

$$Aliphatic\ index = X(Ala) + a * X(Val) + b * \left[X(Ile) + X(Leu)\right] \qquad \textbf{(19)}$$

Where X(Ala), X(Val), X(Ile), and X(Leu) are mole percentages (100 X mole fraction) of alanine, valine, isoleucine, and leucine. The coefficients *a* and *b* are the relative volumes of valine side chain (*a*=2.9) and Leu/Ile side chains (*b*=3.9) with respect to the side chain of alanine.

Grand average of hydropathy: The grand average of hydropathy (GRAVY) value for a peptide or protein is calculated as the sum of hydropathy values of all the amino acids, divided by the number of residues in the sequence. Grand average of hydropathicity index indicates the solubility of the protein: positive GRAVY (hydrophobic), negative GRAVY (hydrophilic) (Kyte and Doolittle, 1982).

III.10. Heterologous expression of NvCI in *P. pastoris*

III.10.1. Construction of pPICZαA-NvCI plasmid

After obtaining the complete amino acid sequence of NvCI by combination of Edman degradation and MALDI-TOF MS, the design and construction of a synthetic gene was performed in order to express this molecule in *P. pastoris* system (Geneart, Germany).

In this design, the DNA sequence of NvCI was fused in frame to the *Saccharomyces cerevisiae* prepro alpha factor signal in *Xho*I site of pPICZαA vector for secretion into the culture medium.

DNA transformation and clone selection: Plasmid DNA stocks of pPICZαA and pMA-T (which contains the DNA encoding for NvCI) were transformed in competent *E.coli* XL1-Blue cells by heat shock according to the protocol described by *Sambrook* and *Russell*, 2001): 30 min on ice, 90 s at 42ºC and finally 5 min on ice.

The cell suspension from the transformation mixture was grown in 300 ml of LB medium for 15 min at 37ºC, followed by centrifugation at 2000g for 2 min. Cells were carefully resuspended and added on a Petri dish with LBA medium (LB medium containing 50 μg/ml ampicillin), IPTG (0.1 gmol/l) and X-Gal (20 mg/ml). The cell suspension was homogeneously deposited on the agar plate using a Drigalsky glass spatula. Then, the agar plate was incubated at 37ºC, overnight. White colonies were selected and these were subsequently grown in LBA medium at 37ºC, 200 rpm, overnight.

Plasmid DNA purification: Plasmid DNA from *E. coli* transformation was purified using the extraction kit GenElute™ Plasmid Miniprep (Sigma-Aldrich, USA)

Double digestion of plasmid DNA with *NotI* and *XhoI* restriction enzymes: Plasmid DNA of pPICZαA and pMA-T was linearized by double digestion with *NotI* and *XhoI* restriction enzymes. Digestions were performed using the following reaction conditions: 4 μl 10X H buffer (500 mM Tris-HCl, 1 M NaCl, 100 mM $MgCl_2$, 10 mM DTT, pH 7.5), 4 μl each restriction enzyme, 2 μl DNA sample, 26 μl sterile water at 37ºC for 1 h. The digestion products were analyzed by agarose gel electrophoresis.

DNA purification from agarose gels: DNA corresponding to linearized pPICZαA and NvCI fragment were cut from the agarose gel under UV light and purified using the DNA extraction kit QIAEX II® (QIAGEN, Germany)

Ligation of DNA fragments: Linearized pPICZαA and NvCI fragment were ligated by employing the following experimental conditions: 4 μl pPICZαA, 2 μl NvCI fragment, 2 μl T4 DNA ligase (Promega, USA), 11 μl 2X rapid ligation buffer (Promega, USA), 3 μl sterile water, at 4ºC for 16 h.

Transformation and clone selection of pPICZαA – NvCI construct: The transformation of ligation products in competent *E.coli* XL1-Blue cells was performed using the method described above. Transformants were grown in a Petri dish with LsLB medium containing 25 μg/ml zeocin™. The agar plate was incubated at 37ºC, overnight. Approximately 10 colonies were selected and these were subsequently grown in LsLB medium containing 25 μg/ml zeocin™ at 30ºC, 200 rpm, overnight.

Storage and purification of pPICZαA – NvCI construct: Prior to DNA purification, an aliquot of 92 μl of each clone was mixed with 400 μl of 80% v/v sterile glycerol and stored at -80ºC. Finally, plasmid DNA from *E. coli* transformation was purified using the extraction kit GenElute™ Plasmid Miniprep (Sigma-Aldrich, USA).

Linearization of pPICZαA – NvCI plasmid with *SacI* restriction enzyme: Plasmid DNA of pPICZαA–NvCI was linearized enzymatic digestion with *SacI* restriction enzyme. Digestion was performed using the following reaction conditions: 5 μl 10X A buffer (330 mM Tris-acetate, 660 mM K-acetate, 100 mM Mg-acetate, 5 mM DTT, pH 7.9), 3 μl *SacI* restriction enzyme, 8 μl DNA sample, 34 μl sterile water at 37ºC for 3 h. The digestion products were analyzed by 1% w/v agarose gel electrophoresis. Subsequently, plasmid DNA was cut and purified from the agarose gel using the method described above.

III.10.2. Transformation of pPICZαA – NvCI plasmid in *P. pastoris*

The linearized plasmid of pPICZαA-NvCI with *Sac* I restriction enzyme was transformed into *P. pastoris* X-33 strain (Mut^+) by electroporation according to the protocol described by Cereghino *et al.*, 2005. Transformants were grown in a Petri dish with YPD medium containing 100 μg/ml zeocin™. The agar plate was incubated at 30ºC, over three days. Afterwards, one clone was selected to grow in YPD medium containing 100 μg/ml zeocin™ at 30ºC, 250 rpm, overnight.

III.10.3. Formation of Zeocin™ hyper-resistant transformants

Selection of Zeocin-resistant transformants at high zeocin concentrations (i.e., Zeocin hyper-resistant transformants) was performed in order to generate enrichment in recombinant strains with multiple copies of the integrated vector. For this purpose, three colonies were grown at 250 rpm and 30ºC, using the following zeocin concentrations (μg/ml): 500, 1000, 2000, 5000 and 10000. Two or the three initial transformants did not grow in YPD medium containing 5000 μg/ml Zeocin™ after 3 days, whereas only one transformant grew at this zeocin concentration, above which not grown was observed. Therefore, the hyper-resistant transformant at 5000 μg/ml Zeocin™ was employed in the bioreactor production of recombinant NvCI.

III.10.4. Recombinant production of NvCI in a plant-pilot-scale agitated bioreactor

III.10.4.1. System preparation

The selected zeocin hyper-resistant transformant was grown in a in a Petri dish with YPD medium containing 500 μg/ml zeocin™ at 30ºC, over 3 days, in order to obtain an isolated colony. Subsequently, one colony from this agar plate was grown overnight in YPD medium without Zeocin™ at 250 rpm and 30ºC. Four Erlenmeyers were filled with 250 ml YPD medium without Zeocin and inoculated with 10 ml of the previous. The cultures were incubated overnight at 30ºC and 250 rpm. Finally, cell cultures were centrifuged at 500 rpm for 5 min and the pellets were resuspended in a total volume of 300 ml using the supernatant (YPD medium) obtained in the previous centrifugation step.

On the other hand, basal salts medium (BSM) for *P. pastoris* high cell density growth was prepared according to Invitrogen Corp., 2000 as follows:

Table 8. Composition of BSM medium for *P. pastoris*

Compound	Concentration
$(NH_4)_2SO_4$	10,0 g/L
H_3PO_4 (85% v/v)	26,7 ml/L
$CaSO_4$	0,47 g/L
K_2SO_4	18,2 g/L
$MgSO_4.7H_2O$	14,9 g/L
KOH	4,13 g/L
Glycerol	50,0 g/L

In addition, a trace minerals solution for *P. pastoris* growth was prepared (table 9) along with a Biotin 500X solution (20 mg/100 ml in 20 mM Tris-HCl pH 8.0). Both solutions were sterile filtered and stored at 4ºC for further use.

Table 9. Composition of trace mineral solution

Compound	Concentration
$CuSO_4$.5H2O	6,00 g/L
KI	0,09 g/L
$MnCl_2$	2,23 g/L
$Na_2MoO_4.2H_2O$	0,20 g/L
H_3BO_3	0,02 g/L
$CoCl_2$.6H2O	0,92 g/L
$ZnSO_4$.7H2O	42,16 g/L
$FeSO_4$.7H2O	65,00 g/L
H_2SO_4	5,00 ml/L

Recombinant production of NvCI was carried out in an autoclavable bioreactor (3-7 l) supplied by Applikon® Biotechnology (Netherlands) and the total fermentation volume was 3 l.

Initially, the pH sensor to be subsequently coupled to the bioreactor was calibrated according to the manufacturer's instructions. Then, the bioreactor was loaded with 2.7 l of BSM medium. Antifoam solution (1/50 v/v aqueous AF-204, Sigma-Aldrich) and inoculum reservoir were sterilized at 121ºC, 20 min.

III.10.4.2. Cell growth in glycerol batch phase

Once stabilized the system properties (T=28ºC, pH=5.0±0.2, 300 rpm, airflow of 1 ml/min) dissolved oxygen (DO) sensor was calibrated at DO=100%. Then, the inoculum was added. At this point, 6 ml of Biotin500X solution and 15 ml of Trace minerals solution were added in sterile conditions to the bioreactor. These solutions were supplied every 24 hours throughout the production process of the protein. The growth in this carbon source was performed in order to generate a desired amount of cell mass prior to protein production (Zhang *et al.*, 2000).

III.10.4.3. Cell growth in methanol fed-batch phase

After approximately 16 hours of growth in glycerol batch phase, a burst increase of DO was observed, as a result of the total consumption of this carbon source. At this point in time, 6 ml of methanol (representing 1.5 g methanol/l culture) were added to the bioreactor. Also, pH value was adjusted to 4 during this growth stage. DO above 20% was maintained whenever possible. When the culture was fully adapted to methanol utilization (approximately after 4 hours), a continuous methanol flow was started. Methanol concentration was maintained during 48 hours of induction time at 2-3 g methanol/l culture (1.96-1.81 V) using a methanol sensor coupled to a DVM830L digital multimeter (Velleman components, Belgium) which was modeled according to the following equation:

$$[Methanol](g/l) = a \cdot \left(1 - e^{-b \cdot V}\right)^c \quad \textbf{(20)}$$

Where,
a=8,2176565x10^{-1}
b=2,49247825
c=-1,1725038546x10^{2}
V=electrical potential difference in volts (V)

III.10.4.4. Monitoring process and yield

Production of the recombinant inhibitor was monitored according to parameters such as wet cell weight (WCW) obtained by centrifugation at 4000 rpm for 15 min, as well as by MALDI-TOF MS, determination of protein concentration in supernatant by BCA method and bCPA inhibitory activity.

III.11. Purification and characterization of recombinant NvCI

III.11.1. Weak cation exchange chromatography

The fermentation supernatant was applied to a weak cation exchange chromatography column Accell™ Plus CM (1.6cmx20cm; Waters Corporation, USA) equilibrated with 20 mM sodium citrate buffer at pH 3.0 using an Äktaprime™ system (GE healthcare, UK). Non-interacting proteins as well as other components from the extract were removed by washing the column with the equilibration buffer for 3 column volumes (CV). Elution was performed using 20 mM Tris-HCl buffer, pH 7.0 for 4 CV, followed by a linear gradient of 20 mM Tris-HCl buffer, pH 7.0, 1 M NaCl for 4 CV, and finally, 100% of the latter buffer for 1 CV. The whole process was carried out at 2 ml/min and room temperature. Elution peak containing bCPA inhibitory activity was analyzed by Tris-Tricine SDS-PAGE and MALDI-TOF MS.

III.11.2. Weak anion exchange chromatography

The active fractions were dialyzed using a CelluSep H1 tubular dialysis membrane (MWCO of 1 kDa) in order to remove salts from this separation step. Sample was equilibrated in 20 mM Tris-HCl pH 8.5 and applied to a weak anion exchange column TSK-GEL® DEAE-5PW (7.5cmx7.5mm; Tosoh Bioscience LLC., Japan) using the Äktaprime™ system mentioned above as a second purification step. Non-retained molecules were removed by washing the column with the equilibration buffer over 5 CV. Elution was performed using a linear gradient from 0% to 100% of 20 mM Tris-HCl pH 8.5, 1 M NaCl for 20 CV, followed by 100% of the latter buffer over 10 CV. The flow rate was 1 ml/min and room temperature.

Functional activity of NvCI was made by measuring bCPA inhibitory activity and purity was evaluated by Tris-tricine SDS-PAGE and MALDI-TOF MS. The N-terminus sequence, number of cysteine residues and tryptic peptide mapping were carried out according to the methods described above for natural NvCI.

III.11.3. Experimental determination of the molar extinction coefficient of rNvCI

The experimental determination of this parameter for rNvCI was performed by measuring the absorbance at 280 nm of inhibitor samples at different protein concentrations, determined by BCA method. A graph representing the absorbance at 280 nm vs. protein concentration was plotted and the molar extinction coefficient was calculated as the slope of the straight line of best fit, according to Lambert-Beer law.

III.12. Analysis of natural and recombinant NvCI by circular dichroism spectroscopy

Circular dichroism (CD) spectra were obtained in a Jasco J-715 spectropolarimeter (Jasco Corporation, Japan) between 25ºC and 90ºC. Far-UV measurements were carried out in cuvettes with 2 mm path length. Initially, the protein (0.05 and 0.06 mg/ml for natural and recombinant NvCI, respectively) was dissolved in 0.1% v/v aqueous TFA pH 2.0, 20 mM sodium phosphate buffer pH 6.5 and 20 mM sodium phosphate buffer pH 8.0, in each case and its CD spectra were acquired at 25ºC. Thermal denaturation analysis was carried out in 20 mM sodium phosphate buffer pH 8.0 at 25, 60, 80 and 90ºC. Reduction and/or denaturation were performed by treating NvCI or rNvCI with 20 mM dithiothreitol (DTT) for 24 h and/or dissolving the protein in 8 M urea.

III.13. Kinetic characterization of natural and recombinant NvCI

III.13.1. Inhibitory specificity study

The specificity of natural and recombinant NvCI (NvCI and rNvCI, respectively) was evaluated by determining the inhibitory activity against proteases from four mechanistic classes:

- **Metalloproteases:** Bovine CPA (bCPA), human CPA-1 (hCPA1, E.C.3.4.17.1), human CPA-2 (hCPA2, E.C.3.4.17.15), human CPA-4 (hCPA4, E.C.3.4.17.1), porcine CPB (pCPB), human CPB (hCPB, E.C.3.4.17.2), bovine thrombin-activable fibrinolysis inhibitor (TAFI) (bTAFI, E.C.4.4.17.20), human TAFI (hTAFI, E.C.3.4.17.20) and CPD domain I from *Drosophila melanogaster* (CPD-I; E.C.3.4.17.22). The inhibitory activity of hCPA1, hCPA2 and hCPA4 was determined using 100 μM AAFP as substrate at 340 nm and bCPA activity buffer; while inhibitory activity of hCPB, bTAFI, hTAFI and CPD-I was evaluated with 100 μM AAFA as substrate at 340 nm and pCPB activity buffer.

hCPB, bTAFI and hTAFI were activated immediately before used as described by Arolas *et al.*, 2005 and Sanglas *et al.*, 2008 as follows:

  - **hCPB:** 15 μg of Pro-hCPB (produced in the enzymology and protein engineering laboratory) and 0.05 μg of trypsin were dissolved in 50 mM Tris-HCl buffer pH 7.5, containing 150 mM NaCl and allowed to incubate at 37ºC for 1 h.
  - **bTAFI:** 15 μg of Pro-bTAFI (produced in the enzymology and protein engineering laboratory) were mixed with 0.075 μg of rabbit thrombomodulin (TM) (American Diagnostica, USA) and 0.015 μg of thrombin (Sigma-Aldrich, USA) using 50 mM Tris-HCl pH 7.5, containing 150 mM NaCl, 10 mM $CaCl_2$ and 0.01% v/v Tween®-20 as activation buffer. The mixture was placed at 37ºC for 30 min.
  - **hTAFI:** 15 μg of Pro-hTAFI (produced in the enzymology and protein engineering laboratory), 1 μg of TM and 0.5 μg of thrombin were dissolved in 50 mM Tris-HCl buffer pH 7.5, containing 150 mM NaCl, 10 mM $CaCl_2$ and 0.01% v/v Tween®-20. The mixture was incubated at room temperature for 30 min.

- **Cysteine proteases:** papain from *C. papaya*
- **Serine proteases:** bovine trypsin and subtilisin from *B. licheniformis*
- **Aspartic proteases:** porcine pepsin.

All assays for MCPs were performed in triplicate, in a final reaction volume of 250 μl at 37ºC, while inhibitory activities for proteases belonging to other mechanistic classes were carried out at 25ºC. Other experimental conditions are described in table 10 and in section III.4.

Table 10. Summary of enzyme and inhibitor concentrations used in the measurement of inhibitory activities in specificity assays

Enzyme (E)	(S_0) (M)	Natural NvCI		Recombinant NvCI	
		[E_0] (nM)	[I_0] (nM)	[E_0] (nM)	[I_0] (nM)
bCPA	$1.0x10^{-4}$	3.5	3.0	2.0	0.7
hCPA1	$1.0x10^{-4}$	2.6	2.5	4.8	0.6
hCPA2	$1.0x10^{-4}$	11.0	16.5	7.4	20.0
hCPA4	$1.0x10^{-4}$	26.2	8.5	22.6	5.0
pCPB	$1.0x10^{-4}$	1.7	10.5	1.9	2.5
hCPB	$1.0x10^{-4}$	8.4	4.5	8.9	2.0
bTAFI	$1.0x10^{-4}$	5.9	3.5	23.0	4.0
hTAFI	$1.0x10^{-4}$	54.2	5.5	31.2	4.0
CPD-I	$1.0x10^{-4}$	16.0	1-160	16.0	1-160
Papain	$4.0x10^{-4}$	45.0	1-500	45.0	1-500
Bovine trypsin	$1.0x10^{-3}$	280.0	1-500	280.0	1-500
Subtilisin	$4.0x10^{-4}$	200.0	1-500	200.0	1-500
Porcine pepsin	$2.0x10^{-4}$	10.0	1-500	10.0	1-500

III.13.2. Active-site titration of proteases and determination of their active concentrations

The strategy for the kinetic study of natural and recombinant NvCI was performed according to the non-Michaelis-Menten kinetics described for tight-biding inhibitors (Bieth, J.G., 1995; Chávez and Gutiérrez, 2012).

Measurement of active-site concentrations of bCPA, hCPA1, hCPA2, hCPA4, pCPB, hCPB, and bTAFI were performed using a tight-binding protease inhibitor such as the Tick Carboxypeptidase Inhibitor (TCI) (Arolas *et al.*, 2005).

Initially, experiments of the effect of time to establish equilibrium on the enzymatic activity were carried out on titration conditions ($[E_0]/K_i \geq 100$). Subsequently, the active-site titrations of proteases were performed by varying the concentration of TCI in each assay using fixed enzyme and substrate concentrations as well as fixed preincubation times (table 11).

Table 11. Summary of enzyme concentrations used in the measurement of inhibitory activities in enzyme titration assays

Enzyme (E)	[E_0] (nM)	[E_0] (nM)
bCPA	25.0	20.0
hCPA1	22.0	26.0
hCPA2	48.0	27.0
hCPA4	45.0	73.5
pCPB	17.5	17.0
hCPB	29.0	NP
bTAFI	95.0	NP
hTAFI	NP	NP

NP: The assay was not performed

Once the equilibrium time was reached, substrate was added in order to initiate the reaction. Then, residual enzymatic activities (v_i/v_0) were plotted against inhibitor concentration, and the initial linear part of the plot was extrapolated to X-intercept at the equivalence point corresponding to the concentration of active enzyme (Bieth,J.G., 1995).

The percentage of active enzyme ($\%E_{active}$) was calculated according to the following equation:

$$\%E_{active} = \frac{[E]_{active}}{[E]_{total}} * 100 \qquad \textbf{(21)}$$

Where $[E]_{active}$ is the concentration of active enzyme and $[E]_{total}$ is the concentration of total enzyme used in the assay.

All assays were performed in triplicate, using 100 μM of substrate in a final reaction volume of 250 μl at 37ºC. Other experimental conditions are described in sections III.4.1, and III.4.2.

III.13.3. Titration of NvCI with standard carboxypeptidases and determination of its active concentration

As a first step, experiments of the effect of time to establish equilibrium on the enzymatic activity were carried out on titration conditions ($[E_0]/K_i \geq 100$). Then, the determination of active concentrations of NvCI and rNvCI was performed by using the active enzyme concentration previously assessed. The experiment was carried out by varying the inhibitor concentration in each assay and using a fixed concentration of active enzyme, substrate and preincubation time. The concentration of active inhibitor was calculated according to the earlier described method, where the X-intercept corresponded to the active inhibitor, assuming a stoichiometry of enzyme-inhibitor interaction 1-to-1.

All assays were performed in triplicate, using 100 μM of AAFP substrate for CPA-like enzymes, and 100 μM AAFA substrate for CPB-like enzymes in a final reaction volume of 250 μl at 37ºC. Other experimental conditions are described in sections III.4.1, and III.4.2.

Table 12. Summary of enzyme concentrations used in the measurement of inhibitory activities in NvCI titration assays

Enzyme (E)	Natural NvCI	Recombinant NvCI
	$[E_0]$ (nM)	$[E_0]$ (nM)
bCPA	25.0	20.0
hCPA1	19.5	26.5
hCPA2	48.0	27.0
hCPA4	45.0	73.0
pCPB	17.5	13.0
hCPB	29.0	NP
bTAFI	95.0	NP
hTAFI	NP	NP

NP: The assay was not performed

III.13.4. Determination of Michaelis-Menten constants (K_M)

The determination of Michaelis-Menten constant (K_M) was performed for all the carboxypeptidases used in this kinetic study with the exception of human TAFI (hTAFI). These experiments were carried out by measuring initial reaction velocities at different substrate concentrations (table 13) using a fixed enzyme concentration $[E_0]$ at 37ºC and the activity buffers described in sections III.4.1 and III.4.2 for CPA and CPB-like enzymes, respectively. A plot of $[v_0]$ vs. $[S_0]$ was performed and their values were adjusted to the Michaelis-Menten equation:

$$v = \frac{v_{max}[S_0]}{K_M + [S_0]} \quad \textbf{(22)}$$

The best-fit values of K_M were obtained by adjusting the experimental values to the Michaelis-Menten equation using the program GraphPad Prism 5 (GraphPad Software, Inc., USA).

Table 13. Summary of enzyme and substrate concentrations used in the determination of K_M values

Enzyme	$[E_0]$ (nM)	*$[S_0]$ range (μM)
Bovine CPA	3.5	0 – 200
Human CPA-1	2.6	0 – 200
Human CPA-2	6.8	0 – 210
Human CPA-4	22.6	0 – 200
Porcine CPB	1.7	0 – 200
Human CPB	8.4	0 – 200
Bovine TAFI	14.0	0 – 200

*100 mM AAFP and 100 mM AAFA substrate were employed for CPA and CPB-like enzymes, respectively

III.13.5. Effect of substrate concentration on residual enzymatic activity

In order to determine the capacity of the substrates to induce dissociation of the EI complex, the effect of substrate concentration on the inhibitory activity of NvCI and rNvCI against bCPA, hCPA1, hCPA2, hCPA4, pCPB, hCPB and bTAFI was carried out. In this experiment, three substrate concentrations were used: 50 μM, 100 μM and 150 μM AAFP for CPA-like enzymes and 50 μM, 100 μM and 150 μM AAFA for CPB-like enzymes.

All assays were performed in triplicate, in a final reaction volume of 250 μl at 37ºC and the activity buffers described in sections III.4.1 and III.4.2 for CPA and CPB-like enzymes, respectively. Other experimental conditions are described in table 14.

Table 14. Summary of enzyme and inhibitor concentrations used during the assays of the effect of substrate concentrations on the inhibitory activity

Enzyme (E)	$[E_0]$ (nM)	$[NvCI_0]$ (nM)	$[E_0]$ (nM)	$[rNvCI_0]$ (nM)
bCPA	3.5	4.0	2.0	1.2
hCPA1	2.6	2.7	4.8	1.2
hCPA2	11.0	25.0	7.4	15.0
hCPA4	26.2	8.5	22.6	7.0
pCPB	1.7	5.5	1.9	1.0
hCPB	8.4	1.5	8.9	1.0
bTAFI	5.9	1.5	23.0	2.5
hTAFI	NP	---	NP	---

NP: The assay was not performed

III.13.6. Determination of inhibition constants (K_i)

Prior to the determination of K_i values, it was necessary to determine the time required to establish equilibrium between NvCI / rNvCI and bCPA, hCPA1, hCPA2, hCPA4, pCPB, hCPB, bTAFI and hTAFI on equilibrium conditions ($[E_0]/Ki \leq 10$). For this purpose, enzyme and inhibitor were preincubated at different times (1, 5, 10, 15, 20, 25 and 30 min) and in each experiment the residual enzymatic activity after addition of substrate was determined (100 μM AAFP/AAFA in all assays).

The K_i values of NvCI and rNvCI against bCPA, hCPA1, hCPA2, hCPA4, pCPB, hCPB, bTAFI and hTAFI were determined by measuring the enzymatic residual activity (**v_i/v_0=a**) at different inhibitor concentrations and using a fixed enzyme concentration. The substrate concentration values were 100 μM of AAFP substrate for CPA-like enzymes and 100 μM AAFA substrate for CPB-like enzymes.

The determination of K_i values of NvCI and rNvCI against bCPA, hCPA1, hCPA2, hCPA4, pCPB, hCPB, bTAFI and hTAFI was carried out on equilibrium conditions ($[E_0]/Ki \leq 10$), using a fixed preincubation time. The K_i values of NvCI and rNvCI against the previously mentioned metallocarboxypeptidases were evaluated by varying the concentration of active inhibitor in each assay and using a fixed concentration of active enzyme. The enzyme and inhibitor concentrations used are summarized in the next table, as follows:

Table 15. Summary of enzyme concentrations used in the determination of K_i values

	K_i for NvCI	**K_i for rNvCI**
Enzyme (E)	**$[E_0]$ (nM)**	**$[E_0]$ (nM)**
bCPA	3.5	2.0
hCPA1	2.6	4.8
hCPA2	11.0	7.4
hCPA4	26.2	22.6
pCPB	1.7	1.9
hCPB	8.4	8.9
bTAFI	5.9	23.0
hTAFI	54.2	31.2

The best estimates of apparent K_i values (K_{iapp}) were obtained by fitting the experimental data to the equation for tight-binding inhibitors (Morrison, 1982) by non-linear regression using the Graphad prisma 5 software (GraphPad Software, Inc., USA):

$$a = 1 - \frac{\left\{\left([E_0]+[I_0]+K_{iapp}\right) - \left([E_0]+[I_0]+K_{iapp}\right)^2 - 4[E_0][I_0]\right\}^{1/2}}{2[E_0]} \quad \textbf{(23)}$$

Where **a**=v_i/v_0 (residual enzymatic activity), **K_{iapp}:** apparent dissociation constant of the enzyme-inhibitor complex, **$[E_0]$:** concentration of active enzyme and **$[I_0]$:** concentration of active inhibitor.

The real K_i values of NvCI and rNvCI against each of these enzymes were calculated according to the equation described by Coopeland, R.A., 1995:

$$K_{iapp} = K_i\left(1 + \frac{[S_0]}{K_M}\right) \quad \textbf{(24)}$$

Where K_{iapp}: apparent dissociation constant of the enzyme-inhibitor complex, **$[S_0]$:** initial substrate concentration in the reaction mixture and **K_M:** Michaelis-Menten constant of the enzyme-substrate complex.

III.14. Heterologous expression of human procarboxypeptidase A4 in *P. pastoris*

Human procarboxypeptidase A4 (PhCPA-4) was over-expressed and secreted to the extracellular medium using the *Pichia pastoris* heterologous system as described by Pallares *et al.*, 2005. The PhCPA4/pPIC9 construct transformed into the *Pichia pastoris* KM71 (mut^S phenotype) cells was kindly provided by Msc. S. Tanco (Institut de Biotecnologia, UAB). Production of PhCPA-4 was performed in an autoclavable bioreactor (3-7 l) supplied by Applikon® Biotechnology (Netherlands) in a total culture volume of 3 l.

Taking into account that at pH=5.0 or even lower values, the zinc ion at the active-site in carboxypeptidases is dissociated with concomitant loss of enzyme activity, it was necessary to carry out the protein expression at pH greater than 5.0. Thus, both the glycerol batch phase and methanol fed-batch phase were performed at pH 6.0. For this purpose, sodium hexametaphosphate was used instead of phosphoric acid in the BSM medium as phosphorus source (Oehler *et al.*, 1998), due to the very low solubility of orthophosphate (HPO_4^{2-}) with Mg^{2+}, Ca^{2+}, and other polyvalent cations presented in trace metals (Curless *et al.*, 1996). When the pH is greater than 5.0, H_3PO_4 forms a heavy precipitate which causes problems such as an unbalanced nutrient supply or nutrient starvation, difficulty in cell density measurement, arduous downstream processing, etc. However, sodium hexametaphosphate does not induce any precipitate below pH=8.5, although this solution must be prepared separately from the BSM medium and filter sterilized, otherwise a heavy precipitate will occur when autoclaved together with other components (Zhang *et al.*, 2000).

The monitoring production of this enzyme was carried out by Tris-glycine SDS-PAGE as well as by determining the wet cell weight (WCW) and protein concentration by BCA method.

III.15. Purification of recombinant human carboxypeptidase A4

III.15.1. Hydrophobic interaction chromatography

The enzyme purification was performed using a combination of hydrophobic interaction chromatography (HIC) and a weak anion-exchange chromatography (Tanco *et al.*, 2010). The supernatant obtained from the bioreactor production was equilibrated with 50 mM Tris-HCl pH 8.0, containing 30% w/v ammonium sulfate and applied to a TOYOPEARL® Butyl-650M column (1.6cmx20cm; Sigma-aldrich, USA), in order to ensure fixation of protein to the matrix through its superficial hydrophobic groups. Retained proteins were eluted using a linear salt gradient from the same equilibration buffer to 0% ammonium sulfate. Identification of PhCPA-4 in the elution peak was performed by Tris-glycine SDS-PAGE and the selected fractions were dialyzed using a cellulose membrane (MWCO of 6-8 kDa; Spectrum laboratories Inc., USA) and equilibrated with 20 mM Tris-HCl pH 9.0. Subsequently, the pro-enzyme form was subjected to an activation process with trypsin: 1:40 w:w trypsin:PhCPA4 at room temperature for 1 h.

III.15.2. Weak anion exchange chromatography

The second purification stage of anion-exchange chromatography enabled the final enzyme purification. TSK-GEL® DEAE-5PW column (7.5cmx7.5mm, Tosoh Bioscience LLC., Japan) was equilibrated with 20 mM Tris-HCl pH 9.0. The elution was performed using a linear gradient from 0% to 100% of 100 mM Tris-HCl pH 8.6, containing 0.4 M ammonium acetate. Finally, the active enzyme was dialyzed and equilibrated in 50 mM Tris-HCl pH 8.5, containing 150 mM NaCl. hCPA4 purity was determined by tris-glycine SDS-PAGE and its functional activity was determined by the hydrolysis of the synthetic substrate N-(4-Methoxyphenylazoformyl)-Phenylalanine (AAFA) (Mock *et al.*, 1996).

III.16. Crystal structure of rNvCI in complex with human carboxypeptidase A4

III.16.1. Formation and purification of hCPA4-rNvCI complex

NvCI complex with hCPA4 was performed by preincubation of both proteins for 30 min in buffer 50 mM Tris-HCl pH 8.5, containing 150 mM NaCl at 37°C. For this purpose, 16 mg of hCPA-4 were incubated with 5.5 mg of rNvCI in a final reaction volume of 70 ml, equivalent to an enzyme: inhibitor molar ratio of 1:2.

The volume of complex solution was reduced to 10 ml by concentration in an Amicon® Ultra-4 centrifugal filter (nominal molecular weight limit of 3 kDa; Millipore, USA). The complex was captured on a molecular exclusion chromatography column (HiLoad Superdex 75 26/60; GE healthcare, UK) equilibrated with the same buffer used for complex formation. Elution peaks corresponding to the complex, free enzyme and inhibitor were analyzed by PAGE using both NvCI and hCPA4 as markers.

In order to perform crystallization assays, the purified hCPA4–NvCI complex was concentrated to 17.6 mg/ml using an Amicon® Ultra-4 centrifugal filter (nominal molecular weight limit of 3 kDa), and at this time a change of size exclusion chromatography buffer to 5 mM Tris-HCl buffer pH 8.5, containing 50 mM NaCl was carried out. Similarly, hCPA4 was also concentrated to 11.8 mg/ml and stored in the same final buffer of the complex.

III.16.2. Crystallization and data collection

In order to obtain protein crystals of NvCI-hCPA4 complex, an initial screening was performed at 18°C by sitting drop vapor diffusion method using a 96-well crystallization plate in a multi-channel PHOENIX™ RE crystallization robot (Art Robbins Instruments, USA). Subsequently, the growth of NvCI-hCPA4 crystals was optimized at 18°C by hanging drop vapor diffusion method using a 24-well crystallization plate. The reservoir solution contained 0.04M ammonium nitrate and 25% w/v PEG3350. Single crystals appeared after 4 days from equal volumes of protein solution (17.6 mg/ml in 5 mM Tris pH 8.5, 50 mM NaCl) and reservoir solution. Crystals were cryo-protected in reservoir buffer containing 12% glycerol and flash-frozen in liquid nitrogen prior to diffraction analysis. Diffraction data were recorded from cryo-cooled crystals (100K) at Grenoble beamline ID23-2. Data were integrated and merged using XDS (Kabsch W., 2010), and scaled, reduced and further analyzed using CCP4 (Winn *et al.*, 2011).

III.16.3. Structure determination and refinement

The structure for the complex NvCI-hCPA4 was determined from the x-ray data at 1.7 Å by molecular replacement using the PDB from hCPA4 (PDB code 2PCU) as a model. The quality of the diffraction data allowed automatic building of the inhibitor using WARP (Langer *et al.*, 2008). Manual building and improvement of the model was performed using the program COOT (Emsley *et al.*, 2010). Refinement was performed by using the CNS (Brunger *et al.*, 2007) and Phenix programs (Adams *et al.*, 2010). Ramachandran analysis shows that 94.70% of residues (661) are in preferred regions, 4.58% of residues (32) are in allowed regions and 0.72% of residues (5) are in outlier regions, for both complexes of NvCI-hCPA4 in the asymmetric unit. All figures were prepared using PyMOL software (Delano W., 2002)

III.17. Enzymatic proteolysis of rNvCI

For analyses of enzymatic proteolysis of rNvCI, 1 μg of inhibitor was used for each enzyme using a 10:1 ratio (w:w) of inhibitor:enzyme, and reaction was carried out at 37ºC. Controls for inhibitor and enzyme stability were also performed, using rNvCI or enzyme in the respective activity buffer. The reaction progress was followed by MALDI-TOF MS, taking aliquots at 1 h, 6 h, and 18 h. For this, 1 μl of five replicates of each digestion mixture was deposited on a MTP 384 target plate polished steel T F (Bruker Daltonics), followed by deposition of 1 μl of DHAP as matrix. The mixture was allowed to dry at room temperature. Sequence analyses were made by using the Biotools software (version 3.2, Bruker Daltonics).

The following proteases and activity buffers were used: Bovine CPA and 20 mM Tris-HCl, pH 7.5, 500 mM NaCl; porcine CPB and 20 mM Tris-HCl, pH 7.5, 100 mM NaCl; papain from *C. papaya* and 20 mM Tris-HCl, pH 7.0, containing 100 M KCl, 0.1 mM EDTA and 3 mM DTT; modified trypsin from bovine pancreas and 100 mM Tris-HCl pH 8.5; subtilisin from *B. licheniformis* and 50 mM Tris-HCl, pH 8.5; porcine pepsin and 100 mM acetate buffer, pH 4.4; thermolysin from *Bacillus thermoproteolyticus rokko* (E.C.3.4.24.27) and 50 mM Tris-HCl, pH 7.5, 10 mM $CaCl_2$; Lys-C and 25 mM Tris-HCl pH 8.5, 1 mM EDTA; Glu-C and 100 mM ammonium bicarbonate pH 7.8; endoproteinase Asp-N from *Pseudomonas fragi* (E.C.3.4.24.33) and 50 mM sodium phosphate buffer, pH 8.0; a-chymotrypsin from bovine pancreas (E.C.3.4.21.1) and 100 mM Tris-HCl, pH 7.8, 10 mM $CaCl_2$; carboxypeptidase Y (CPY) from *Saccharomyces cerevisiae* (E.C.3.4.16.1) and 50 mM sodium phosphate buffer, pH 6.5, 150 mM NaCl; leucine aminopeptidase (LAP), microsomal from porcine kidney (E.C.3.4.11.1) and 20 mM Tris-HCl, pH 8.0, 5 mM $MgCl_2$. LAP was previously activated for two hours at 37ºC in 20 mM Tris-HCl, pH 8.0, 1 mM $MnCl_2$.

III.18. Preliminary study of antimalarial activity of rNvCI

All assays for antimalarial activity of rNvCI were performed in the laboratory of Prof. José María Bautista Santa Cruz, by members of his research unit at the department of Biochemistry and Molecular Biology IV, Veterinary Faculty, Universidad Complutense de Madrid, Spain.
Plasmodium falciparum Dd2 strains (chloroquine resistant) were used in this study. The cultures were maintained in cell culture flasks by mixing infected and uninfected erythrocytes up to a maximum hematocrit of 1.5%, depending on the percentage of parasitemia, culture volume and intraerythrocytic stage of the cycle, according to the methodology established by Radfar *et al.*, 2009). Cultures were incubated at 37°C in an atmosphere of 5% CO_2. Culture synchronization in different intraerythrocytic stages was performed as described by Radfar *et al.*, 2009. The strategy for the determination of antimalarial activity in rNvCI was carried out according to that described by Moneriz *et al.*, 2009.

The concentration value that is able to inhibit 50% of parasite growth (IC_{50}) was calculated by determining the percentage of parasitemia (%P) or parasite growth in the assay susceptible to the compound by the following equation:

$$\%P = \frac{\left(F_p - F_{cn}\right)}{\left(F_{cp} - F_{cn}\right)} x100 \qquad \textbf{(25)}$$

Where,
Fp is the unit of fluorescence of infected erythrocytes treated with the compound
Fcp is the unit of fluorescence of untreated infected erythrocytes and cultivated
Fcn is the unit of fluorescence of infected erythrocytes that did not grow (GRI)

The IC_{50} value was calculated by fitting the experimental data to a nonlinear regression using the Graphad prisma 5 software (GraphPad Software, Inc., USA) at $p<0.05$. Data are means (n=3 cultures) ± S.E.M.

III.2.19. Statistical analyses

In the statistical processing of experimental data were employed a simple linear regression analysis and a simple analysis of variance (ANOVA). Linear regression analysis was performed by using the SigmaPlot sotfware (version 10.0, Systat software Inc., USA).

The best-fit value of IC_{50} was obtained by adjusting the experimental values to the non-linear equation (Eq. 2) using the program GraphPad Prism 5 (GraphPad Software, Inc.) at $p<0.05$.

The best-fit value of K_M was obtained by adjusting the experimental values to the non-linear equation of Michaelis-Menten (Eq. 22) using the program GraphPad Prism 5 (GraphPad Software, Inc.) at $p<0.05$.

The best-fit value of K_i was obtained by adjusting the experimental values to the non-linear equation of Morrison (Eq. 23) using the program GraphPad Prism 5 (GraphPad Software, Inc.) at $p<0.05$.

Experimental data of the effect of preincubation time of enzyme-inhibitor and substrate concentration on residual enzymatic activity were processed by a simple ANOVA and comparison of sample means was carried out by a Tukey's HSD (honest significant difference) test, using the XLSTAT's statistical analysis software in Microsoft Office Excel 2003 (Microsoft Corp., USA). All assays were performed in triplicate and the results were presented as the mean ± standard deviation (SD)

IV. RESULTS AND DISCUSSION

IV.1 Identification in marine invertebrates of protease inhibitory activities. Dose-response relationships and determination of IC_{50} values

The marine Caribbean fauna is characterized by its richness and diversity, which makes it a natural source very attractive for the identification of biomolecules with biomedical potentialities. Among them, proteases and their inhibitors constitute the most promising group of biomolecules due to the increasing of effective protease targets identified for systemic and infectious diseases and efficient and selective inhibitors as drugs. Potential of marine invertebrates as source of these biomolecules has been well demonstrated by several authors who have found, particularly, protease inhibitors (PI) with exceptional structural and functional properties (Delfin *et al.*, 1996; Lenarcic et al., 1997; González *et al.*, 2007).

These reasons led to select 30 species of marine invertebrates, collected on the north coasts of Cuba, representing the main Phyla of marine invertebrates. These species were identified and classified by specialists from the Institute of Oceanology of Havana (Cuba).

Selection of marine invertebrate species into different Phyla was made on the basis of their representativeness and bioavailability. In order to obtain a better understanding of selected species, a taxonomic review was made, as shown in figure 30. Of them, two species of Annelida (segmented worms), two species of Bryozoa (Ectoprocts), six species of Chordata (sea squirts), eight species of Cnidaria (corals, flower animals, jellyfish, sea anemones), five species of Echinodermata (sea urchins, sand dollars, sea stars, sea cucumbers), four species of Mollusca (snails and slugs), and two species of Porifera (sponges) were used.

This work was devoted to the identification of protease inhibitors, such as metallocarboxypeptidase, aspartic, cysteine and serine protease inhibitors. Selection of the enzyme models was based on their potential in Biomedicine as targets and/or homologous targets in systemic and infectious diseases (Abbenante and Fairlie, 2005).

Given the above considerations, model enzymes belonging to four mechanistic classes of proteases were selected: bovine carboxypeptidase A and porcine carboxypeptidase B as models for metallo-exopeptidases, porcine pepsin, an aspartic enzyme representative of this mechanistic class, papain from *Carica papaya* as model enzyme for cysteine proteases, and bovine trypsin and subtilisin from *B. licheniformis* representing two families of serine proteases.

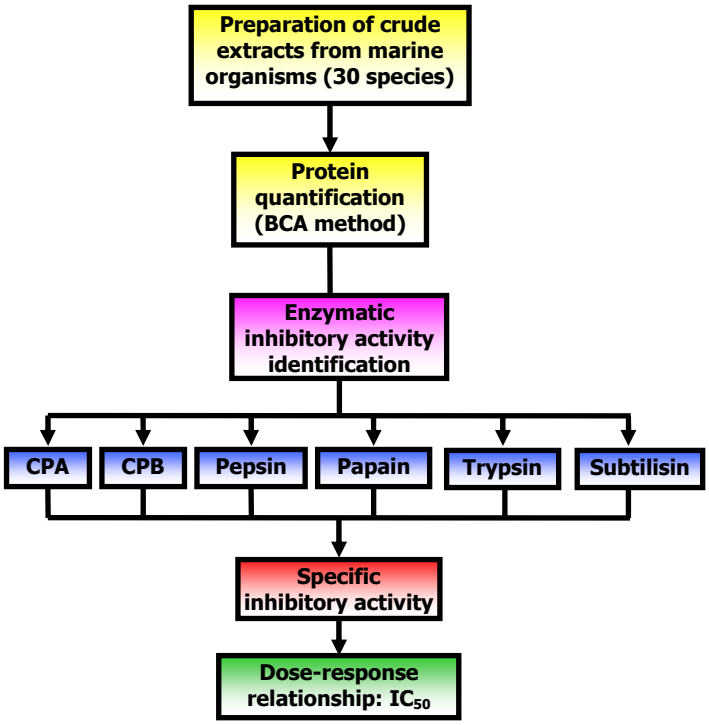


Figure 29. Flowchart process of the strategy used for kinetic identification of CPA, CPB, pepsin, papain, trypsin and subtilisin inhibitors

The strategy used (figure 29) after preparing crude extracts of marine invertebrates, was based in a first step, on using kinetic assays with specific substrates and experimental conditions for the identification of CPA, CPB, pepsin, papain, trypsin and subtilisin inhibitory activities, followed by the determination of their specific inhibitory activities and the establishment of dose-response curves, which allowed to calculate the IC_{50} values.

Following the strategy above mentioned, a general qualitative and quantitative screening was performed to detect inhibitory activities in the 30 marine extracts selected. Qualitative results are summarized in table 16, showing that fourteen extracts exhibited trypsin inhibitory activity, nine extracts displayed CPA, CPB, papain and subtilisin inhibitory activities, whilst only four extracts were identified for pepsin inhibitory activity.

It should be noted that these results could be affected by the presence of non-dissociated endogenous protease-inhibitor complexes during the preparation of the extracts, which do not allow the real detection of protease inhibitory activities. This fact was previously confirmed in different marine extracts, using agents allowing the dissociation of complexes, such as trichloroacetic acid (TCA) (Delfín *et al.*, 1996). On the other hand, endogenous PI are generally tight-binding inhibitors, with values of Ki 10^{-7} - 10^{-8} M or even lower. Thus, they cannot be easily dissociated by the presence of the substrate in the assay, particularly at the low substrate concentrations used (~1 Km) (Bieth, J.G., 1995).

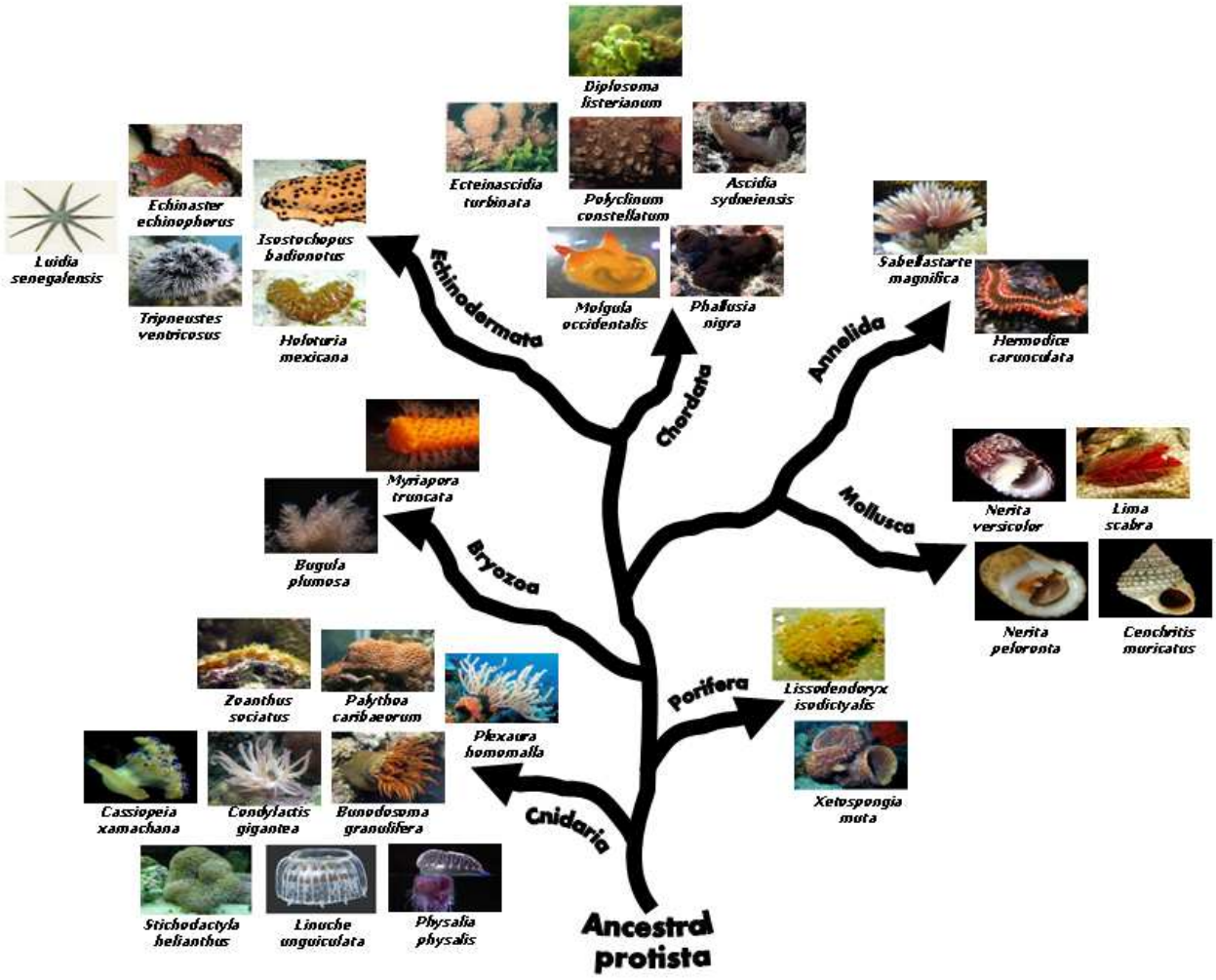


Figure 30. Marine species used in the screening for CPA, CPB, pepsin, papain trypsin and subtilisin inhibitors

Quantification of inhibitory activity in positive extracts was performed by assessing their specific inhibitory activities using appropriate experimental conditions for the detection of tight-binding inhibition. Taking into account that the inhibition degree of these inhibitors varies not only with the inhibitor concentration, but also the enzyme concentration, in most cases, high enzyme concentrations were employed in order to achieve a linear relationship between the inhibition degree (residual activity) and inhibitor concentration. However, in these conditions, linearity was not achieved in all cases due to several factors, such as the strength of the inhibitor and/or the sensitivity of the substrate. High sensitivity substrates do not allow obtaining good measurements of initial velocities at high enzyme concentrations (Bieth, J.G., 1995; Chávez and Gutiérrez, 2012). On the other hand, substrate concentrations were kept low to avoid possible dissociation of the enzyme-inhibitor complexes induced by the substrate, in the case of competitive inhibitors.

Table 16. Screening of 30 marine invertebrate extracts for CPA, CPB, pepsin, papain, trypsin and subtilisin inhibitory activities

Species	Phyllum	CPA inhibitory activity	CPB inhibitory activity	Pepsin inhibitory activity	Papain inhibitory activity	Trypsin inhibitory activity	Subtilisin inhibitory activity
Hermodice carunculata	Annelida	+	+	+		+	
*Sabellastarte magnifica**	Annelida						+
Bugula plumosa	Bryozoa						
Myriapora truncata	Bryozoa						
Ascidia sydneiensis	Chordata						
Diplosoma listerianum	Chordata			+			
Ecteinascidia turbinata	Chordata						
Molgula occidentalis	Chordata	+	+	+		+	
Phallusia nigra	Chordata					+	
Polyclinum constellatum	Chordata						
Bunodosoma granulifera	Cnidaria	+	+	+		+	+
Cassiopeia xamachana	Cnidaria						
Condylactis gigantea	Cnidaria	+	+		+	+	+
Linuche unguiculata	Cnidaria						
Palythoa caribbaeorum	Cnidaria						
Physalia physalis	Cnidaria				+	+	
Plexaura homomalla	Cnidaria	+	+		+		
Stichodactyla helianthus	Cnidaria	+	+		+	+	
Zoanthus sociatus	Cnidaria						
Echinaster echinophorus	Echinodermata					+	+
Holothuria mexicana	Echinodermata						
Isostochopus badionotus	Echinodermata				+	+	+
Luidia senegalensis	Echinodermata						
Tripneustes ventricosus	Echinodermata					+	
Cenchritis muricatus	Mollusca	+	+		+	+	+
Lima scabra	Mollusca						
Nerita peloronta	Mollusca	+	+		+	+	+
Nerita vericolor	Mollusca	+	+		+	+	+
Lissodendorix isodyctialis	Porifera				+	+	+
Xetospongia muta	Porifera						

*In the case of marine invertebrate *S. magnifica*, the animal was separated in two parts: tentacle crown and body. The body section was used in this screening, in which was previously described and characterized a metallo-carboxypeptidase (Alonso del Rivero *et al.*, 2009).

Table 17 shows the values of total protein concentration, inhibitory activity (mU/ml extract), specific inhibitory activity and IC_{50} values of the extracts with CPA inhibitory activity. Among them, *N. versicolor* extract showed the highest CPA inhibitory activity and lowest IC_{50} value, followed by *P. homomalla* and *S. helianthus* crude extracts. Other extracts with IC_{50} values less than 1 mg/ml such as *C. muricatus*, *M. occidentalis* and *N. peloronta* could be also good candidates for the isolation of efficient MCP inhibitors. However, results obtained with extracts of the Phylum Chordata, such as *M. occidentalis* which are known as metal deposits, must be taken in caution (Chen and Mayer, 2000).

Table 17. Results of the screening for CPA inhibitory activity in marine invertebrates

Species	Protein concentration (mg/ml)	Inhibitory activity (mU/ml)	Specific inhibitory activity (mU/mg)	IC50 (mg/ml)
H. carunculata	6.92 ± 0.25	18.89 ± 3.06	2.73 ± 0.44	0.98 ± 0.04
M. occidentalis	17.56 ± 0.42	107.06 ± 27.61	6.10 ± 1.57	0.44 ± 0.04
B. granulifera	13.41 ± 0.20	12.78 ± 2.45	0.95 ± 0.18	1.80 ± 0.29
C. gigantea	31.40 ± 1.49	37.66 ± 4.80	1.20 ± 0.15	2.28 ± 0.12
P. homomalla	11.42 ± 0.11	250.99 ± 45.66	21.98 ± 4.00	0.10 ± 0.01
S. helianthus	16.64 ± 0.35	269.28 ± 96.3	16.19 ± 5.8	0.15 ± 0.01
C. muricatus	18.67 ± 0.49	153.60 ± 5.38	8.23 ± 0.29	0.374 ± 0.003
N. peloronta	20.66 ± 0.19	66.91 ± 15.8	3.24 ± 0.77	0.57 ± 0.02
N. versicolor	27.35 ± 1.13	814.29 ± 178.52	29.77 ± 6.53	0.051 ± 0.002

Data are means (n=3) ± S.D.

Dose-response relationships of CPA inhibitory activity for the nine positive extracts displayed residual activity decreases (increase of inhibitory activity) as a function of extract concentration (figure 31), which is usual in the presence of inhibitors. Those extracts that showed the most significant decrease in residual activity (higher slopes) led to lower IC_{50} values. These results were consistent with those obtained in terms of specific inhibitory activity, which are typical behaviours of tight-binding inhibitors, where a linear dependence between residual activity and inhibitor concentration is established (Bieth, J.G., 1995). However, as this dependence is also determined by the possibility of using very high enzyme concentrations and substrates with low sensitivity, as well as a right equilibrium time, it is not possible to affirm that a non-linear dependence corresponds to a classical inhibitor instead of a tight-binding inhibitor.

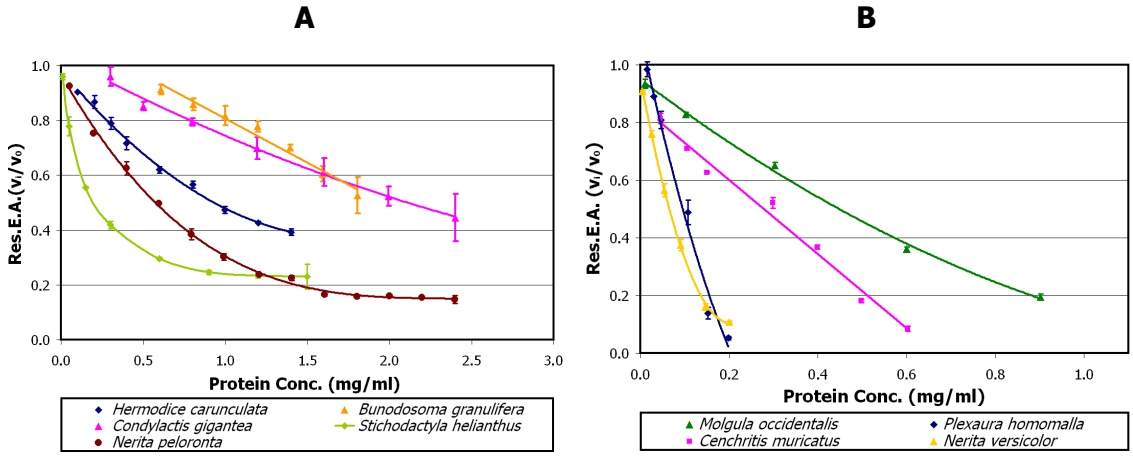


Figure 31. Dose-response relationships for CPA inhibitory activity in marine crude extracts. Effects of different doses of crude extracts on CPA activity: (A) *H. carunculata, B. granulifera, C. gigantea, S. helianthus,* and *N. peloronta;* and (B) *M. occidentalis, P. homomalla, C. muricatus,* and *N. versicolor.* $7.0x10^{-9}$ M CPA, $1.0x10^{-4}$ M AAFP substrate in a final volume of 250 µl, pH=7.5, T=25 ºC. Pre-incubation time, 10 min at 25 ºC. Res.E.A.: Residual Enzymatic Activity (Mock *et* al., 1996). The best-fit value of IC_{50} was performed by adjusting the experimental values to the equation (Eq. 2, Materials and Methods) using the program GraphPad Prism 5 (GraphPad Software, Inc.) at $p<0.05$. Data are means (n=3) ± S.D.

Regarding CPB inhibitory activity, as mentioned above, nine extracts were positive (table 18) and dose-response curves are shown in figure 32. The crude extract of *N. versicolor* as in the case of CPA was the most promising with a remarkable specific inhibitory activity and very low IC_{50} value, followed by *C. muricatus* and *P. homomalla* crude extracts. In addition, extracts of *N. peloronta* and *H. carunculata* could be also considered for the isolation of CPB inhibitors (IC_{50} less than 1 mg/ml).

Table 18. Results of the screening for CPB inhibitory activity in marine invertebrates

Species	Protein concentration (mg/ml)	Inhibitory activity (mU/ml)	Specific inhibitory activity (mU/mg)	IC50 (mg/ml)
H. carunculata	6.92 ± 0.25	27.34 ± 4.28	3.95 ± 0.62	0.70 ± 0.02
M. occidentalis	17.56 ± 0.42	233.88 ± 44.55	13.32 ± 2.54	0.20 ± 0.03
B. granulifera	13.41 ± 0.20	14.86 ± 3.52	1.11 ± 0.26	1.50 ± 0.09
C. gigantea	31.40 ± 1.49	24.03 ± 4.17	0.77 ± 0.13	1.65 ± 0.22
P. homomalla	11.42 ± 0.11	128.37 ± 27.64	11.24 ± 2.42	0.20 ± 0.01
S. helianthus	16.64 ± 0.35	26.79 ± 4.90	1.61 ± 0.29	1.49 ± 0.08
C. muricatus	18.67 ± 0.49	442.30 ± 103.68	23.69 ± 5.55	0.123 ± 0.002
N. peloronta	20.66 ± 0.19	153.16 ± 24.62	7.41 ± 1.19	0.29 ± 0.01
N. versicolor	27.35 ± 1.13	2276.68 ± 530.26	83.24 ± 19.39	0.025 ± 0.002

Data are means (n=3) ± S.D.

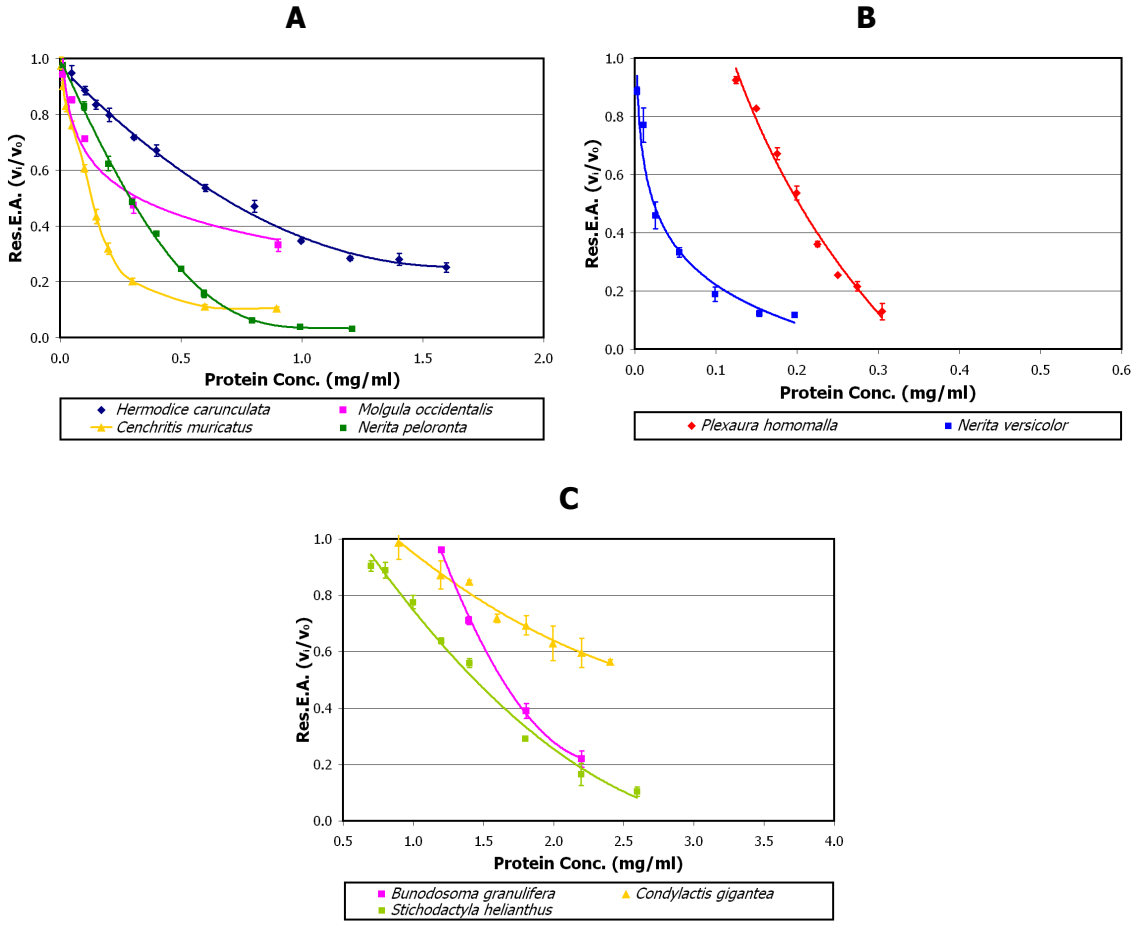


Figure 32. Dose-response relationships for CPB inhibitory activity in marine crude extracts. Effects of different doses of crude extracts on CPB activity: **(A)** *H. carunculata, M. occidentalis, C. muricatus,* and *N. peloronta.* **(B)** *P. homomalla,* and *N. versicolor.* **(C)** *B. granulifera, C. gigantea,* and *S. helianthus.* $3.0x10^{-9}$ M CPB, $1.0x10^{-4}$ AAFA substrate in a final volume of 250 µl, pH=7.5, T=25 ºC. Pre-incubation time, 10 min at 25 ºC. Res.E.A.: Residual Enzymatic Activity (Mock and Stanford, 2002). The best-fit value of IC_{50} was performed by adjusting the experimental values to the equation (Eq. 2, Materials and Methods) using the program GraphPad Prism 5 (GraphPad Software, Inc.) at $p<0.05$. Data are means (n=3) ± S.D.

The presence of pepsin inhibitory activity was not frequent in the extracts tested, since only four extracts were positive, with relatively low values of specific inhibitory activity (table 19). The extract showing the highest specific inhibitory activity and the lowest IC_{50} value was *B. granulifera*, followed by *H. carunculata*, *D. listerianum* and *M. occidentalis*. It is important to highlight the low IC_{50} value obtained for *B. granulifera* that makes it the most promising extract.

The other three extracts could be also considered interesting for the isolation of pepsin inhibitors. Dose-response relationships for pepsin inhibitory activity are shown in figure 33.

Table 19. Results of the screening for pepsin inhibitory activity in marine invertebrates

Species	Protein concentration (mg/ml)	Inhibitory activity (mU/ml)	Specific inhibitory activity (mU/mg)	IC50 (mg/ml)
H. carunculata	6.07 ± 0.31	55.19 ± 10.71	9.10 ± 1.77	0.153 ± 0.003
D. listerianum	1.15 ± 0.01	8.75 ± 1.44	7.59 ± 1.25	0.230 ± 0.010
M. occidentalis	17.42 ± 0.84	56.62 ± 4.68	3.25 ± 0.27	0.429 ± 0.022
B. granulifera	15.14 ± 0.15	90.97 ± 21.08	13.30 ± 0.36	0.059 ± 0.006

Data are means (n=3) ± S.D.

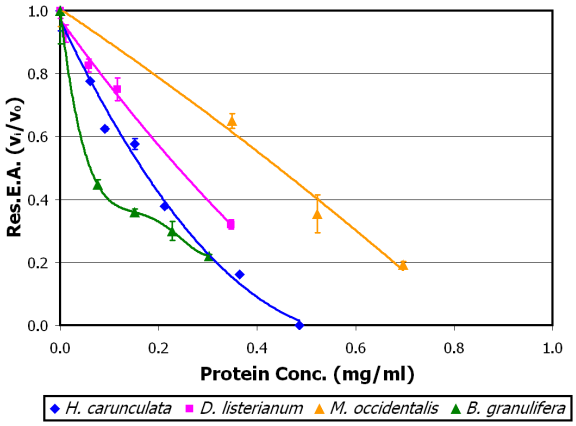


Figure 33. Dose-response relationships for pepsin inhibitory activity in marine crude extracts.

Effects of different doses of crude extracts on pepsin activity: *M. occidentalis* and *B. granulifera*. $1.1x10^{-8}$ M pepsin, $2.0x10^{-4}$ LSPNPLAL substrate in a final volume of 1 ml, pH=4.4, T=25 ºC. Pre-incubation time, 10 min at 25 ºC. Res.E.A.: Residual Enzymatic Activity (Martín *et al.*, 1980). The best-fit value of IC_{50} was performed by adjusting the experimental values to the equation (Eq. 2, Materials and Methods) using the program GraphPad Prism 5 (GraphPad Software, Inc.) at $p<0.05$. Data are means (n=3) ± S.D.

Papain inhibitory activity was detected in nine extracts, from which *P. homomalla* extract stands out for its high value of specific inhibitory activity and low IC_{50} value, compared to the rest of the positive extracts (table 20). Other good results were obtained with *P. physalis, C. muricatus, S. helianthus* and *N. peloronta* crude extracts. Dose-response curves for papain inhibitory activity are shown in figure 34.

It is interesting to remark that a high protein concentration of an extract, negatively affects both, specific inhibitory activity and IC_{50} value, as in the case of *N. peloronta* extract. Its lower value of specific inhibitory activity could be mainly explained by its high protein content, unlike the other extracts, affecting the IC_{50} value, since this parameter is based on the dose-response relationship. Another factor to be considered for the selection of the most promising extracts is the bioavailability of the species in nature, which is the case of this sea snail.

Table 20. Results of the screening for papain inhibitory activity in marine invertebrates

Species	Protein concentration (mg/ml)	Inhibitory activity (mU/ml)	Specific inhibitory activity (mU/mg)	IC50 (mg/ml)
C. gigantea	31.40 ± 1.49	112.83 ± 21.47	3.59 ± 0.68	1.86 ± 0.03
P. physalis	4.90 ± 0.03	62.92 ± 7.44	12.83 ± 1.52	0.338 ± 0.001
P. homomalla	11.42 ± 0.11	984.01 ± 185.61	86.16 ± 16.25	0.09 ± 0.01
S. helianthus	16.64 ± 0.35	144.24 ± 26.63	8.67 ± 1.60	0.41 ± 0.03
I. badionotus	4.20 ± 0.03	19.23 ± 3.23	4.58 ± 0.77	1.26 ± 0.13
C. muricatus	18.67 ± 0.49	200.03 ± 42.38	10.72 ± 2.27	0.54 ± 0.03
N. peloronta	20.66 ± 0.19	166.58 ± 41.05	8.06 ± 1.99	0.46 ± 0.04
N. versicolor	27.35 ± 1.13	71.41 ± 16.09	2.61 ± 0.59	3.46 ± 0.13
L. isodyctialis	28.00 ± 0.33	64.42 ± 15.39	2.30 ± 0.55	2.95 ± 0.06

Data are means (n=3) ± S.D.

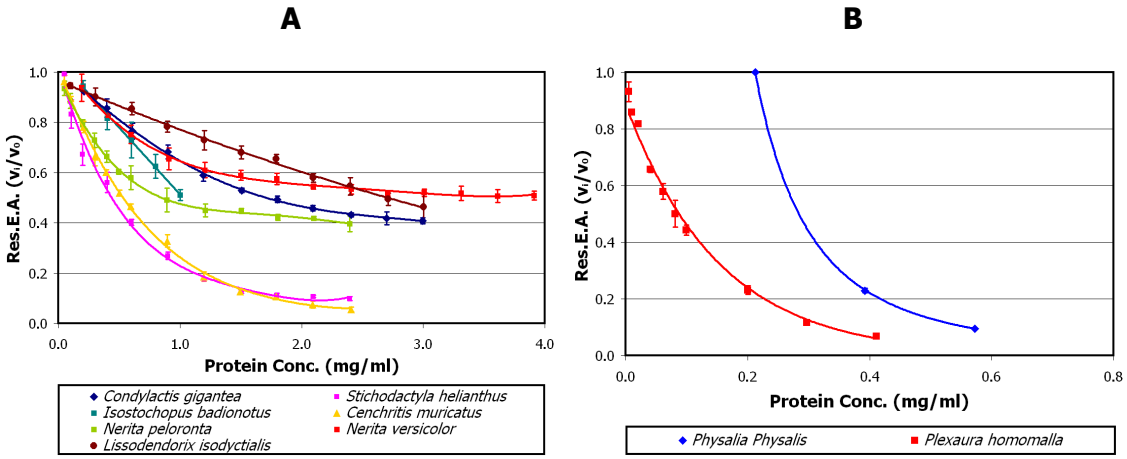


Figure 34. Dose-response relationships for papain inhibitory activity in marine crude extracts.
Effects of different doses of crude extracts on papain activity: (A) *C. gigantea, S. helianthus, I. badionotus, C. muricatus, N. peloronta, N. versicolor* and *L. isodyctialis;* and (B) *P. physalis,* and *P. homomalla.* $4.5x10^{-8}$ M papain, $4.0x10^{-4}$ PFLNA substrate in a final volume of 250 µl, pH=6.5, T=25 ºC. Pre-incubation time, 10 min at 25 ºC. Res.E.A.: Residual Enzymatic Activity (Filippova *et al.*, 1984). The best-fit value of IC_{50} was performed by adjusting the experimental values to the equation (Eq. 2, Materials and Methods) using the program GraphPad Prism 5 (GraphPad Software, Inc.) at $p<0.05$.
Data are means (n=3) ± S.D.

The presence of trypsin inhibitory activity was detected in fourteen extracts (table 21). The best results in terms of specific inhibitory activity and IC_{50} value were obtained for *S. helianthus* extract followed with a high difference by *N. versicolor, C. muricatus, N. peloronta* and in a less extent the rest of crude extracts wit IC_{50} values less than 1 mg/ml.

It is important to note that the IC_{50} value shown by *S. helianthus* extract is 8.7 times lower than the obtained for *N. versicolor* extract. This is well observed in the behaviours of the dose-response curves (figure 35), where the sea anemone extract revealed a ten to one hundred times higher slope compared to all positive extracts identified.

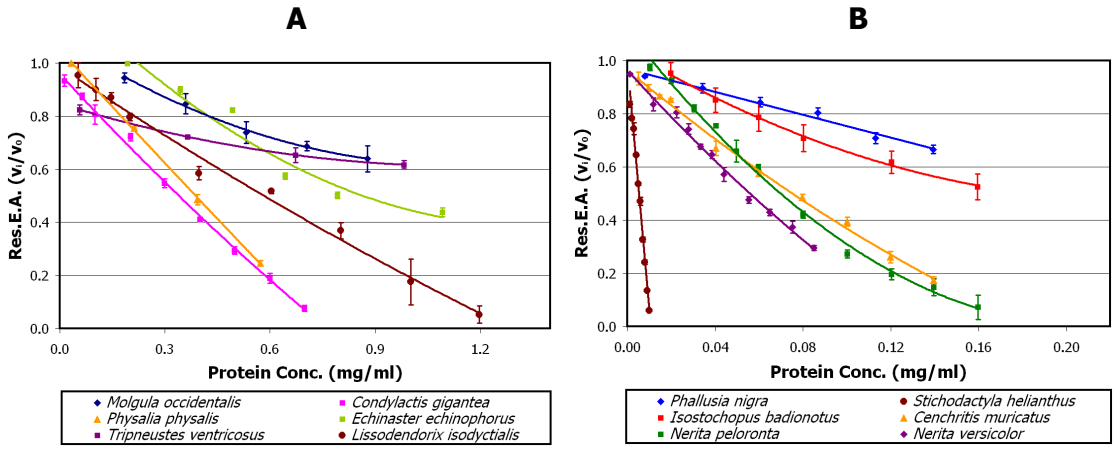
Table 21. Results of the screening for trypsin inhibitory activity in marine invertebrates

Species	Protein concentration (mg/ml)	Inhibitory activity (mU/ml)	Specific inhibitory activity (mU/mg)	IC50 (mg/ml)
H. carunculata	6.92 ± 0.25	19.69 ± 4.25	2.85 ± 0.61	2.02 ± 0.12
M. occidentalis	17.56 ± 0.42	90.34 ± 13.30	5.14 ± 0.76	1.25 ± 0.04
P. nigra	1.19 ± 0.01	33.41 ± 3.19	27.99 ± 2.67	0.28 ± 0.03
B. granulifera	13.41 ± 0.20	32.72 ± 6.00	2.44 ± 0.45	3.06 ± 0.61
C. gigantean	31.40 ± 1.49	538.94 ± 105.56	17.17 ± 3.36	0.365 ± 0.003
P. physalis	4.90 ± 0.03	75.77 ± 5.18	15.45 ± 1.06	0.365 ± 0.004
S. helianthus	16.64 ± 0.35	16057.59 ± 1182.95	965.16 ± 71.10	0.0062 ± 0.0001
E. echinophorus	13.10 ± 0.08	101.85 ± 13.82	7.78 ± 1.06	0.86 ± 0.02
I. badionotus	4.20 ± 0.03	143.22 ± 28.89	34.10 ± 6.88	0.19 ± 0.01
T. ventricosus	8.41 ± 0.29	52.91 ± 8.69	6.29 ± 1.03	3.58 ± 0.12
C. muricatus	18.48 ± 0.57	1452.37 ± 225.71	78.60 ± 12.22	0.078 ± 0.002
N. peloronta	20.66 ± 0.19	1263.64 ± 217.64	61.15 ± 10.53	0.070 ± 0.003
Nerita versicolor	25.78 ± 1.21	2866.17 ± 438.78	111.19 ± 17.02	0.054 ± 0.001
L. isodyctialis	28.00 ± 0.33	241.93 ± 45.64	8.64 ± 1.63	0.582 ± 0.01

Data are means (n=3) ± S.D.

Trypsin inhibitory activity detected in the *S. helianthus* extract could be probably related to the presence of three strong trypsin inhibitors earlier described by Chávez *et al.*, 1988; Antuch *et al.*, 1993; Delfin *et al.*, 1996; García Fernández *et al.*, 2011).

The main inhibitor, ShPI-I (UniProt ID: ISH1_STOHE) is a Kunitz-type inhibitor, which inhibits serine proteases such as trypsin (K_i = 1.3x10^{-10} M), plasmin, chymotrypsin and kallikrein and proteases belonging to other mechanistic classes such as cysteine proteases (papain) and aspartic proteases (pepsin). This fact is in agreement with the positive result obtained with this extract against papain. However, pepsin inhibitory activity was not detected in this crude extract as a result of using a low molecular weight substrate such as the hexapeptide employed in this work. Pepsin inhibitory activity of ShPI-I was only detected in the presence of high molecular weight substrates such as hemoglobin (Delfín *et al.*, 1996). The two other protease inhibitors isolated from this organism are called ShPI-II (ISH2_STOHE) (Díaz *et al.*, 1998) and ShPI-III. On the other hand, the positive result displayed by *C. muricatus* extract could be ascribed to the presence in this extract of a Kazal-type trypsin inhibitor (CMPI-II, UniProt ID: IPK2_CENMR) isolated from this extract, which is able to inhibit trypsin, subtilisin and elastase with K_i values in the nanomolar range (González *et al.*, 2007a).



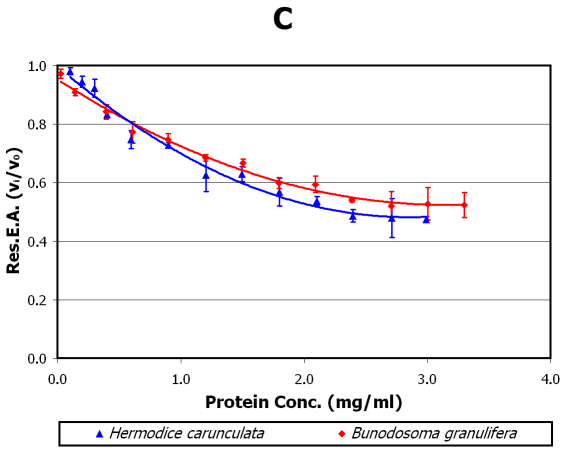


Figure 35. Dose-response relationships for trypsin inhibitory activity in marine crude extracts.
Effects of different doses of crude extracts on trypsin activity: (A) *M. occidentalis, C. gigantea, P. physalis, E. echinophorus, T. ventricosus,* and *L. isodyctialis;* and (B) *P. nigra, S. helianthus, I. badionotus, C. muricatus, N. peloronta,* and *N. versicolor.* (C) *H. carunculata* and *B. granulifera*. $2.8x10^{-7}$ M trypsin, $1.0x10^{-3}$ BAPNA substrate in a final volume of 250 µl, pH=8.0, T=25 ºC. Pre-incubation time, 10 min at 25 ºC. Res.E.A.: Residual Enzymatic Activity (Erlanger *et al.*, 1961). The best-fit value of IC_{50} was performed by adjusting the experimental values to the equation (Eq. 2, Materials and Methods) using the program GraphPad Prism 5 (GraphPad Software, Inc.) at $p<0.05$.
Data are means (n=3) ± S.D.

Subtilisin inhibitory activity, with nine positive extracts (table 22), was much less frequent than trypsin inhibitory activity. This result is related to the major presence in nature of serine proteases of chymotrypsin family (to which trypsin belongs), compared to the subtilisin family, at least in terms of number of entries in the MEROPS database (Rawlings *et al.*, 2002): 13408 and 4666 annotated sequences for chymotrypsin and subtilisin families, respectively. The most promising marine invertebrate extracts were *L. isodyctialis*, *I. badionotus* and *C. muricatus*. Despite of the 10 times higher IC_{50} value compared to the above selected extracts, *N. peloronta*, *N. versicolor*, *C. gigantea* and *E. echinophorus* could be good extracts as sources of subtilisin inhibitors.

Table 22. Results of the screening for subtilisin inhibitory activity in marine invertebrates

Species	Protein concentration (mg/ml)	Inhibitory activity (mU/ml)	Specific inhibitory activity (mU/mg)	IC50 (mg/ml)
S. magnifica	17.05 ± 1.76	150.54 ± 30.53	8.83 ± 1.79	0.44 ± 0.01
B. granulifera	8.11 ± 0.53	37.75 ± 7.09	4.66 ± 0.88	0.55 ± 0.02
C. gigantea	24.30 ± 0.76	252.33 ± 37.22	10.39 ± 1.53	0.254 ± 0.004
E. echinophorus	13.48 ± 0.35	161.13 ± 11.55	11.95 ± 0.86	0.37 ± 0.01
I. badionotus	4.20 ± 0.03	485.33 ± 138.94	115.55 ± 33.08	0.026 ± 0.001
C. muricatus	18.48 ± 0.57	1618.36 ± 322.81	87.59 ± 17.47	0.05654 ± 0.00005
N. peloronta	19.84 ± 1.43	362.66 ± 64.30	18.28 ± 3.24	0.199 ± 0.003
N. versicolor	25.78 ± 1.21	610.91 ± 155.81	23.70 ± 6.04	0.22 ± 0.02
L. isodyctialis	28.00 ± 0.33	3314.14 ± 425.41	118.36 ± 15.19	0.0260 ± 0.0003

Data are means (n=3) ± S.D.

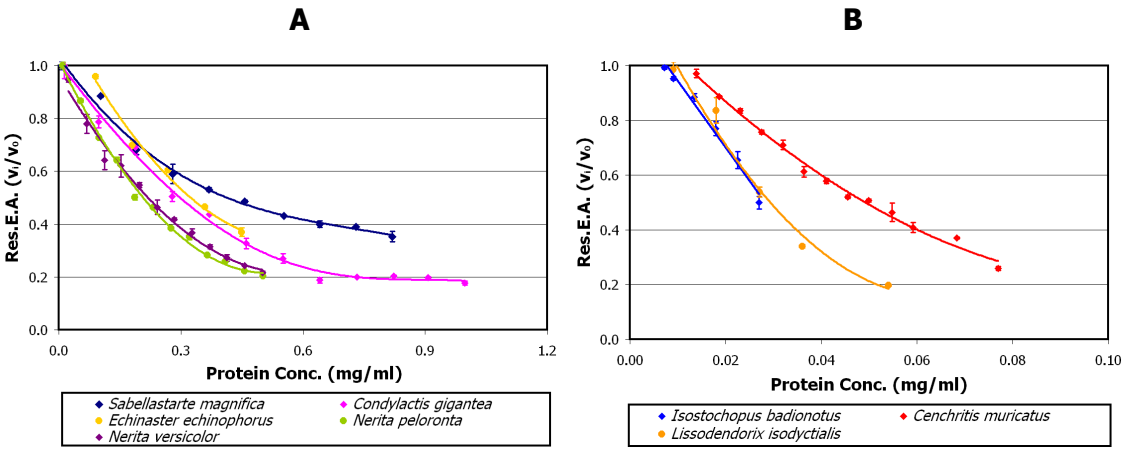


Figure 36. Dose-response relationships for subtilisin inhibitory activity in marine crude extracts.

Effects of different doses of crude extracts on subtilisin activity: (A) *S. magnifica, C. gigantea, E. echinophorus, N. peloronta,* and *N. versicolor;* and (B) *I. badionotus, C. muricatus,* and *L. isodyctialis.* $2.0x10^{-7}$ M subtilisin, $4.0x10^{-4}$ GGLPNA substrate in a final volume of 250 µl, pH=8.6, T=25 ºC. Pre-incubation time, 10 min at 25 ºC. Res.E.A.: Residual Enzymatic Activity (Lyublinskaya *et al.*, 1974). The best-fit value of IC_{50} was performed by adjusting the experimental values to the equation (Eq. 2, Materials and Methods) using the program GraphPad Prism 5 (GraphPad Software, Inc.) at $p<0.05$.
Data are means (n=3) ± S.D.

Experimental conditions affect inhibitory activities, particularly enzyme concentrations. It has been described that IC_{50} value for tight-binding inhibitors linearly depends on the enzyme concentration (Copeland, R.A., 2000). However, these experimental conditions, as well as others (substrate concentration, pH, ionic strength, equilibrium time, etc.) were the same for all the extracts evaluated against one specific enzyme. Moreover, substrate concentrations were maintained low to avoid substrate induced dissociation. Nevertheless, it has been very well demonstrated that even if the experimental conditions are the same, time to reach equilibrium may be different depending on the inhibitor nature and its mechanism (fast or slow tight-binding inhibitor) (Bieth, J.G., 1995).

On the other hand, it has been earlier mentioned that specific inhibitory activities and IC_{50} values could be affected by the inhibitor content/total protein ratio in the crude extracts. In addition, inhibitory activity could increase during the purification process due to the dissociation of endogenous complexes (Delfín *et al.*, 1996).

Therefore, other behaviours of dose-response relationships with lower slopes can not be discarded because they also depend, not only on the strength of the inhibitor, but also on several other factors (Bieth, J.G., 1995; Copeland *et al.*, 1995, Chávez and Gutiérrez, 2012).

IV.2. Immobilization of target proteases

In order to detect protease inhibitors in complex biological mixtures by IF MALDI-TOF MS as well as their subsequent purification by affinity chromatography, the immobilization-stabilization of the target proteases on glyoxal Sepharose support was performed, according to the method described by Guisán, J.M. 1988.

Glyoxal Sepharose supports display suitable properties for immobilization-stabilization of proteins by covalent attachment. Based on these properties, the immobilization of bovine CPA, porcine CPB, papain, bovine trypsin and subtilisin were performed on this solid support at pH 10.5. This pH value allowed a high immobilization degree due to the additional presence of non charged ε-NH_2 groups of Lysine ($pK\varepsilon_{NH3+}$ about 10.5) able to interact with the support, in addition to the N-terminus groups.

In the case of porcine pepsin, NHS-Sepharose was used as support. Coupling of this enzyme to the support was performed at pH 6.0, since pepsin, although not active in the range of pH 6-7, is stable in solution, while increasing pH above 8.0, the enzyme is irreversibly inactivated (Ryle, A.P., 1970). Therefore, the glyoxal-Sepharose support is not appropriated for the immobilization of this enzyme, because coupling, as discussed above, is performed at alkaline pH (10.5-10.7).

The strategy followed for the immobilization-stabilization of target proteases included different steps, such as the inclusion of a spacer arm and activation of the support and finally the enzyme coupling on the activated Sepharose, using three different enzyme loads. The time progression of this stage was monitored by quantifying protein concentration, as well as enzymatic activity, both in the supernatants. Once established the final time of protease coupling, enzyme stabilization on the support was carried out by reducing and/or blocking reactive groups. Then, the characterization of immobilized derivatives was performed through the determination of protein concentration, enzymatic activity and SDS-PAGE analysis. Immobilization parameters such as degree of enzyme immobilization in terms of protein and enzymatic activity, as well as retention of functional activity were evaluated.

Finally, the effect of enzyme loading on external and internal mass transfer was studied in order to establish the best loading with minimal steric hindrances and diffusional limitations. Figure 37 summarizes the strategy.

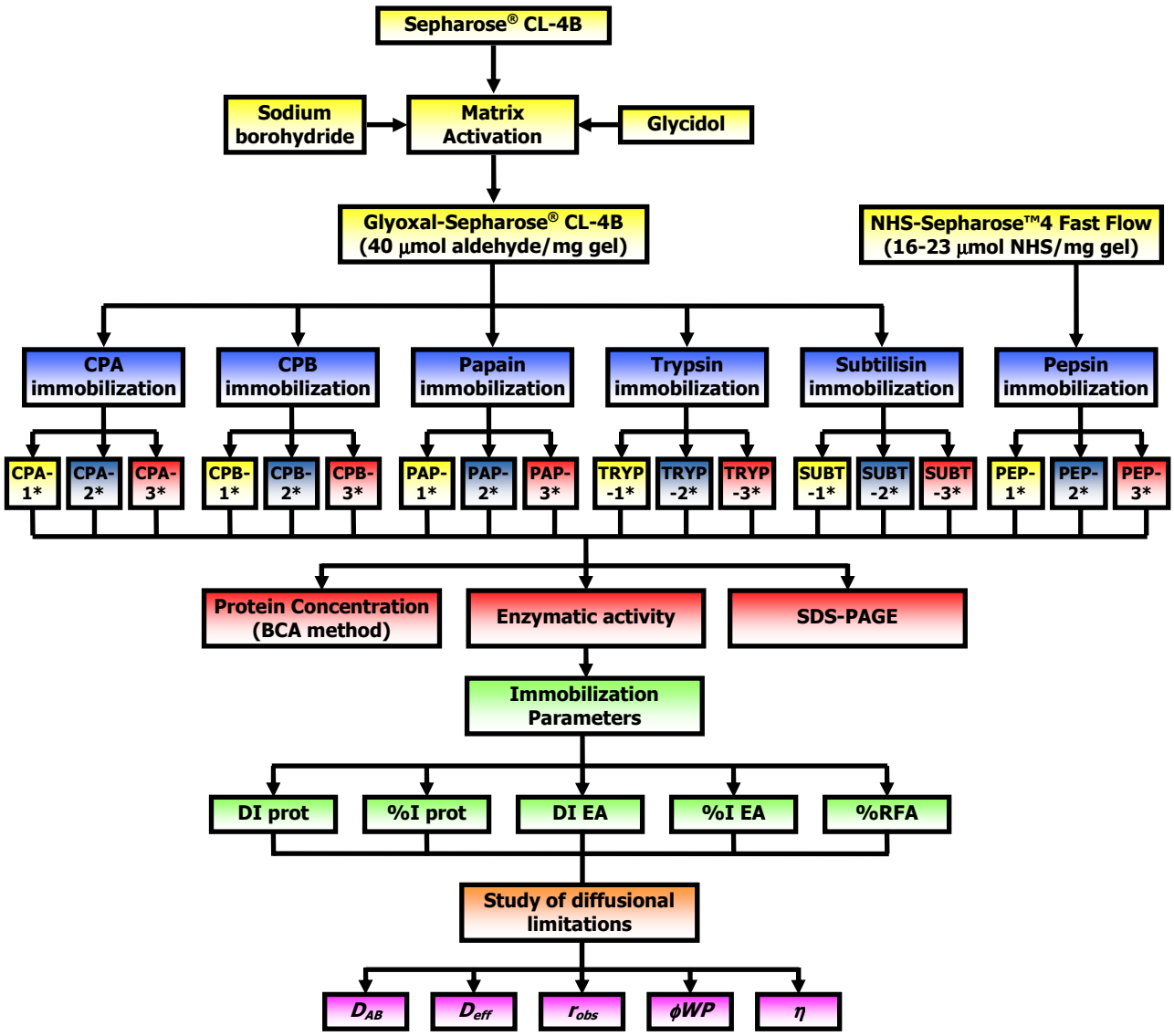


Figure 37. Flowchart process for the immobilization of target proteases on activated Sepharose® supports.

* 1, *2 and *3 represent different enzymatic loads. DI prot: degree of enzyme immobilization in terms of protein. %I prot: percentage of enzyme immobilization in terms of protein. DI EA: degree of enzyme immobilization in terms of enzymatic activity. %I EA: percentage of enzyme immobilization in terms of enzymatic activity. %RFA: percentage of retention of functional activity. ϕWP: Weisz-Prater modulus. D_{AB}: diffusion coefficient of solute A in solvent B (cm^2/s). D_{eff}: effective diffusion coefficient (cm^2/s) r_{obs}: observed reaction rate (mmol/h.g catalyst). η: effectiveness factor.

IV.2.1. Activation degree of Sepharose® support with glyoxal groups

The support employed for the immobilization of target porteases, with the exception of pepsin, was Sepharose® CL-4B. The overall activation (etherification to obtain glyceryl groups from glycidol, and periodate oxidation to produce aldehyde groups from glyceryl groups) was performed in mild experimental conditions to obtain intact sepharose gels, containing monolayers of identical aldehyde groups (Pedroche *et al.*, 2002). The amount of aldehyde groups in the support was controlled by the addition of a determined amount of periodate. Its quantification was performed by measuring the Na_IO_4 that was not consumed in the reaction by titration with sodium iodide (KI), using the calibration curve [Na_IO_4] *vs.* DO_{390} shown in figure 38. Thus, 37.7 μmol aldehyde/ml gel were obtained which represents a good activation degree of the support in order to obtain enzyme-support multi-interaction (table 23) (Grazu *et al.*, 2006; Mateo *et al.*, 2006).

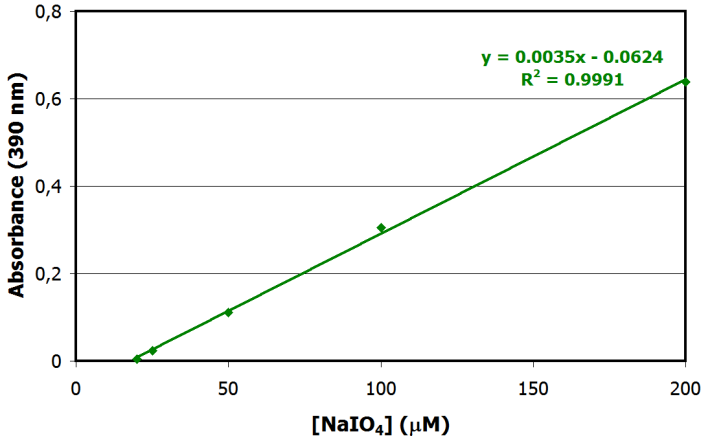


Figure 38. Calibration curve of [$NaIO_4$] *vs.* OD_{390} for the determination of activation degree of Sepharose® support with glyoxal groups

The immobilization of target proteases was performed at 4ºC with the exception of subtilisn, where the coupling step was carried out at room temperature, since as will be explained later, subtilisin coupling in the first condition of temperature did not produce high/moderate degrees of immobilization. The enzyme immobilization was performed by end-over-end rotation of the reaction suspension during a period of two and six hours, depending on the coupling rate of each enzyme on the support. In the case of glyoxal Sepharose support, a pH=10.5 was used in the coupling step, while in the case of NHS Sepharose, employed in the pepsin immobilization, a pH=6.0 was used during the enzyme immobilization. Other experimental conditions are described in Materials and Methods (paragraphs III.6.1 to III.6.3).

Table 23. Summary of results obtained for the activation degree of glyoxal Sepharose® support

[$NaIO_4$]final (μM)	1168.7
V total solution (ml)	130
μmol $NaIO_4$ final	151.9
μmol $NaIO_4$ initial	2600
V gel (ml)	65
Activation degree (μmol aldehyde /ml gel)	**37.7**

IV.2.2. Immobilization of bovine CPA on glyoxal Sepharose® CL-4B

Immobilization of bovine CPA on glyoxal Sepharose® support was carried out at 4ºC for 4.5 h. In the time course of immobilization, a decrease in both, the amount of protein and the total units of enzymatic activity were observed, indicative of the enzyme immobilization on the support (figure 39). At the immobilization time of 1.5 h, percentages of enzyme immobilization in terms of protein (%I prot) of 80.0, 55.7 and 53.0% were reached for CPA-1, CPA-2 and CPA-3 derivatives, respectively. After this time, the immobilization rate decreased until to obtain final values of %I prot of 82.8, 74.3 and 64.8% for CPA-1, CPA-2 and CPA-3, respectively (table 24).

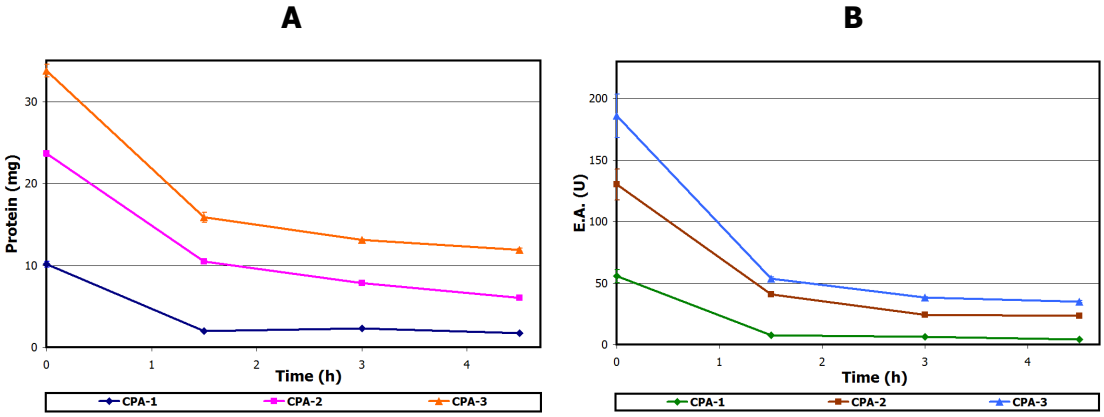


Figure 39. Reaction progress curves for the loss of protein and enzyme activity during the bovine CPA immobilization on glyoxal Sepharose® CL-4B.
A. Protein profile. Protein concentration in supernatants determined by BCA method. **B. Enzymatic activity profile.** The enzymatic activity in supernatants was followed by its ability to hydrolyse $1.0x10^{-4}$ M AAFP synthetic substrate in a final volume of 1 ml, pH 7.5 and 25 ºC (Mock *et al.*, 1996). E.A.: enzymatic activity. U: international Units. Data are means (n=3) ± S.D.

Regarding to the total units of enzymatic activity, the progression curves of the enzyme activity showed similar behaviours to those obtained in terms of protein, achieving percentages of enzyme immobilization in terms of activity (%I EA) of 92.1, 82.2 and 81.2 for CPA-1, CPA-2 and CPA-3 derivatives, respectively (table 24).

In both cases (protein and enzymatic activity units), dependences between percentage of immobilization and initial enzyme loading were observed. Thus, for the lowest enzyme loading used, a lower quantity of enzyme was immobilized in terms of protein and enzyme activity but the corresponding percentages were 1.3 and 1.1 times higher compared to those values obtained with higher enzyme loads employed. Therefore, for higher enzyme loads, higher immobilized enzyme was obtained although the immobilized percentage regarding the initial enzyme used was lower. These results could be related to the presence of steric hindrances and mass transfer limitations at higher enzyme loads. These results are in agreement with the hyperbolic behaviour obtained for the dependence of degree of immobilization vs. enzyme loading (Montes *et al.*, 2006).

Table 24. Immobilization parameters of bovine CPA on glyoxal Sepharose® CL-4B

	$Prot_{initial}$ (mg)	$TU_{initial}$ (U)	V support (ml)	DI prot (mg prot/ml gel)	%I prot	DI EA (units/ml gel)	%EA	Direct DI EA (units/ml gel)	%RFA
CPA-1	10.1 ± 0.4	55.8 ± 5.3	5	1.68 ± 0.07	82.8 ± 4.7	10.3 ± 1.1	92.1 ± 13.0	4.88 ± 0.07	43.7 ± 4.2
CPA-2	23.6 ± 0.4	130.2 ± 12.4	5	3.52 ± 0.08	74.3 ± 2.0	21.4 ± 2.5	82.2 ± 12.4	6.8 ± 0.3	26.1 ± 2.8
CPA-3	33.8 ± 0.8	186.0 ± 17.8	5	4.4 ± 0.2	64.8 ± 2.9	30.2 ± 3.6	81.2 ± 12.3	7.6 ± 0.1	20.5 ± 2.0

Prot: protein. TU: Total units of enzymatic activity (International units). V=volumen. DI prot: degree of enzyme immobilization in terms of protein. %I prot: percentage of enzyme immobilization in terms of protein. DI EA: degree of enzyme immobilization in terms of enzymatic activity. %I EA: percentage of enzyme immobilization in terms of enzymatic activity. %RFA: percentage of retention of functional activity**.**
Data are means (n=3) ± S.D.

The immobilization reaction progress was also monitored by Tris-glycine SDS-PAGE analysis. A decrease in the intensity of the band at approximately 30 kDa (corresponding to the CPA) at different times was visualized in the supernatants of all preparations, indicating the progress of the immobilization reaction. It was also observed the presence of a band corresponding to the enzyme after treating the immobilized derivative with ethanolamine/SDS/β-mercaptoethanol, indicating the release of the protein bound to the support (figure 40).

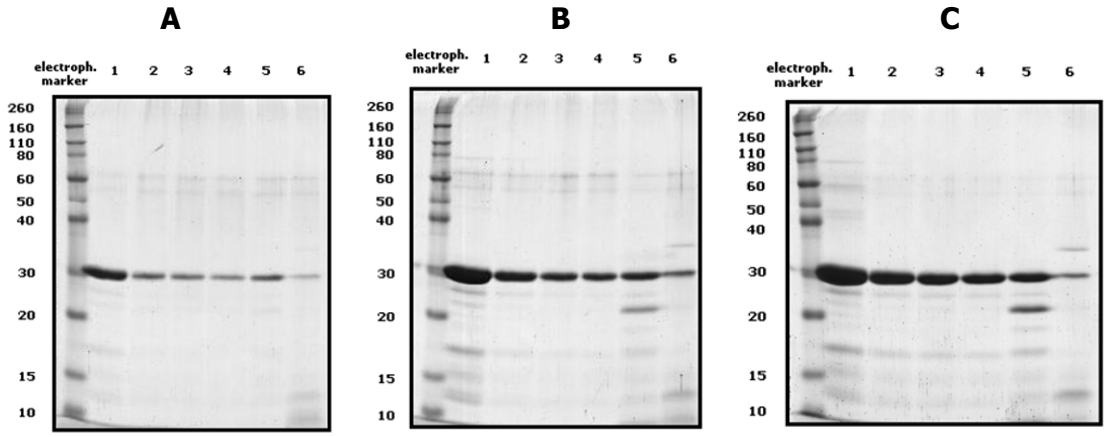


Figure 40. SDS-PAGE of bovine CPA immobilization on glyoxal Sepharose® CL-4B.
A. CPA-1 (1.07 mg CPA/ml gel). Lane 1: initial preparation, lane 2: supernatant at t = 1 1/2 h, lane 3: supernatant at t = 3 h, lane 4: supernatant at t = 4 1/2 h, lane 5: supernatant of the filtrated suspension (at final time), lane 6: CPA-Sepharose derivative. **B.** CPA-2 (3.5 mg CPA/ml gel). Lane 1: initial preparation, lane 2: supernatant at t = 1 1/2 h, lane 3: supernatant at t = 3 h, lane 4: supernatant at t = 4 ½ h, lane 5: supernatant of the filtrated suspension (at final time), lane 6: CPA-Sepharose derivative. **C.** CPA-3 (4.8 mg CPA/ml gel). Lane 1: initial preparation, lane 2: supernatant at t = 1 1/2 h, lane 3: supernatant at t = 3 h, lane 4: supernatant at t = 4 1/2 h, lane 5: supernatant of the filtrated suspension (at final time), lane 6: CPA-Sepharose derivative

The low percentages of retention of functional activity (RFA) obtained mainly for CPA-2 and CPA-3 Sepharose derivatives are also related to the same causes mentioned in the case of percentages of enzyme immobilization that is steric hindrances, conformational changes during the immobilization process and/or the ocurrence of internal and external mass transfer limitations. The sole difference between RFA and percentages of enzyme immobilization is the measurement of enzyme activity, which is directly determined on the immobilized enzyme in the case of RFA and by difference (indirect method) on the supernatant in the case of percentages.

The analysis of the stirring speed on the enzymatic activity in CPA-sepharose derivatives displayed a dependence of the overall reaction rate with respect to this variable, indicating the presence of external diffusional limitations, which were more pronounced in the immobilized derivative with higher enzyme loading (figure 41). The analysis of these factors implies the existence of an unstirred layer of fluid around the particle (Nernst's diffusion layer) where the Fick's law is fulfilled. The thickness of this layer decreases with increasing flow rate of fluid around the particle. Thus, increasing the stirring speed, the external diffusional limitations decreased and a process controlled by diffusion becomes in a kinetically controlled. Another factor which favours these external mass transfer limitations is the low substrate concentration as that used in these experiments (Chávez *et al.*, 2012).

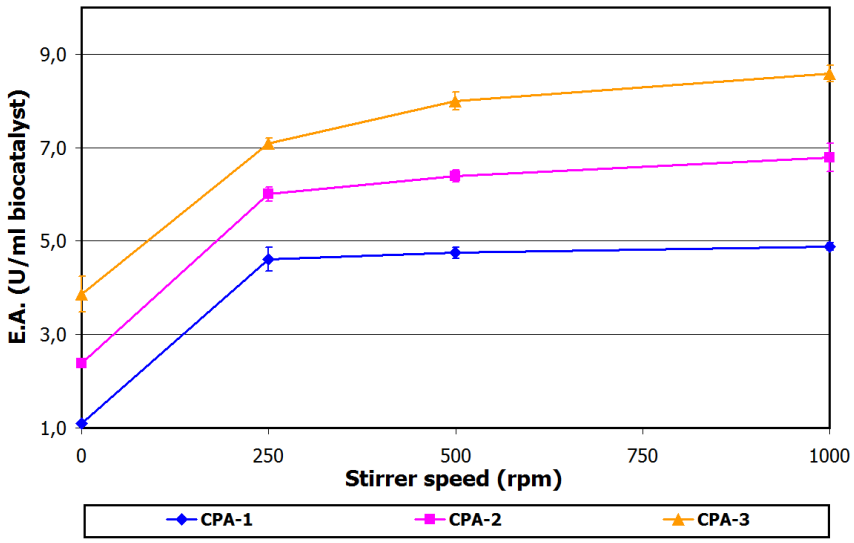


Figure 41. Effect of stirring speed on external mass transfer limitations of bovine CPA–glyoxal Sepharose® CL-4B. The enzymatic activity in the immobilized derivatives was followed by its ability to hydrolyse $1.0x10^{-4}$ M AAFP synthetic substrate in a final volume of 2.5 ml, pH 7.5 and 25 ºC (Mock *et al.*, 1996). E.A.: enzymatic activity. U: international Units. Data are means (n=3) ± S.D.

Based on the results obtained above, the study of internal mass transfer limitations was performed using a stirring speed of 1000 rpm in order to minimize the external diffusional limitations. The effective diffusion coefficient for the substrate was $3.1x10^{-6}$ cm^2/s, calculated from equation 14 (Materials and Methods) and taking into account the parameters described for both, the substrate and the support. The Weis-Prater modulus, also known as the observable Thiele modulus, were 0.58, 0.81 and 0.91 for CPA-1, CPA-2 and CPA-3 sepharose derivatives, respectively. For all CPA-sepharose derivatives, internal transfer limitations were significant, leading to effectiveness factors of 0.7, 0.6 and 0.6 for CPA-1, CPA-2 and CPA-3, respectively (table 25).

These results indicate that the overall reaction rates were controlled by internal mass transfer limitations, which may contribute to the low percentages of RFA obtained specially for CPA-2 and CPA-3 Sepharose derivatives. It is very well described that diffusional limitations are favoured at high concentration of immobilized enzyme. The results obtained confirmed this behaviour because as the enzyme load increased (CPA-2 and CPA-3), a decrease in the effectiveness factor was obtained (0.65 and 0.60, respectively), as well as in the percentages of RFA. Other factors such as low substrate concentration, particle size and support porosity also contribute to internal mass transfer limitations but in this case these parameters were the same for the three derivatives (Chávez *et al.*, 2012).

It is also important to take into account that the effect of reduction of CPA-Sepharose derivatives with sodium borohydride ($NaBH_4$) on the enzymatic activity has been described (Tardiolo *et al.*, 2003). These authors found a decrease in the enzymatic activity, adscribed to the reduction of Zn^{2+} (located at the active site of the enzyme and necessary to perform the catalysis) by $NaBH_4$. This phenomenon was not present in other enzymes (Guisán, J.M., 1988; Fernández-Lafuente, *et al.*, 1995).

Table 25. Estimation of internal mass transfer limitations in bovine CPA-glyoxal Sepharose® CL-4B derivatives

	CPA-1	CPA-2	CPA-3
V_M AAFP (cm^3/gmol)	367.8		
D_{AB} (cm^2/s)	$4.9x10^{-6}$		
D_{eff} (cm^2/s)	$3.1x10^{-6}$		
C_s (mmol/l)	$1.0x10^{-1}$		
MW_s (g/gmol)	365.43		
MW_E (g/gmol)	34800		
γ_{obs} (mmol/h.g_c)	0.276 ± 0.004	0.385 ± 0.017	0.411 ± 0.007
ρ_c (g/ml)	1.06 ± 0.02	1.06 ± 0.03	1.12 ± 0.05
ϕ_{WP}	0.58 ± 0.01	0.81 ± 0.04	0.91 ± 0.01
η	0.75	0.65	0.60

V_M: molar volume of the substrate at normal boiling point, D_{AB}: diffusion coefficient of solute A in solvent B, D_{eff}: effective diffusion coefficient, C_s: substrate concentration, MW_s: molecular weight of the substrate, MW_E: molecular weight of the enzyme, γ_{obs}: observed reaction rate, ρ_c: biocatalyst density, ϕ_{WP}: Weisz-Prater modulus, η: effectiveness factor. Data are means (n=3) ± S.D.

IV.2.3. Immobilization of porcine CPB on glyoxal Sepharose® CL-4B

Reaction progress curve of CPB immobilization on glyoxal Sepharose showed that immobilization rate was faster compared to CPA due to the lower amount of CPB applied in comparison with CPA. In this case, during the first hour of enzyme coupling, both, protein and enzymatic activity decreased rapidly. After this time, the immobilization rate was significantly reduced until final immobilization time of 3 h (figure 42). At this point, percentages of protein immobilization of 80.9, 86.6 y 83.8% for CPB-1, CPB-2 and CPB-3 derivatives were reached respectively, whereas a %I EA of 99.8, 99.7 y 99.0% were achieved in each case (table 26).

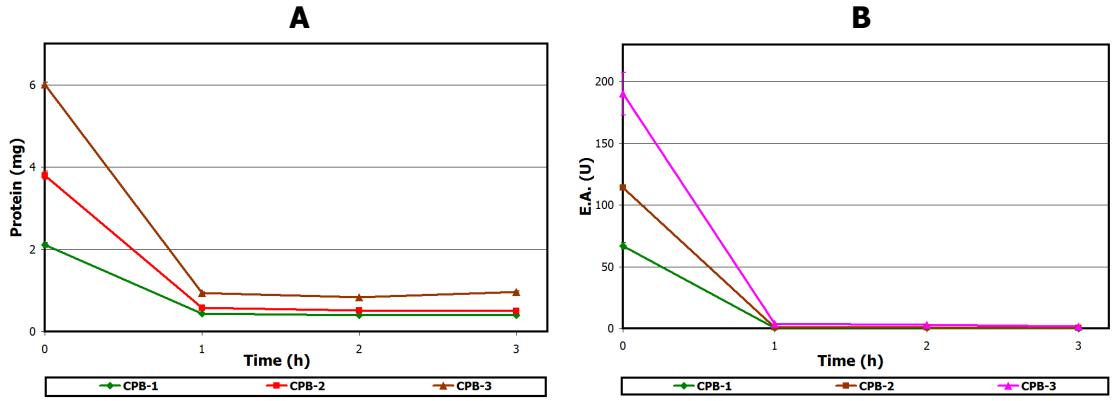


Figure 42. Reaction progress curves for the loss of protein and enzyme activity during porcine CPB immobilization on glyoxal Sepharose® CL-4B.

A. Protein profile. Protein concentration in supernatants determined by BCA method. **B. Enzymatic activity profile.** The enzymatic activity in supernatants was followed by its ability to hydrolyse $1.0x10^{-4}$ M AAFA synthetic substrate in a final volume of 1 ml, pH 7.5 and 25 ºC (Mock and Stanford, 2002). E.A.: enzymatic activity. U: international Units. Data are means (n=3) ± S.D.

Table 26. Immobilization parameters of porcine CPB on glyoxal Sepharose® CL-4B

	$Prot_{initial}$ (mg)	$TU_{initial}$ (U)	V support (ml)	DI prot (mg prot/ml gel)	%I prot	DI EA (units/ml gel)	%EA	Direct DI EA (units/ml gel)	%RFA
CPB-1	2.11 ± 0.06	67.0 ± 2.5	2	0.85 ± 0.03	80.9 ± 3.6	33.4 ± 1.3	99.8 ± 5.4	7.6 ± 0.2	22.7 ± 1.1
CPB-2	3.8 ± 0.1	114.3 ± 0.6	2	1.65 ± 0.05	86.6 ± 3.4	57.0 ± 0.3	99.7 ± 0.8	9.26 ± 0.03	16.2 ± 0.1
CPB-3	6.01 ± 0.05	190.2 ± 17.1	2	2.52 ± 0.03	83.8 ± 1.1	94.2 ± 8.5	99.0 ± 12.6	13.6 ± 0.2	14.3 ± 1.3

Prot: protein. TU: Total units of enzymatic activity (International units). V=volumen. DI prot: degree of enzyme immobilization in terms of protein. %I prot: percentage of enzyme immobilization in terms of protein. DI EA: degree of enzyme immobilization in terms of enzymatic activity. %I EA: percentage of enzyme immobilization in terms of enzymatic activity. %RFA: percentage of retention of functional activity**.** Data are means (n=3) ± S.D.

Analysis of supernatants and immobilized derivatives by Tris-glycine SDS-PAGE also showed the time progress of the immobilization process of CPB on the glyoxal sepharose support (figure 43). It should be noted that both, the initial preparation of the free enzyme as well as the immobilized derivative showed additional bands to the CPB (visualized approximately at 35 kDa), which revealed that these impurities in the original preparation were also immobilized on the support.

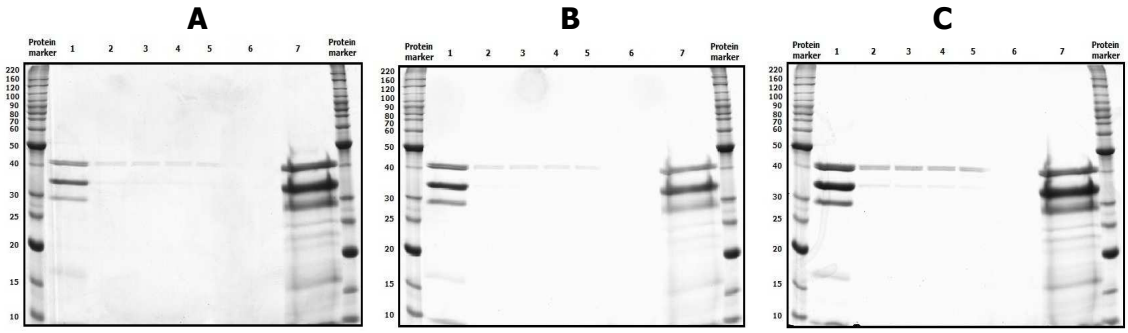


Figure 43. SDS-PAGE of porcine CPB immobilization on glyoxal Sepharose® CL-4B. **A.** CPB-1 (0.9 mg CPB/ml gel). Lane 1: initial preparation, lane 2: supernatant at t = 1 h, lane 3: supernatant at t = 2 h, lane 4: supernatant at t = 3 h, lane 5: supernatant of the filtrated suspension (at final time), lane 6: supernatant of the washing step (at final time), lane 7: CPB-Sepharose derivative. **B.** CPB-2 (1.6 mg CPB/ml gel). Lane 1: initial preparation, lane 2: supernatant at t = 1 h, lane 3: supernatant at t = 2 h, lane 4: supernatant at t = 3 h, lane 5: supernatant of the filtrated suspension (at final time), lane 6: supernatant of the washing step (at final time), lane 7: CPB-Sepharose derivative. **C.** CPB-3 (2.4 mg CPB/ml gel). Lane 1: initial preparation, lane 2: supernatant at t = 1 h, lane 3: supernatant at t = 2 h, lane 4: supernatant at t = 3 h, lane 5: supernatant of the filtrated suspension (at final time), lane 6: supernatant of the washing step (at final time), lane 7: CPB-Sepharose derivative

The effect of stirring speed on the enzymatic activity of the immobilized derivatives (CPB-1 and CPB-2) did not reveal differences in the reaction rate at 250, 500 and 1000 rpm (figure 44). These results indicated that external mass transfer limitations were not rate-limiting in these immobilized derivatives.

In contrast, the highest enzyme loaded derivative (CPB-3) displayed a slight dependence of stirring speed on the overall reaction rate. Therefore, for the subsequent analysis of internal diffusional limitations, a stirring speed of 1000 rpm was used for determining the hydrolysis rate of the synthetic substrate AAFA in all derivatives. The low substrate concentration used as in the case of CPA did not promote external mass transfer limitations, at least in the less loaded derivatives, probably due to the presence of a moderate enzyme activity.

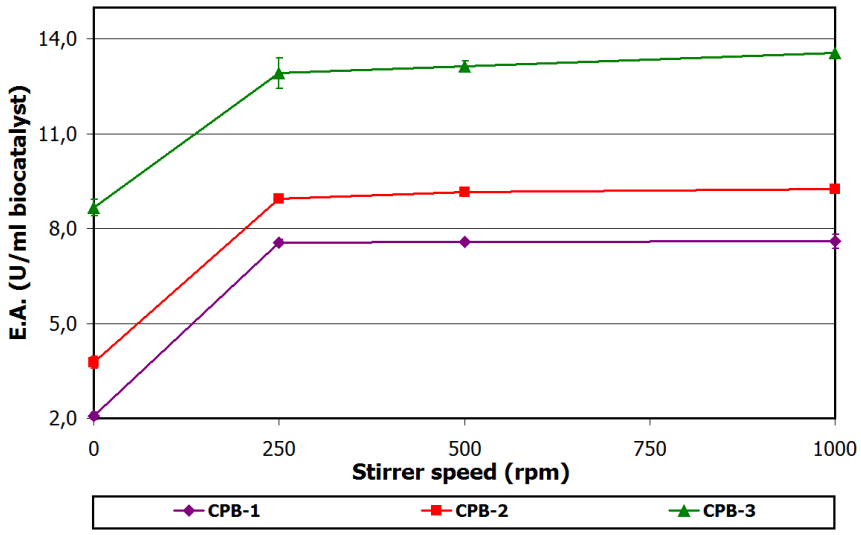


Figure 44. Effect of stirring speed on external mass transfer limitations of porcine CPB–glyoxal Sepharose® CL-4B. The enzymatic activity in the immobilized derivatives was followed by its ability to hydrolyse $1.0x10^{-4}$ M AAFA synthetic substrate in a final volume of 2.5 ml, pH 7.5 and 25 ºC (Mock and Stanford, 2002). E.A.: enzymatic activity. U: international Units. Data are means (n=3) ± S.D.

The analysis of internal mass transfer limitations for the three CPB-sepharose derivatives revealed effectiveness factors of 0.6, 0.5 and 0.4 for CPB-1, CPB-2 and CPB-3 derivatives, respectively (table 27), indicating that in all cases, internal diffusional limitations were the rate-controlling step on the overall reaction rate. These results are in agreement with the low percentages of RFA obtained: 22, 16 and 14% for CPB-1, CPB-2 and CPB-3, in each case.

The effectiveness factor is lower in comparison with CPA, which may be due to the fact that a higher degree of immobilization in terms of enzyme activity was obtained with a lower amount of immobilized protein. This is probably a result of the higher reaactivity of CPB toward its substrate. Therefore, by using the same substrate concentration (0.1 mM) for both enzymes and effective diffusivities in the same order of magnitude, probably the reaction rate for CPB is much higher than the substrate diffusivity flux through the pores in the support, where the enzyme is immobilized, leading to higher values of Weisz-Prater modulus and lower effectiveness factors.

Table 27. Estimation of internal mass transfer limitations in porcine CPB-glyoxal Sepharose® CL-4B derivatives

	CPB-1	CPB-2	CPB-3
V_M AAFA (cm^3/gmol)	387.6		
D_{AB} (cm^2/s)	$4.8x10^{-6}$		
D_{eff} (cm^2/s)	$3.0x10^{-6}$		
C_s (mmol/l)	$1.0x10^{-1}$		
MW_s (g/gmol)	372.81		
MW_E (g/gmol)	34000		
γ_{obs} (mmol/h.g_c)	0.424 ± 0.012	0.521 ± 0.001	0.815 ± 0.009
ρ_c (g/ml)	1.07 ± 0.10	1.07 ± 0.06	0.998 ± 0.002
ϕ_{WP}	0.94 ± 0.03	1.142 ± 0.003	1.67 ± 0.02
η	0.60	0.55	0.45

V_M: molar volume of the substrate at normal boiling point, D_{AB}: diffusion coefficient of solute A in solvent B, D_{eff}: effective diffusion coefficient, C_s: substrate concentration, MW_s: molecular weight of the substrate, MW_E: molecular weight of the enzyme, γ_{obs}: observed reaction rate, ρ_c: biocatalyst density, ϕ_{WP}: Weisz-Prater modulus, η: effectiveness factor. Data are means (n=3) ± S.D.

IV.2.4. Immobilization of papain from *C. papaya* on glyoxal Sepharose® CL-4B

The immobilization of papain on glyoxal sepharose supports was carried out for 3 h. During the first hour of enzyme-support interaction, a high coupling rate was observed, both in terms of protein and enzyme activity, after which the process rate decreased over the next two hours of immobilization (figure 45).

Degrees of immobilization in terms of protein (DI prot) of 2.7, 5.9 and 7.9 mg papain/ml gel for PAP-1, PAP-2 and PAP-3, respectively were obtained. These values represent 96.7, 91.3 and 89.5% percentages of protein immobilization in each case (table 28). Degrees of immobilization in terms of enzymatic activity (DI EA) were 8.9, 21.1 and 26.7 U/ml for PAP-1, PAP-2 and PAP-3, respectively.

The percentages of RFA for these immobilized derivatives were low as in the previous studied cases. For this enzyme, 23.8, 24.5 and 30.2% of enzyme activity were recovered in the PAP-1, PAP-2 and PAP-3 derivatives, respectively. It should be noted that in contrast with the two earlier enzymes studied, the lowest value of %RFA was obtained for the derivative with lower enzyme loading. Taking into account that the percentages of immobilization in terms of enzyme activity measured by difference in the supernatant were very high, it is probably that in the low %RFA obtained (based on the direct determination of enzymatic activity in the immobilized derivative) could contribute other factors, such as mass transfer limitations, enzyme denaturation, among others.

Monitoring of papain immobilization was also followed by Tris-glycine SDS-PAGE analysis, where as in the previous cases, results showed a decrease in the band intensities of the supernatants as a function of time, and finally the protein recovery from the support (figure 46).

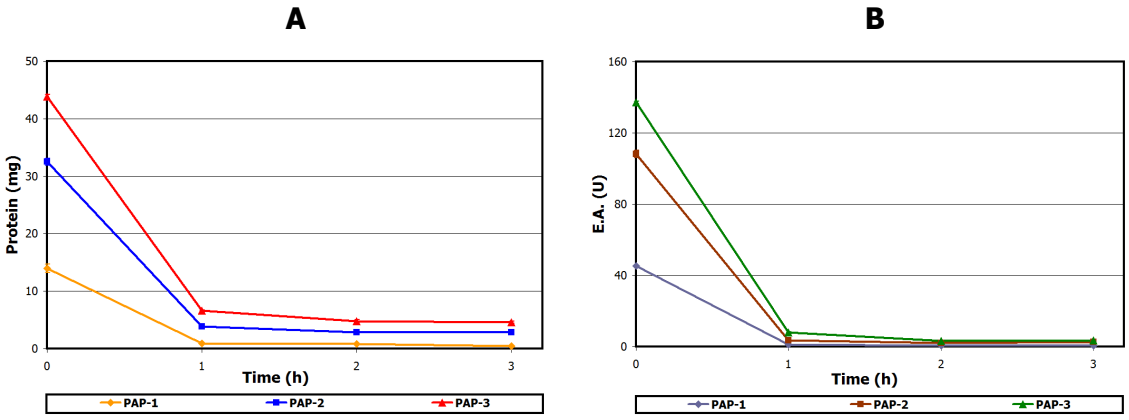


Figure 45. Reaction progress curves for the loss of protein and enzyme activity during papain immobilization on glyoxal sepharose® CL-4B.
A. Protein profile. Protein concentration in supernatants determined by BCA method. **B. Enzymatic activity profile.** The enzymatic activity in supernatants was followed by its ability to hydrolyse $4.0x10^{-4}$ M PFLNA synthetic substrate in a final volume of 1 ml, pH 6.5 and 25 ºC (Filippova *et al.*, 1984). E.A.: enzymatic activity. U: international Units. Data are means (n=3) ± S.D.

Table 28. Immobilization parameters of papain on glyoxal Sepharose® CL-4B

	$Prot_{initial}$ (mg)	$TU_{initial}$ (U)	V support (ml)	DI prot (mg prot/ml gel)	%I prot	DI EA (units/ml gel)	%EA	Direct DI EA (units/ml gel)	%RFA
PAP-1	14.0 ± 0.8	45.2 ± 0.5	5	2.7 ± 0.2	96.7 ± 7.5	8.9 ± 0.1	98.6 ± 1.4	2.1 ± 0.1	23.8 ± 0.7
PAP-2	32.5 ± 0.6	108.3 ± 2.1	5	5.9 ± 0.1	91.3 ± 2.5	21.1 ± 0.4	97.6 ± 2.7	7.5 ± 0.1	34.5 ± 0.8
PAP-3	43.8 ± 0.3	137.2 ± 0.8	5	7.9 ± 0.1	89.5 ± 1.1	26.7 ± 0.2	97.4 ± 0.8	8.3 ± 0.3	30.2 ± 1.0

Prot: protein. TU: Total units of enzymatic activity (International units). V=volumen. DI prot: degree of enzyme immobilization in terms of protein. %I prot: percentage of enzyme immobilization in terms of protein. DI EA: degree of enzyme immobilization in terms of enzymatic activity. %I EA: percentage of enzyme immobilization in terms of enzymatic activity. %RFA: percentage of retention of functional activity**.** Data are means (n=3) ± S.D.

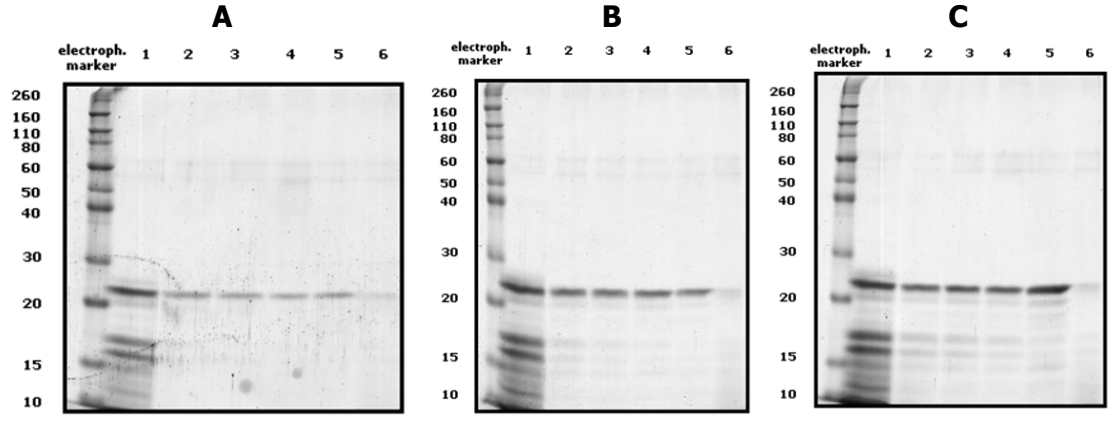


Figure 46. SDS-PAGE of papain immobilization on glyoxal Sepharose® CL-4B.
A. PAP-1 (2.3 mg papain/ml gel). Lane 1: initial preparation, lane 2: supernatant at t = 1 h, lane 3: supernatant at t = 2 h, lane 4: supernatant at t = 3 h, lane 5: supernatant of the filtrated suspension (at final time), lane 6: papain-Sepharose derivative. **B.** PAP-2 (5.5 mg papain/ml gel). Lane 1: initial preparation, lane 2: supernatant at t = 1 h, lane 3: supernatant at t = 2 h, lane 4: supernatant at t = 3 h, lane 5: supernatant of the filtrated suspension (at final time), lane 6: papain-Sepharose derivative. **C.** PAP-3 (7.2 mg papain/ml gel). Lane 1: initial preparation, lane 2: supernatant at t = 1 h, lane 3: supernatant at t = 2 h, lane 4: supernatant at t = 3 h, lane 5: supernatant of the filtrated suspension (at final time), lane 6: papain-Sepharose derivative

Analysis of the effect of stirring speed on the overall reaction rate revealed a dependency between these two variables, which becomes more pronounced as the enzyme load is increased (figure 47). For the immobilized derivative with the lowest enzyme loading, an atypical behaviour was observed in the analysis of the effect of speed stirring on enzyme activity. Bearing in mind that the experimental conditions such as substrate concentration and stirring speed were the same for all derivatives, is likely as discussed above, this behaviour can be adscribed to other factors. This result is in agreement with the lowest RFA value obtained for this immobilized derivative (PAP-1) when compared to the immobilized derivatives with higher enzyme loadings.

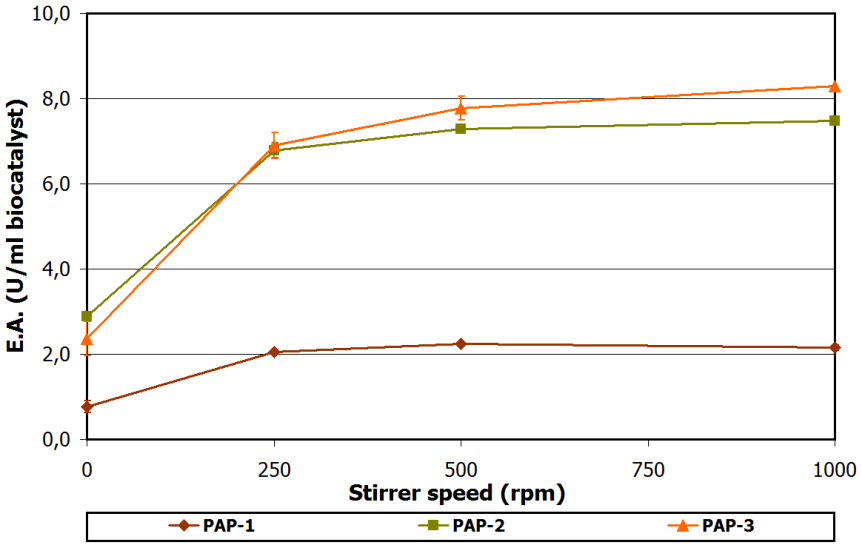


Figure 47. Effect of stirring speed on external mass transfer limitations of papain–glyoxal Sepharose® CL-4B. The enzymatic activity in the immobilized derivatives was followed by its ability to hydrolyse $4.0x10^{-4}$ M PFLNA synthetic substrate in a final volume of 2.5 ml, pH 6.5 and 25 ºC (Filippova *et al.*, 1984). E.A.: enzymatic activity. U: international Units. Data are means (n=3) ± S.D.

The Weisz-Prater module obtained in papain-sepharose derivatives were 0.85, 2.95 and 3.27 for PAP-1, PAP-2 and PAP-3, respectively, leading to effectiveness factors of 0.6, 0.3 and 0.3 (table 29). These low values of efficiency factor, particularly for the two immobilized derivatives with higher enzyme loadings, indicate that the overall reaction rate is a process controlled by mass internal limitations, favoured by high concentrations of immobilized enzyme and low substrate concentration, among others.

Table 29. Estimation of internal mass transfer limitations in papain-glyoxal Sepharose® CL-4B derivatives

	PAP-1	PAP-2	PAP-3
V_M PFLNA ($cm^3/gmol$)	584.3		
D_{AB} (cm^2/s)	$3.7x10^{-6}$		
D_{eff} (cm^2/s)	$2.4x10^{-6}$		
C_s (mmol/l)	$4.0x10^{-2}$		
MW_s (g/gmol)	509.56		
MW_E (g/gmol)	23000		
γ_{obs} (mmol/h.g_c)	0.123 ± 0.004	0.398 ± 0.005	0.491 ± 0.017
ρ_c (g/ml)	1.05 ± 0.02	1.13 ± 0.04	1.01 ± 0.03
ϕ_{WP}	0.85 ± 0.02	2.95 ± 0.04	3.27 ± 0.11
η	0.60	0.30	0.27

V_M: molar volume of the substrate at normal boiling point, D_{AB}: diffusion coefficient of solute A in solvent B, D_{eff}: effective diffusion coefficient, C_s: substrate concentration, MW_s: molecular weight of the substrate, MW_E: molecular weight of the enzyme, γ_{obs}: observed reaction rate, ρ_c: biocatalyst density, ϕ_{WP}: Weisz-Prater modulus, η: effectiveness factor. Data are means (n=3) ± S.D.

IV.2.5. Immobilization of bovine trypsin on glyoxal Sepharose® CL-4B

The immobilization of trypsin on glyoxal Sepharose was performed for 2 hours, although for practical purposes, the enzyme-support contact of 1 hour was enough to achieve good results of immobilization. At the end of immobilization time, %I prot of 79.0, 69.9 and 65.4% were achieved for TRYP-1, TRYP-2 and TRYP-3, respectively. On the other hand, the percentages of immobilization in terms of enzymatic activity were found to be 100% for all the immobilized derivatives. The differences observed between the protein and activity could be explained by the differential method used (determination of enzymatic activity by difference in the supernatant) and the heterogeneity and/or denaturation of the enzyme sample in the supernatant. If some non immobilized enzyme is denatured in the supernatant, a result of 100% immobilized activity will be obtained. On the other hand, the high %RFA obtained (directly measured activity) and the absence of internal diffusional limitations (see below) does not completely support this explanation. However, the %RFA is 10 to 20 percent lower than the % of enzyme activity measured in the supernatant, suggesting that some enzyme denaturation could take place. Monitoring results of trypsin immobilization by SDS-PAGE are shown in figure 48.

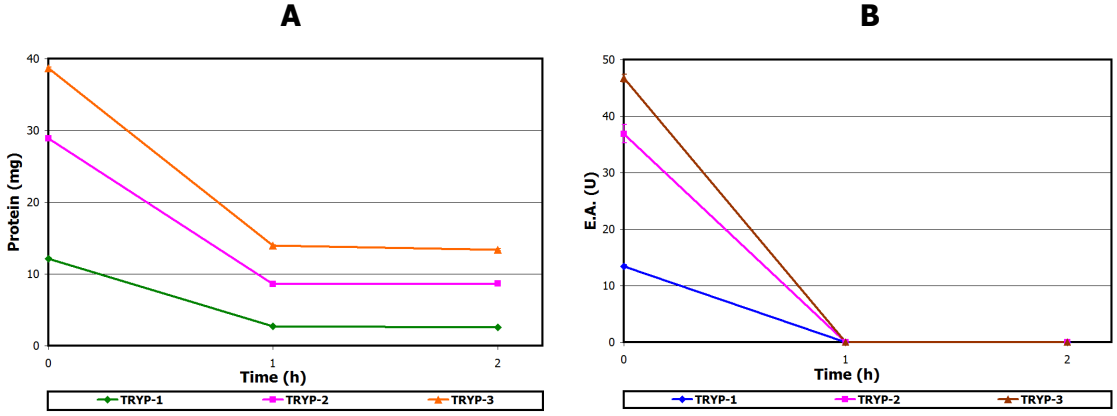


Figure 48. Reaction progress curves for the loss of protein and enzyme activity during bovine trypsin immobilization on glyoxal sepharose® CL-4B.
A. Protein profile. Protein concentration in supernatants determined by BCA method. **B. Enzymatic activity profile.** The enzymatic activity in supernatants was followed by its ability to hydrolyse $1.0x10^{-3}$ M BAPNA synthetic substrate in a final volume of 1 ml, pH 8.0 and 25 ºC (Erlanger *et al.*, 1961). E.A.: enzymatic activity. U: international Units. Data are means (n=3) ± S.D.

Table 30. Bovine trypsin immobilization parameters on glyoxal Sepharose® CL-4B

	$Prot_{initial}$ (mg)	$TU_{initial}$ (U)	V support (ml)	DI prot (mg prot/ml gel)	%I prot	DI EA (units/ml gel)	%EA	Direct DI EA (units/ml gel)	%RFA
TRYP-1	12.1 ± 0.1	13.5 ± 0.3	5	1.92 ± 0.02	79.0 ± 1.2	2.7 ± 0.1	100.0 ± 2.9	2.43 ± 0.01	90.1 ± 1.9
TRYP-2	28.93 ± 0.05	36.9 ± 1.7	5	4.05 ± 0.04	69.9 ± 0.6	7.4 ± 0.3	100.0 ± 6.3	6.1 ± 0.1	82.8 ± 3.9
TRYP-3	38.7 ± 0.1	46.7 ± 0.7	5	5.06 ± 0.04	65.4 ± 0.6	9.3 ± 0.1	100.0 ± 2.1	7.6 ± 0.1	81.1 ± 1.7

Prot: protein. TU: Total units of enzymatic activity (International units). V=volumen. DI prot: degree of enzyme immobilization in terms of protein. %I prot: percentage of enzyme immobilization in terms of protein. DI EA: degree of enzyme immobilization in terms of enzymatic activity. %I EA: percentage of enzyme immobilization in terms of enzymatic activity. %RFA: percentage of retention of functional activity**.**
Data are means (n=3) ± S.D.

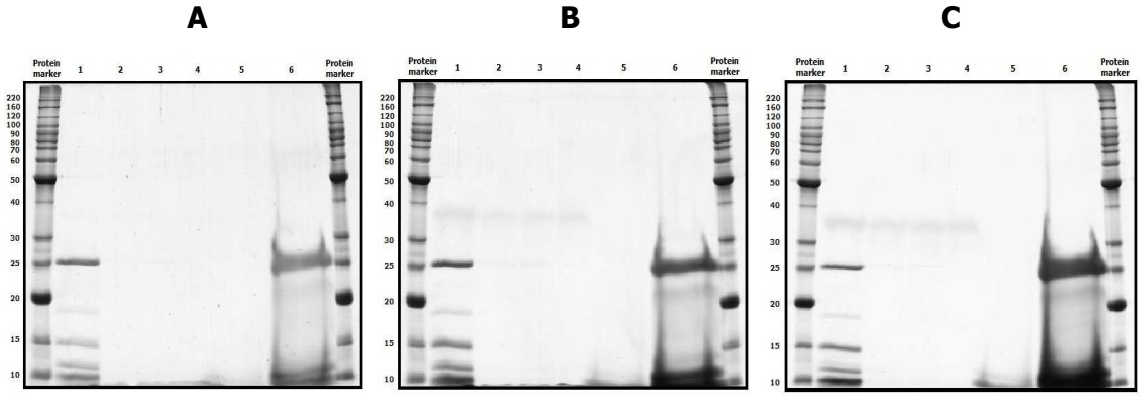


Figure 49. SDS-PAGE of bovine trypsin immobilization on glyoxal Sepharose® CL-4B.
A. TRYP-1 (1.4 mg trypsin/ml gel). Lane 1: initial preparation, lane 2: supernatant at t = 1 h, lane 3: supernatant at t = 2 h, lane 4: supernatant of the filtrated suspension (at final time), lane 5: supernatant of the washing step (at final time), lane 6: trypsin-Sepharose derivative. **B.** TRYP-2 (4.0 mg trypsin/ml gel). Lane 1: initial preparation, lane 2: supernatant at t = 1 h, lane 3: supernatant at t = 2 h, lane 4: supernatant of the filtrated suspension (at final time), lane 5: supernatant of the washing step (at final time), lane 6: trypsin-Sepharose derivative. **C.** TRYP-3 (4.6 mg trypsin/ml gel). Lane 1: initial preparation, lane 2: supernatant at t = 1 h, lane 3: supernatant at t = 2 h, lane 4: supernatant of the filtrated suspension (at final time), lane 5: supernatant of the washing step (at final time), lane 6: trypsin-Sepharose derivative

The study of external mass transfer limitations in the trypsin-sepharose derivatives showed that at least for TRYP-1 and TRYP-2, the overall reaction rate was kinetically controlled, while TRYP-3 showed a slight dependence of the stirring speed on the global rate.

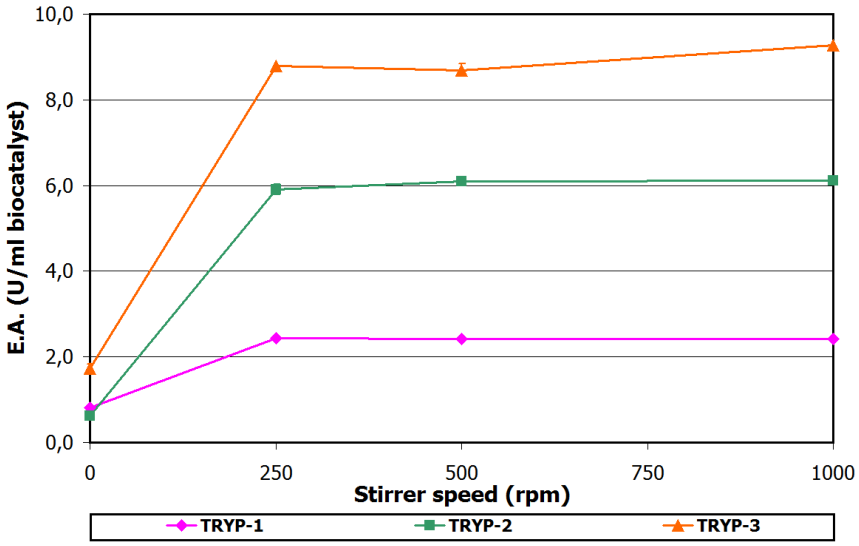


Figure 50. Effect of stirring speed on external mass transfer limitations of bovine trypsin–glyoxal Sepharose® CL-4B. The enzymatic activity in the supernatants was followed by its ability to hydrolyse $1.0x10^{-3}$ M BAPNA synthetic substrate in a final volume of 2.5 ml, pH 8.0 and 25 ºC (Erlanger *et al.*, 1961). E.A.: enzymatic activity. U: international Units. Data are means (n=3) ± S.D.

In all trypsin-Sepharose derivatives analyzed, an effectiveness factor of 1 was obtained, indicating that the overall reaction rate of the process is free of mass transfer limitations and the process is only controlled by the chemical reaction. This result is in agreement with the percentage of RFA obtained for each derivative: 90.1, 82.8 and 81.1% for TRYP-1, TRYP-2 and TRYP-3, respectively (table 30). In this case, the external constraints, as in all cases, were alleviate by a proper system agitation, while the mass internal limitations were diminished by different factors, such as the higher substrate concentration used (one order of magnitude higher than the concentration used for the rest of substrates), the low reaction rate of BAPNA hydrolysis by trypsin, given by the low sensibility of this substrate, compared with other synthetic substrates (see BRENDA (BRaunschweig ENzyme DAtabase), http://www.brenda-enzymes.org).

Table 31. Estimation of internal mass transfer limitations in bovine trypsin-glyoxal Sepharose® CL-4B derivatives

	TRYP-1	TRYP-2	TRYP-3
V_M BAPNA (cm^3/gmol)	443.3		
D_{AB} (cm^2/s)	$4.4x10^{-6}$		
D_{eff} (cm^2/s)	$2.8x10^{-6}$		
C_s (mmol/l)	1.0		
MW_s (g/gmol)	434.88		
MW_E (g/gmol)	24000		
γ_{obs} (mmol/h.g_c)	0.140 ± 0.001	0.324 ± 0.005	0.436 ± 0.005
ρ_c (g/ml)	1.04 ± 0.03	1.13 ± 0.03	1.04 ± 0.04
ϕ_{WP}	0.0324 ± 0.0002	0.082 ± 0.001	0.101 ± 0.001
η	1.00	1.00	1.00

V_M: molar volume of the substrate at normal boiling point, D_{AB}: diffusion coefficient of solute A in solvent B, D_{eff}: effective diffusion coefficient, C_s: substrate concentration, MW_s: molecular weight of the substrate, MW_E: molecular weight of the enzyme, γ_{obs}: observed reaction rate, ρ_c: biocatalyst density, ϕ_{WP}: Weisz-Prater modulus, η: effectiveness factor. Data are means (n=3) ± S.D.

IV.2.6. Immobilization of subtilisin from *B. licheniformis* on glyoxal Sepharose® CL-4B

Immobilization of subtilisin on glyoxal Sepharose supports was carried out for 4 hours and at room temperature, unlike the immobilization of other proteases. Immobilization was initially performed at 4°C, obtaining a percentage of enzyme immobilization in terms of protein about 20% (data not shown). For this reason, immobilization at room temperature was carried out in order to increase the immobilization rate. During the first hour of coupling, high immobilization rates were observed, leading to the immobilization of 27, 26 and 26% of subtilisin for SUBT-1, SUBT-2, and SUBT-3 derivatives, respectively (figure 51). After this time, immobilization rates decreased and at a final immobilization time of 4 h percentages of immobilization in terms of protein of 53.8, 32.8 and 29.2% were obtained for SUBT-1, SUBT SUBT-2 and-3 (table 32).

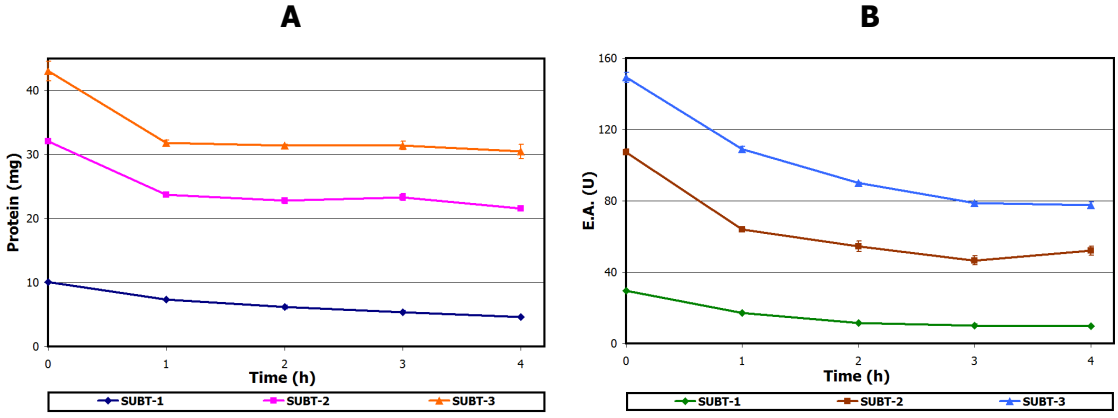


Figure 51. Reaction progress curves for the loss of protein and enzyme activity durin subtilisin immobilization on glyoxal Sepharose® CL-4B.
A. Protein profile. Protein concentration in supernatants determined by BCA method. **B. Enzymatic activity profile.** The enzymatic activity in supernatant was followed by its ability to hydrolyse $4.0x10^{-4}$ M GGLPNA synthetic substrate in a final volume of 1 ml, pH 8.5 and 25 ºC (Lyublinskaya *et al.*, 1974). E.A.: enzymatic activity. U: international Units. Data are means (n=3) ± S.D.

Table 32. Subtilisin immobilization parameters on glyoxal Sepharose® CL-4B (T_{immob} = 25ºC)

	$Prot_{initial}$ (mg)	$TU_{initial}$ (U)	V support (ml)	DI prot (mg prot/ml gel)	%I prot	DI EA (units/ml gel)	%EA	Direct DI EA (units/ml gel)	%RFA
SUBT-1	10.1 ± 0.1	29.7 ± 0.3	5	1.08 ± 0.03	53.8 ± 1.8	4.02 ± 0.06	67.5 ± 1.2	1.7 ± 0.1	27.8 ± 1.6
SUBT-2	32.1 ± 0.1	107.4 ± 1.6	5	2.11 ± 0.06	32.8 ± 0.9	11.1 ± 0.6	51.5 ± 2.8	5.1 ± 0.2	23.8 ± 0.9
SUBT-3	43.1 ± 1.5	149.2 ± 2.8	5	2.5 ± 0.4	29.2 ± 4.5	14.3 ± 0.7	48.0 ± 2.4	6.6 ± 0.1	22.0 ± 0.5

Prot: protein. TU: Total units of enzymatic activity (International units). V=volumen. DI prot: degree of enzyme immobilization in terms of protein. %I prot: percentage of enzyme immobilization in terms of protein. DI EA: degree of enzyme immobilization in terms of enzymatic activity. %I EA: percentage of enzyme immobilization in terms of enzymatic activity. %RFA: percentage of retention of functional activity**.**
Data are means (n=3) ± S.D.

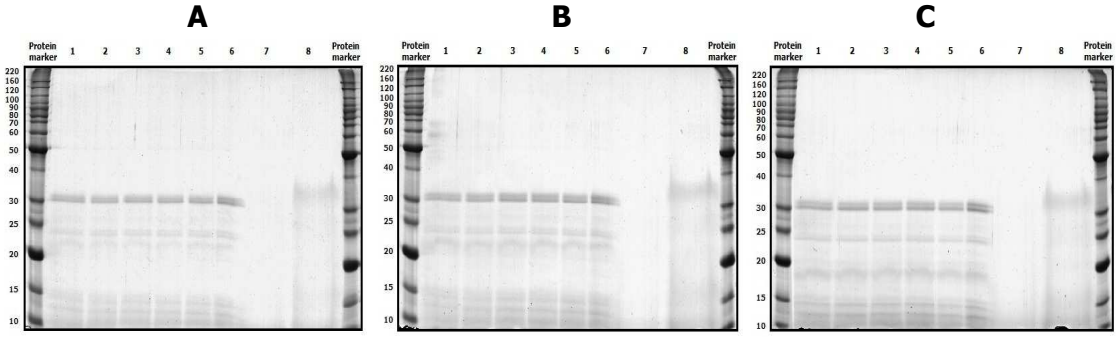


Figure 52. SDS-PAGE of subtilisin immobilization on glyoxal Sepharose® CL-4B.
A. SUBT-1 (0.3 mg subtilisin/ml gel). Lane 1: initial preparation, lane 2: supernatant at t = 1 h, lane 3: supernatant at t = 2 h, lane 4: supernatant at t = 3 h, lane 5: supernatant at t = 4 h, lane 6: supernatant of the filtrated suspension (at final time), lane 7: supernatant of the washing step (at final time), lane 8: subtilisin-Sepharose derivative. **B.** SUBT-2 (1.2 mg subtilisin/ml gel). Lane 1: initial preparation, lane 2: supernatant at t = 1 h, lane 3: supernatant at t = 2 h, lane 4: supernatant at t = 3 h, lane 5: supernatant at t = 4 h, lane 6: supernatant of the filtrated suspension (at final time), lane 7: supernatant of the washing step (at final time), lane 8: subtilisin-Sepharose derivative. **C.** SUBT-3 (2.1 mg subtilisin/ml gel). Lane 1: initial preparation, lane 2: supernatant at t = 1 h, lane 3: supernatant at t = 2 h, lane 4: supernatant at t = 3 h, lane 5: supernatant at t = 4 h, lane 6: supernatant of the filtrated suspension (at final time), lane 7: supernatant of the washing step (at final time), lane 8: subtilisin-Sepharose derivative

Percentages of RFA obtained in the three immobilized derivatives were 27.8, 23.8 and 22.0% for SUBT-1, SUB-SUB-T-2 and-3, respectively. Analysis of the immobilization process by SDS-PAGE confirmed the moderate to low affinity of subtilisin coupling on Sepharose support (figure 52).

Regarding the study of the external mass transfer limitations, a slight effect of the stirring speed on the overall reaction rate was only observed in the SUBT-3 derivative (highest enzyme loading). The external diffusional limitations in this derivative were alleviated by increasing the stirring speed to 1000 rpm.

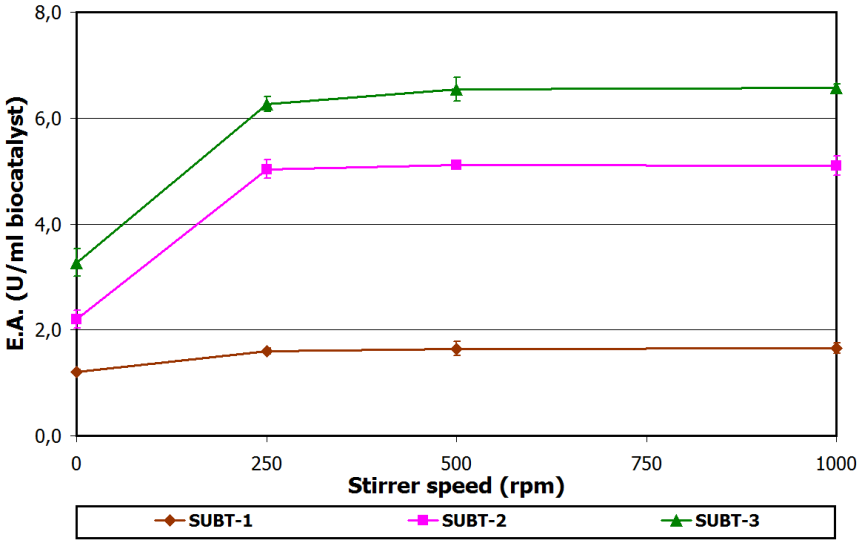


Figure 53. Effect of stirring speed on external mass transfer limitations of subtilisin–glyoxal Sepharose® CL-4B. The enzymatic activity in the immobilized derivatives was followed by its ability to hydrolyse $4.0x10^{-4}$ M GGLPNA synthetic substrate in a final volume of 2.5 ml, pH 8.5 and 25 ºC (Lyublinskaya *et al.*, 1974). E.A.: enzymatic activity. U: international Units. Data are means (n=3) ± S.D.

The Weisz-Prater modules obtained for the immobilized subtilisin derivatives were 0.62, 1.92 and 2.47 for SUBT-1, SUB-SUB-T-2 and-3, respectively, leading to effectiveness factors of 0.7, 0.4 and 0.3 (table 33). Thus, in all cases, the overall reaction rate of the process was controlled by internal diffusional limitations.

Table 33. Estimation of internal mass transfer limitations in subtilisin-glyoxal Sepharose® CL-4B derivatives(T_{immob} = 25ºC)

	SUBT-1	SUBT-2	SUBT-3
V_M GGLPNA (cm^3/gmol)	538.2		
D_{AB} (cm^2/s)	$3.9x10^{-6}$		
D_{eff} (cm^2/s)	$2.5x10^{-6}$		
C_s (mmol/l)	$4.0x10^{-2}$		
MW_s (g/gmol)	499.52		
MW_E (g/gmol)	27290		
γ_{obs} (mmol/h.g_c)	0.096 ± 0.002	0.290 ± 0.010	0.377 ± 0.004
ρ_c (g/ml)	1.04 ± 0.04	1.06 ± 0.03	1.04 ± 0.03
ϕ_{WP}	0.62 ± 0.01	1.92 ± 0.07	2.47 ± 0.03
η	0.75	0.40	0.30

V_M: molar volume of the substrate at normal boiling point, D_{AB}: diffusion coefficient of solute A in solvent B, D_{eff}: effective diffusion coefficient, C_s: substrate concentration, MW_s: molecular weight of the substrate, MW_E: molecular weight of the enzyme, γ_{obs}: observed reaction rate, ρ_c: biocatalyst density, ϕ_{WP}: Weisz-Prater modulus, η: effectiveness factor. Data are means (n=3) ± S.D.

IV.2.7. Immobilization of porcine pepsin on NHS activated Sepharose™ 4 fast flow

Reaction progress curves of pepsin immobilization on NHS activated Sepharose support showed that during the first hour of coupling all the amount of initial protein was immobilized on the support, although the coupling step was extended to 6 hours. At the end of the immobilization process, both %I prot and %I EA were 100% for all the immobilized derivatives. SDS-PAGE analysis of the immobilization process confirmed the fast coupling of the enzyme on this support, and furthermore the release of the protein bound to the Sepharose matrix, besides other components not seen in the initial preparation (figure 55). Therefore, the presence of some autoproteolysis can not be discarded, particularly for enzymes with wide specificities such as pepsin (Hamuro *et al.*, 2008). Nevertheless, it is important to highlight the apparent high efficiency of the immobilization reaction, measured by difference in the supernatant, in

comparison with the value directly obtained in the immobilized derivative. These differences could be adscribed to other factors such as the inactivation of enzyme in the supernatant, probably by autoproteolysis and overestimation of the immobilized enzyme, and/or other factors related to diffusional limitations of the substrate in the immobilized derivative.

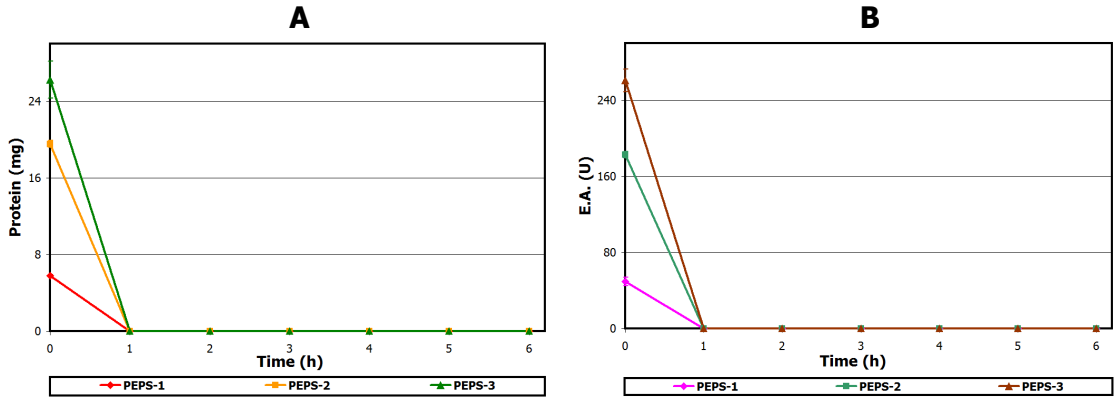


Figure 54. Reaction progress curves for the loss of protein and enzyme activity during pepsin immobilization on NHS activated Sepharose™ 4 fast flow.
A. Protein profile. Protein concentration in supernatants determined by BCA method. **B. Enzymatic activity profile.** The enzymatic activity in supernatants was followed by its ability to hydrolyse $2.0x10^{-4}$ M LSPNPLAL synthetic substrate in a final volume of 1 ml, pH 4.4 and 25 ºC (Martín *et al.*, 1980). E.A.: enzymatic activity. U: international Units. Data are means (n=3) ± S.D.

Table 34. Porcine pepsin immobilization parameters on NHS activated Sepharose™ 4 fast flow

	$Prot_{initial}$ (mg)	$TU_{initial}$ (U)	V support (ml)	DI prot (mg prot/ml gel)	%I prot	DI EA (units/ml gel)	%EA	Direct DI EA (units/ml gel)	%RFA
PEPS-1	5.8 ± 0.1	49.4 ± 4.3	5	1.15 ± 0.03	100.0 ± 3.5	9.9 ± 0.9	100.0 ± 12.4	1.5 ± 0.1	14.8 ± 1.7
PEPS-2	19.5 ± 0.4	182.8 ± 3.1	5	3.91 ± 0.07	100.0 ± 2.6	36.6 ± 0.6	100.0 ± 2.4	1.7 ± 0.2	4.6 ± 0.5
PEPS-3	26.3 ± 1.9	261.0 ± 11.9	5	5.3 ± 0.4	100.0 ± 10.5	52.2 ± 2.4	100.0 ± 6.5	1.82 ± 0.01	3.5 ± 0.2

Prot: protein. TU: Total units of enzymatic activity (International units). V=volumen. DI prot: degree of enzyme immobilization in terms of protein. %I prot: percentage of enzyme immobilization in terms of protein. DI EA: degree of enzyme immobilization in terms of enzymatic activity. %I EA: percentage of enzyme immobilization in terms of enzymatic activity. %RFA: percentage of retention of functional activity**.** Data are means (n=3) ± S.D.

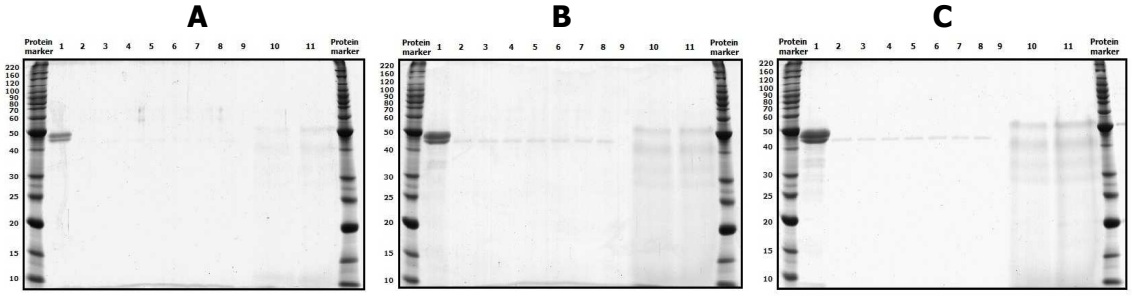


Figure 55. SDS-PAGE of porcine pepsin immobilization on NHS activated Sepharose™ 4 fast flow.
A. PEPS-1 (X mg pepsin/ml gel). Lane 1: initial preparation, lane 2: supernatant at t = 1 h, lane 3: supernatant at t = 2 h, lane 4: supernatant at t = 3 h, lane 5: supernatant at t = 4 h, lane 6: supernatant at t = 5 h, lane 7: supernatant at t = 6 h, lane 8: supernatant of the filtrated suspension (before the ethanolamine blocking step), lane 9: supernatant of the filtrated suspension (after the ethanolamine blocking step), lane 10: pepsin-sepharose derivative before treatment with ethanolamine, lane 11: final pepsin-Sepharose derivative. **B.** PEPS-2 (X mg pepsin/ml gel). Lane 1: initial preparation, lane 2: supernatant at t = 1 h, lane 3: supernatant at t = 2 h, lane 4: supernatant at t = 3 h, lane 5: supernatant at t = 4 h, lane 6: supernatant at t = 5 h, lane 7: supernatant at t = 6 h, lane 8: supernatant of the filtrated suspension (before the ethanolamine blocking step), lane 9: supernatant of the filtrated suspension (after the ethanolamine blocking step), lane 10: pepsin-sepharose derivative before treatment with ethanolamine, lane 11: pepsin-Sepharose derivative. **C.** PEPS-3 (X mg pepsin/ml gel). Lane 1: initial preparation, lane 2: supernatant at t = 1 h, lane 3: supernatant at t = 2 h, lane 4: supernatant at t = 3 h, lane 5: supernatant at t = 4 h, lane 6: supernatant at t = 5 h, lane 7: supernatant at t = 6 h, lane 8: supernatant of the filtrated suspension (before the ethanolamine blocking step), lane 9: supernatant of the filtrated suspension (after the ethanolamine blocking step), lane 10: pepsin-sepharose derivative before treatment with ethanolamine, lane 11: pepsin-Sepharose derivative

The analysis of external diffusional limitations in all the pepsin-Sepharose derivatives evidenced a remarkable dependence of the stirring speed on the overall reaction rate. In this case, unlike the other immobilized enzymes studied above, a stabilization of the global process rate was not reached. Therefore, it is difficult to ensure that the external diffusional limitations were minimized at 1000 rpm, and consequently, there is no guarantee they were completely removed at this stirring state. In this case, the higher molecular weight of the substrate (hexapeptide) and its lower concentration in comparison with the rest of the substrates used against other immobilized proteases are responsible for external mass transfer limitations in the range of the stirring speed analyzed.

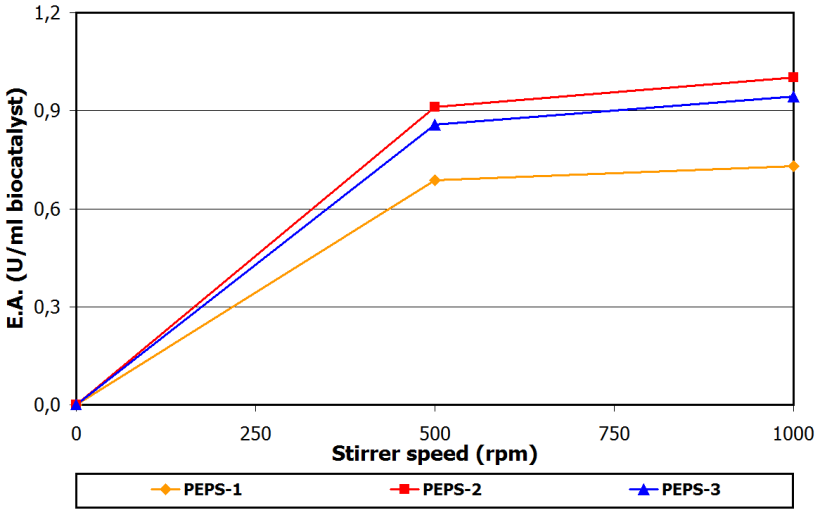


Figure 56. Effect of stirring speed on external mass transfer limitations of porcine pepsin– NHS activated Sepharose™ 4 fast flow.

The enzymatic activity in the immobilized derivatives was followed by its ability to hydrolyse $2.0x10^{-4}$ M LSPNPLAL synthetic substrate in a final volume of 2.5 ml, pH 4.4 and 25 ºC (Martín *et al.*, 1980). E.A.: enzymatic activity. U: international Units.

An effectiveness factor of 0.4, obtained by interpolating the values of the Weisz-Prater modulus in figure 28, was determined for the PEPS-1, PEPS-2 and PEPS-3 derivatives (table 35). As in most of the immobilized proteases studied in this work, the overall reaction rate was controlled by internal mass transfer limitations. In this case, the major contribution to these diffusional limitations is probably related to the fact that on this activated Sepharose matrix is possible the presence of a large amount of immobilized enzyme in terms of protein, such as the indirect method indicated, but most of it is inactive, as %RFA revealed. It should be noted that in spite of having Sepharose matrices with the same pore size, in this NHS activated Sepharose matrix, a larger sparcer arm is anchored, probably promoting internal mass transfer limitations.

Table 35. Estimation of internal mass transfer limitations in porcine pepsin-NHS activated Sepharose™ 4 fast flow derivatives

	PEPS-1	PEPS-2	PEPS-3
V_M S* (cm^3/gmol)	868.3		
D_{AB} (cm^2/s)	$2.9x10^{-6}$		
D_{eff} (cm^2/s)	$1.9x10^{-6}$		
C_s (mmol/l)	$2.0x10^{-2}$		
MW_s (g/gmol)	835.88		
MW_E (g/gmol)	35000		
γ_{obs} (mmol/h.g_c)	0.085 ± 0.005	0.097 ± 0.008	0.1085 ± 0.0003
ρ_c (g/ml)	1.02 ± 0.03	1.05 ± 0.05	1.00 ± 0.03
ϕ_{WP}	1.46 ± 0.08	1.70 ± 0.14	1.817 ± 0.005
η	0.45	0.40	0.40

V_M: molar volume of the substrate at normal boiling point, D_{AB}: diffusion coefficient of solute A in solvent B, D_{eff}: effective diffusion coefficient, C_s: substrate concentration, MW_s: molecular weight of the substrate, MW_E: molecular weight of the enzyme, γ_{obs}: observed reaction rate, ρ_c: biocatalyst density, ϕ_{WP}: Weisz-Prater modulus, η: effectiveness factor. Data are means (n=3) ± S.D.

Mass transfer limitations were relevant in all target protease-Sepharose derivatives, with the exception of trypsin, limiting their catalytic function. External diffusional limitations were in most cases minimized by a proper stirring of the reaction mixture, whereas the internal diffusional limitations were largely determined by the length of the spacer arm, a high enzyme loading and the immobilized enzyme, in terms of protein and/or enzyme activity, the characteristics of the substrate in terms of high sensitivity, molecular weight and concentration (only in the case of trypsin the substrate concentration is one order of magnitude higher).

It is important to highlight that the inhibitors selected for their further identification by the proteomic method Intensity Fading MALDI-TOF MS and their purification are probably tight-binding inhibitors. These inhibitors strongly interact with their target proteases with K_i values in the range of enzyme concentrations used in the assays. Therefore, as both methods are based on the association and dissociation of the inhibitor, immobilization degrees must be low in order to ensure a mild interaction and to avoid drastic elutions which can damage inhibitor molecules (Chávez *et al.*, 2012).

IV.3. Identification and analysis of protease inhibitors by proteomic methods

IV.3.1. Identification of protease inhibitors by Intensity Fading MALDI-TOF MS

The presence of inhibitory activity against CPA, CPB, pepsin, papain, trypsin and subtilisin was evaluated using a medium-throughput proteomic method of molecular recognition called Intensity Fading (IF) MALDI TOF MS. The selection of extracts was based on their inhibitory efficiencies in terms of specific inhibitory activity, IC_{50} value and the abundance of the sources.

It should be noted that this technique was initially established to detect CPA and trypsin inhibitors by direct and indirect IF MALDI-TOF MS (Villanueva *et al.*, 2003; Yanes *et al.*, 2005, 2007). During this research work, this technique was extended to other target proteases such as porcine carboxypeptidase B (CPB), aspartic and cysteine proteases like porcine pepsin and papain from *C. papaya*, respectively, and a serine protease from the subtilisin family (subtilisin from *B. licheniformis*). These proteases as well as bovine CPA and bovine trypsin were used to identify in the selected extracts interacting molecules with this immobilized enzymes.

The extracts submitted to IF MALDI-TOF MS were *N. versicolor*, *P. homomalla* and *S. helianthus* for CPA; *N. versicolor*, *C. muricatus* and *P. homomalla* for CPB; *B. granulifera*, *H. carunculata* and *D. listerianum* for pepsin; *P. homomalla*, *P. physalis* and *C. muricatus* for papain; *S. helianthus*, *N. versicolor*, *C. muricatus* and *N. peloronta* for trypsin; and *L. isodyctialis*, *I. badionotus* and *C. muricatus* for subtilisin.

Since marine extracts are biological samples with high salt content, they were initially desalted and concentrated using ZipTip®C_4 pipette tips, prior to MALDI-TOF analysis. The desalting process was made in order to minimize signal suppression effects in MALDI-MS analysis (Yanes *et al.*, 2005). This treatment, not only allowed the removal of salts, but also decreased the complexity of extracts before being used as controls in MALDI-TOF MS.

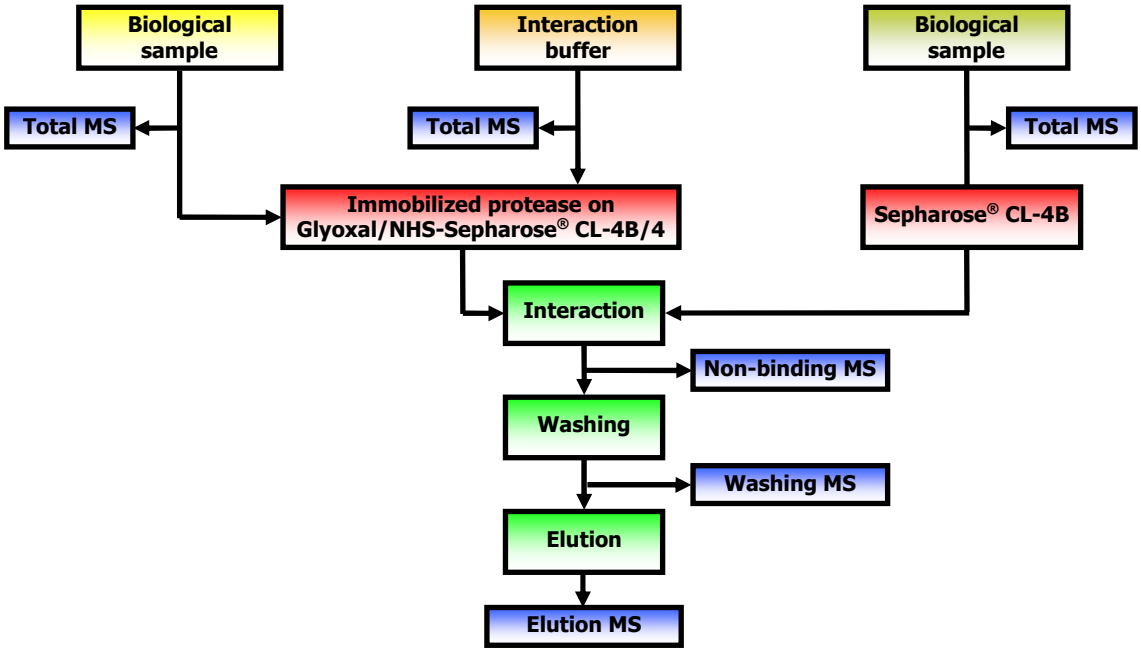


Figure 57. Flowchart process for the identification of protease inhibitors by IF MALDI-TOF MS.

Duplicate IF MALDI-TOF MS analyses of all selected extracts were performed by using an indirect method based on an affinity step with the target proteases immobilized, as described by Yanes *et. al.*, 2007 and shown in figure 17. After optimizing mass spectra of the control extracts, experimental conditions were set in order to achieve retention of inhibitor molecules on the affinity matrices, as well as their elution, followed at each step by MALDI-TOF MS analyses (5 replicates of each desalted fraction). IF MALDI-TOF MS results are only shown for one selected extract against each enzyme.

IV.3.1.1. IF MALDI-TOF MS of a control mixture (PCI and insulin) against CPA-glyoxal Sepharose® CL-4B

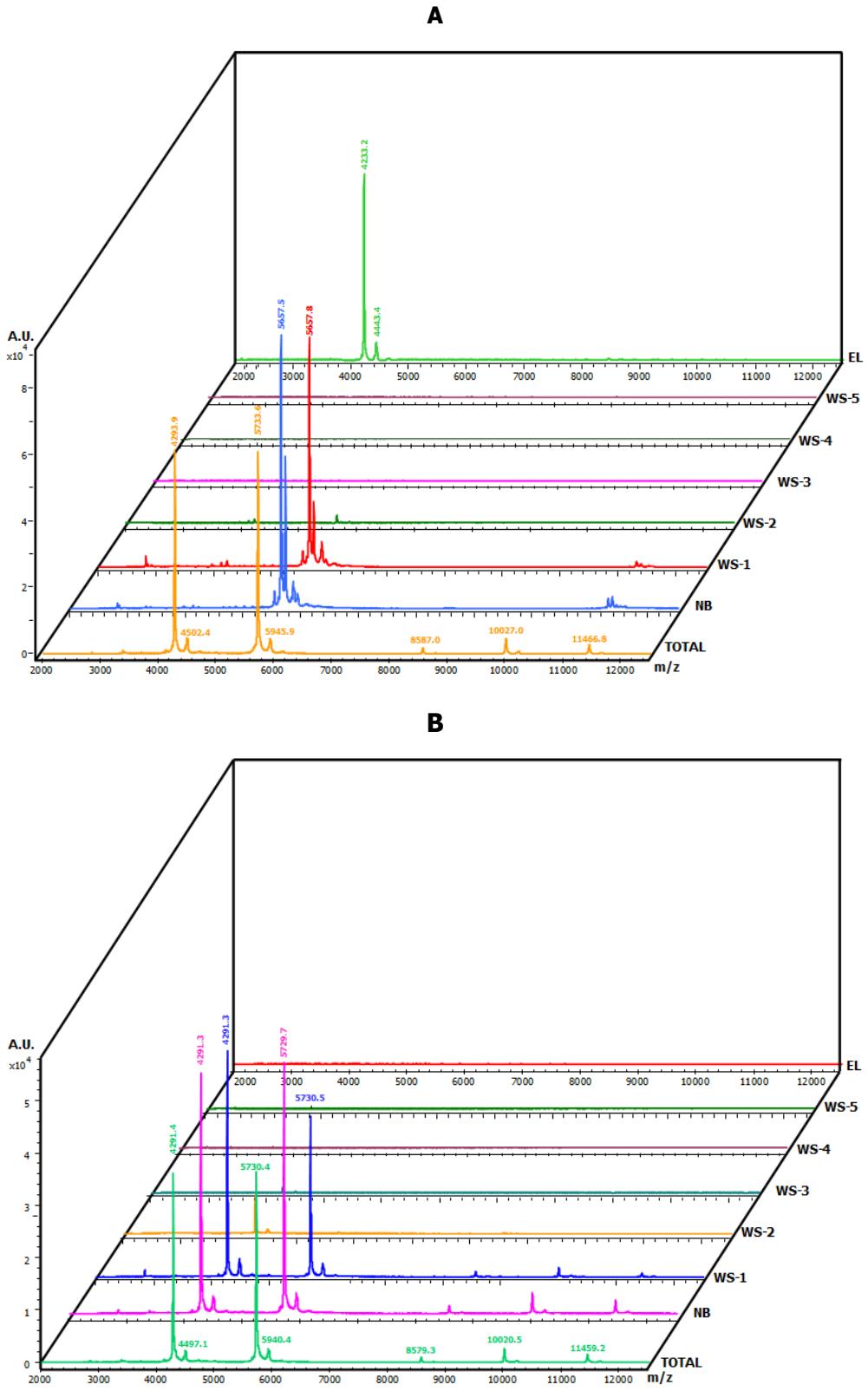


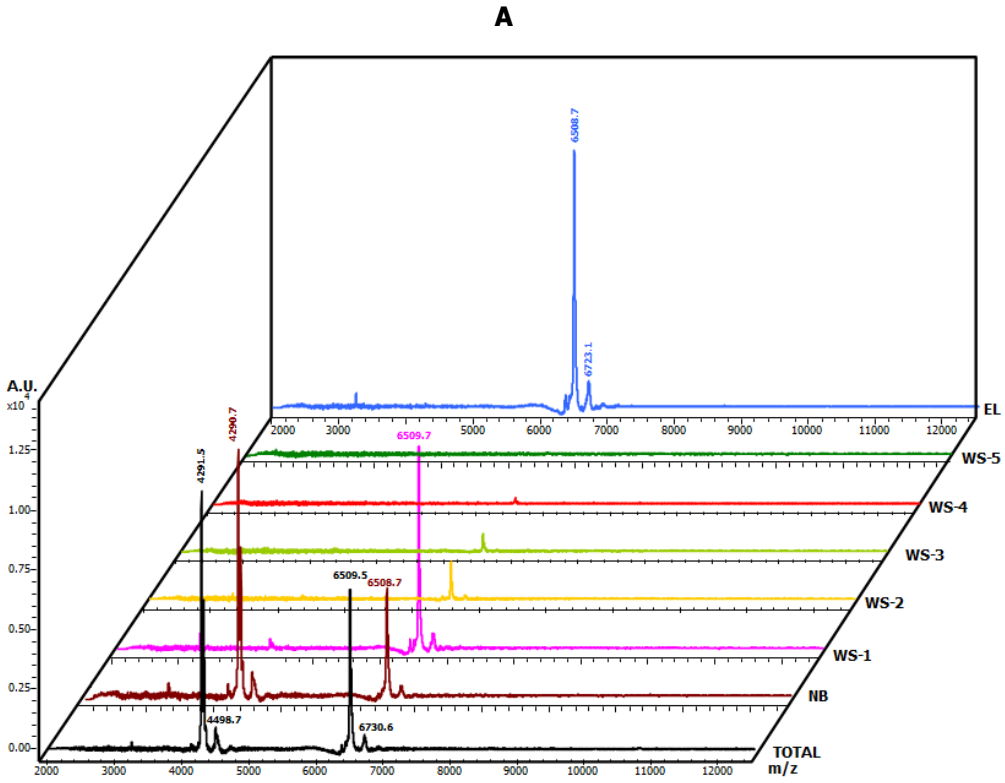
Figure 58. Intensity Fading MALDI-TOF MS of control mixture (PCI and Insulin) against CPA-glyoxal Sepharose® CL-4B (A) and Sepharose® CL-4B (B). TOTAL: PCI and Insuline in 20 mM Tris-HCl buffer pH 7.5, 500 mM NaCl. **NB:** non-binding molecules, interaction at pH 7.5, 10 min. **WS-1 to WS-5:** washing at pH 7.5. **EL:** elution with 0.5% v/v TFA, 10 min. Interaction assays were carried out in duplicate at room temperature. The experimental conditions of MALDI-TOF MS are described in section III.7.2**.**

Figure 58A shows the results of IF MALDI-TOF MS obtained with a control mixture of insulin and PCI (carboxypeptidase inhibitor, Hass *et al.*, 1975) against CPA-Sepharose. MALDI mass spectrum of the total extract displayed clearly the presence of two peaks of 4293.9 and 5733.6 Da for PCI and insulin, respectively. In a second step of the experiment (NB: non-binding molecules), the peak corresponding to PCI completely disappeared from the mass spectrum after the sample was placed to interact with the immobilized CPA. During the subsequent washing step (five times), insulin in excess or bound by weak and non-specific interactions with the matrix was removed. The latest wash did not indicate any signal. Finally, dissociation of the protease-inhibitor complex was performed and the previously disappeared PCI signal was recovered in the MALDI spectrum (elution). Spectrum of elution fraction revealed the presence of a single peak of 4233.2 Da which was not detected in the total extract. This mass difference is explained by the removal of the valine residue at the C-terminal tail of PCI by the action of CPA (earlier described in PCI purification processes using immobilized CPA, Hass *et al.*, 1975). Molecular mass obtained corresponds to the loss of the terminal valine in PCI and the addition of a hydroxyl group at the C-terminal tail of this molecule.

IF MALDI-TOF MS analysis of the same mixture on Sepharose (in absence of the immobilized target protease) confirmed that the interaction between PCI and the immobilized enzyme was exclusively given by the protease-inhibitor interaction and not by non-specific interactions between the inhibitor and Sepharose matrix (figure 58B). Disappearance of signals in the mass spectrum of the non-retained fraction was not observed, and therefore, at the final elution step, any peak was not recovered in the MALDI spectrum.

This result indicates the effectiveness and selectivity of the affinity matrix along with the experimental conditions used for retention and elution of binding molecules present in complex biological samples.

IV.3.1.2. IF MALDI-TOF MS of a control mixture (aprotinin and PCI) against trypsin-glyoxal Sepharose® CL-4B



B

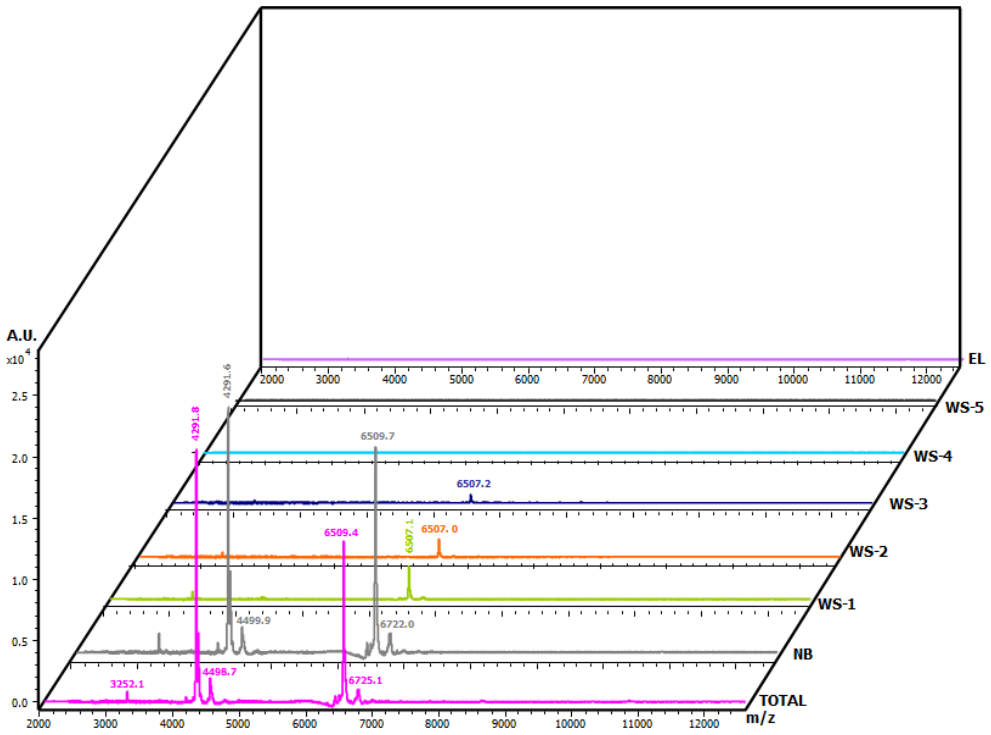


Figure 59. Intensity Fading MALDI-TOF MS of control mixture (PCI and aprotinin) against trypsin-glyoxal Sepharose® CL-4B (A) and Sepharose® CL-4B (B). TOTAL: PCI and aprotinin in 20 mM Tris-HCl buffer pH 8.0, 150 mM NaCl, 20 mM $CaCl_2$. **NB:** non-binding molecules, interaction at pH 8.0, room temperature, 10 min. **WS-1 to WS-5:** washing at pH 8.0. **EL:** elution with 0.5% v/v TFA, 10 min. Interaction assays were carried out in duplicate at room temperature. The experimental conditions of MALDI-TOF MS are described in section III.7.2**.**

IF MALDI-TOF MS of a mixture of PCI and aprotinin (trypsin inhibitor, Huber *et al*., 1970) against trypsin-Sepharose are shown in figure 59A. MALDI mass spectrum of the total extract clearly showed the presence of two peaks of 4291.5 and 6509.5 Da for PCI and aprotinin, respectively. In a second step of the assay, the sample was placed to interact with the immobilized enzyme. Slight evidence of decay of the signal corresponding to aprotinin was observed in the mass spectrum of the non-retained fraction, which initially indicates that there was no strong interaction between inhibitor and protease. In addition, during the subsequent washing steps, aprotinin in excess or weakly interacting with the immobilized protease was removed. However, a MALDI signal corresponding to aprotinin was recovered in the elution stage. This result confirms molecular recognition of trypsin by this inhibitor.

IF MALDI-TOF MS analysis of the same sample using only Sepharose is shown in figure 59B. In this experiment was confirmed that the interaction between aprotinin and the immobilized enzyme was exclusively given by the protease-inhibitor interaction and not by non-specific interactions between the inhibitor and the Sepharose matrix. Mass spectrum of the non-retained fraction in this experiment displayed a similar profile of that obtained with the immobilized enzyme. However, there was no evidence of any signal in the mass spectrum of the elution fraction, as expected. Both experiments confirm IF MALDI-TOF MS efficiency for molecular recognition of trypsin inhibitors.

IV.3.1.3. IF MALDI-TOF MS of *N. versicolor* crude extract against CPA-glyoxal Sepharose® CL-4B

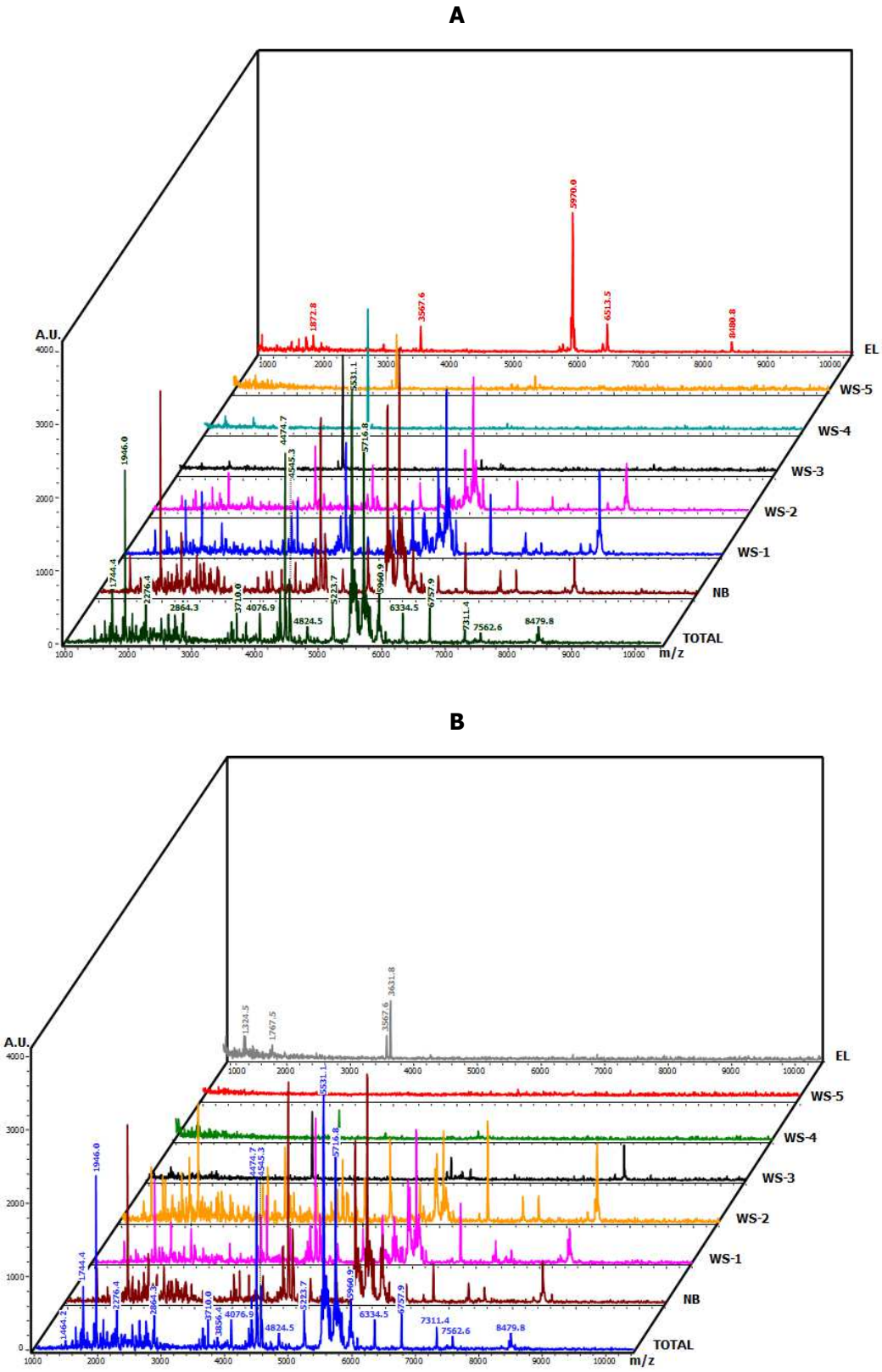


Figure 60. Intensity Fading MALDI-TOF MS of *N. versicolor* crude extract against CPA-glyoxal Sepharose® CL-4B (A) and Sepharose® CL-4B (B). TOTAL: crude extract of *N. versicolor* in 20 mM Tris-HCl buffer pH 7.5, 500 mM NaCl. **NB:** non-binding molecules, interaction at pH 7.5, 10 min. **WS-1 to WS-5:** washing at pH 7.5. **EL:** elution with 0.5% v/v TFA, 10 min. Interaction assays were carried out in duplicate at room temperature. The experimental conditions of MALDI-TOF MS are described in section III.7.2**.**

IF MALDI-TOF MS results of *N. versicolor* crude extract with CPA-Sepharose are shown in figure 60A. In this case, mass spectrum of the crude extract is quite heterogeneous, as expected for these samples. Therefore, it was difficult to clearly observe disappearance or decay of a signal in the mass spectrum of the sample after interaction with the immobilized enzyme. Due to the complexity of the extract, effects of signal overlapping and/or suppression in the MALDI analysis are generated (Yanes *et al.*, 2005). However, after washing steps, MALDI spectrum was simplified obtaining a mass spectrum without any signal. After elution step, recovery of a major peak at approximately 6 kDa containing 3 MALDI signals (5946.5, 5961.3 and 5970.0) and minor peaks of 1872.8, 3567.6, 6513.5 and 8480.8 Da were observed, due to dissociation of complexes between the target protease and binding molecules.

It is important to note that a detailed view of the IF MALDI-TOF MS experiment of *N. versicolor* crude extract and CPA-glyoxal Sepharose within the range of 5750-6100 Da allowed to observe the fading and subsequent recovery of two signals at 5946.3 and 5960.9 Da (figure 61). In addition, an intense peak at 5970.0 Da not previously observed either in the total mass spectrum or in the subsequent fractions was visualized in the elution fraction.

Control analysis of *N. versicolor* crude extract on Sepharose without any immobilized enzyme displayed the presence of two molecules of 3567.6 and a 3631.8 Da, indicating non-specific interactions between these molecules and the support (figure 60B). These signals must be discarded from the mass spectrum of the elution fraction obtained with CPA-Sepharose.

These results confirm the presence of at least one molecule in *N. versicolor* extract able to interact with CPA, corresponding to the initially observed CPA inhibitory activity in this extract by kinetic assays and corroborated by measuring CPA inhibitory activity in the elution fraction analysed by IF MALDI-TOF MS (CPA inhibitory activity = 255.3 ± 21.0 mU/ml). Other minor peaks in the same range of molecular mass may be related to isoforms, which has been described for inhibitors isolated from marine invertebrates (Delfín *et al.*, 2006, Gonzalez *et al.*, 2007a).

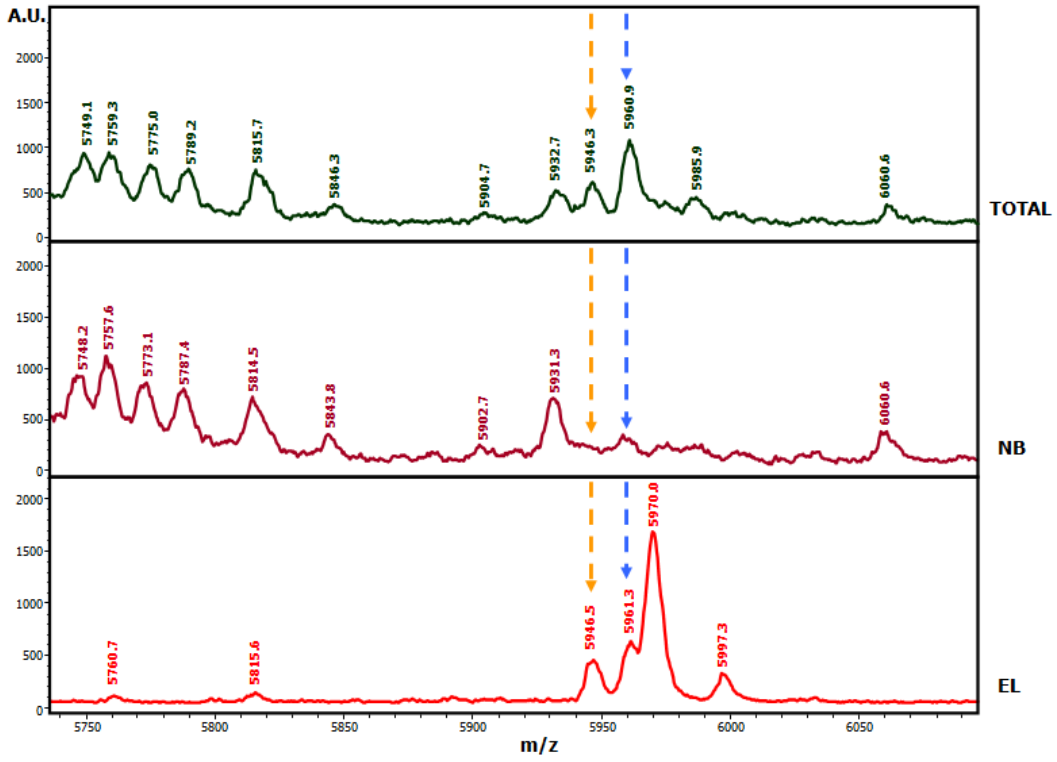


Figure 61. Detailed view of the Intensity Fading MALDI-TOF MS of *N. versicolor* crude extract against CPA-glyoxal Sepharose® CL-4B. TOTAL: crude extract of *N. versicolor* in 20 mM Tris-HCl buffer pH 7.5, 500 mM NaCl. **NB:** non-binding molecules, interaction at pH 7.5, 10 min. **EL:** elution with 0.5% v/v TFA, 10 min. Interaction assays were carried out in duplicate at room temperature. The experimental conditions of MALDI-TOF MS are described in section III.7.2.

IV.3.1.4. IF MALDI-TOF MS of *N. versicolor* crude extract against CPB-glyoxal Sepharose® CL-4B

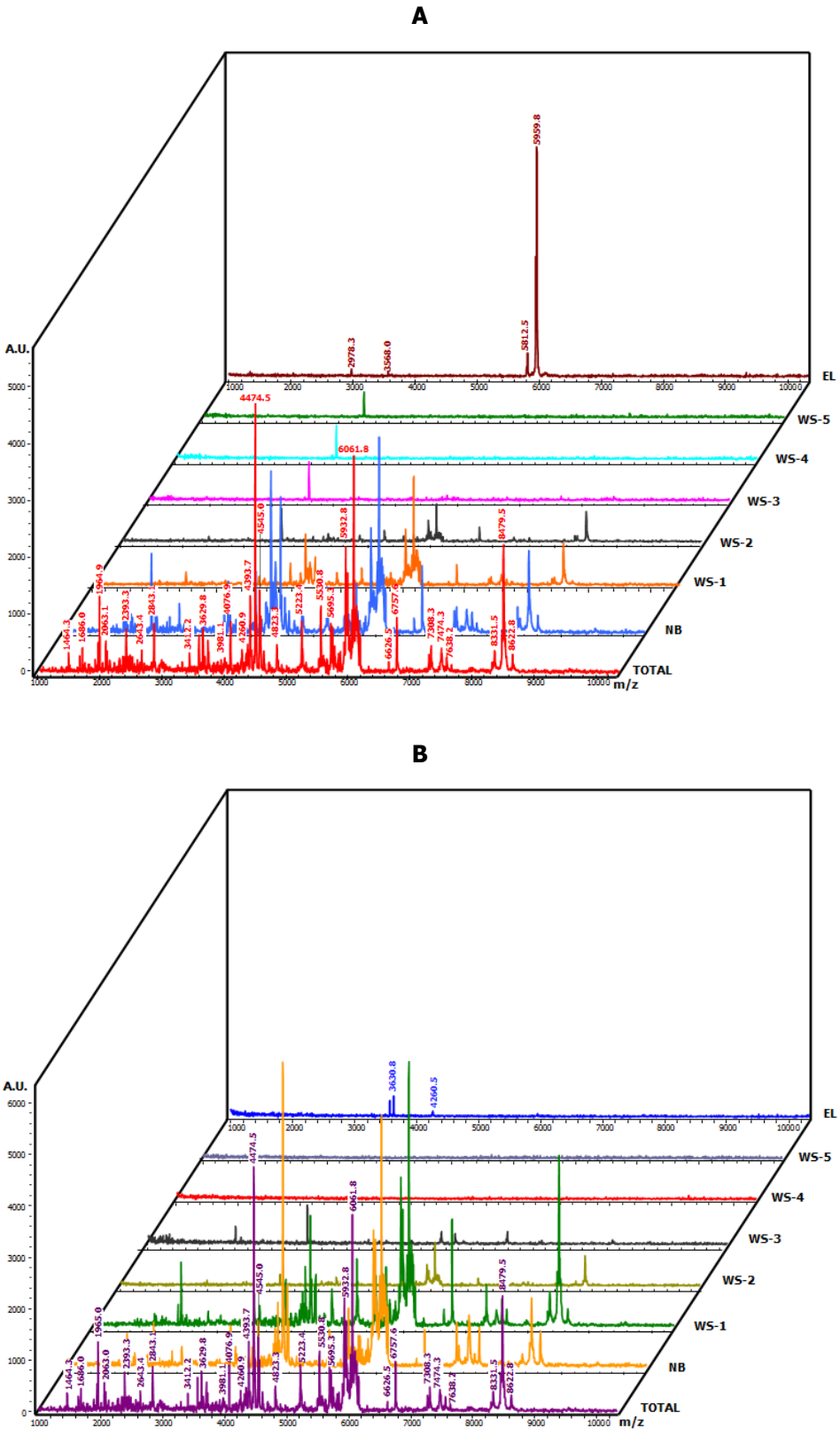


Figure 62. Intensity Fading MALDI-TOF MS of *N. versicolor* crude extract against CPB-glyoxal Sepharose® CL-4B. (A) and Sepharose® CL-4B (B). TOTAL: crude extract of *N. versicolor* in 20 mM Tris-HCl buffer pH 7.5, 100 mM NaCl. **NB:** non-binding molecules, interaction at pH 7.5, 10 min. **WS-1 to WS-5:** washing at pH 7.5. **EL:** elution with 0.5% v/v TFA, 10 min. Interaction assays were carried out in duplicate at room temperature. The experimental conditions of MALDI-TOF MS are described in section III.7.2.

Figure 62A shows the results with *N. versicolor* crude extract on CPB-Sepharose. As previously noted, total extract was found to be quite heterogeneous, making difficult to clearly visualize the decay or disappearance of any signal in the mass spectrum of the non-retained fraction, after interaction with the affinity matrix. However, during the washing steps, mass spectrum of the sample was simplified to a final spectrum with a single peak at 3568.0 Da.

After dissociation of the complexes between the enzyme and binding molecules in the extract, the recovery of two major peaks not completely resolved at 5945.7 and m/z 5959.8 Da and other minor peaks at m/z 2978.3, 3568.0 and 5812.5 Da were observed. In addition, a detailed view of the IF MALDI-TOF MS experiment of *N. versicolor* crude extract and CPB-glyoxal Sepharose within the range of 5700-6150 Da allowed to clearly observe the fading and subsequent recovery of the two signals at 5945.7 and 5959.8 Da (figure 63). Peaks at m/z 3568.0 and 3630.8 observed in the control experiment on Sepharose can be discarded (figure 62B). CPB inhibitory activity in the elution fraction (649.5 ± 6.8 mU/ml) corroborates the presence of at least one CPB inhibitor present in the *N. versicolor* crude extract.

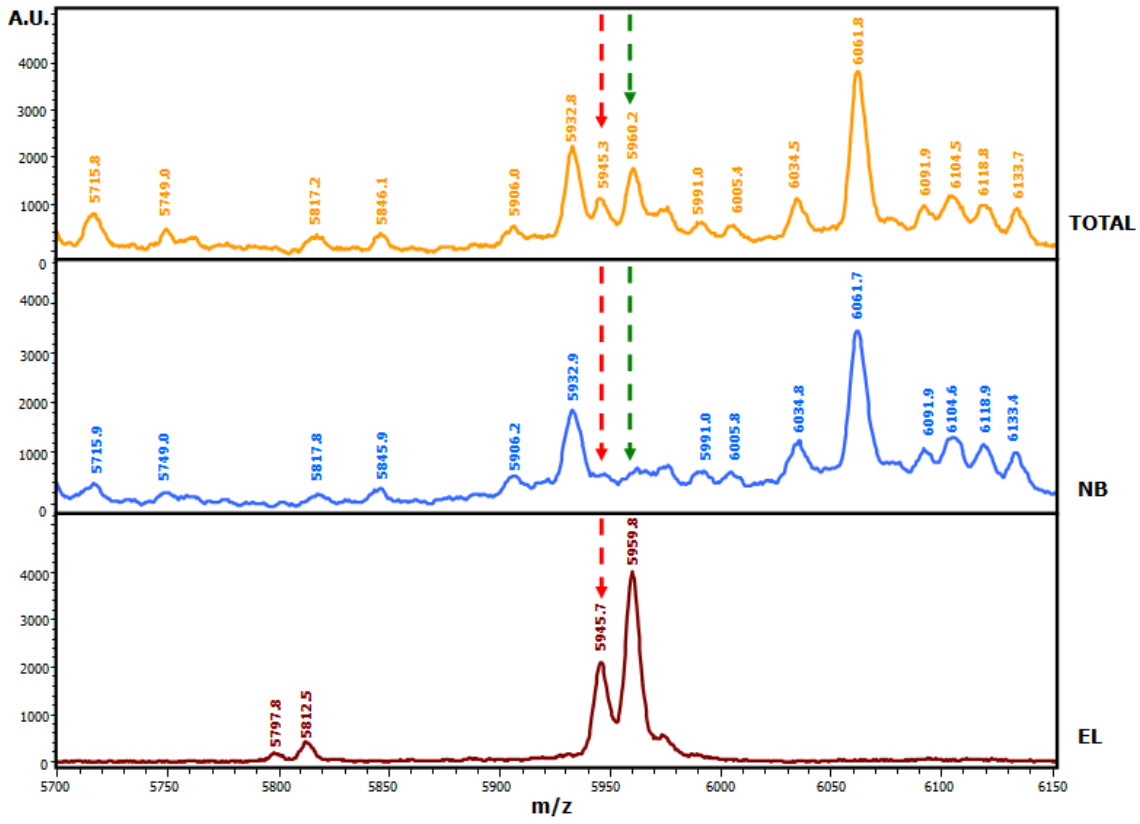


Figure 63. Detailed view of the Intensity Fading MALDI-TOF MS of *N. versicolor* crude extract against CPB-glyoxal Sepharose® CL-4B. TOTAL: crude extract of *N. versicolor* in 20 mM Tris-HCl buffer pH 7.5, 100 mM NaCl. **NB:** non-binding molecules, interaction at pH 7.5, 10 min. **EL:** elution with 0.5% v/v TFA, 10 min. Interaction assays were carried out in duplicate at room temperature. The experimental conditions of MALDI-TOF MS are described in section III.7.2.

IV.3.1.5. IF MALDI-TOF MS of *H. carunculata* crude extract against pepsin-NSH activated Sepharose™ 4 fast flow

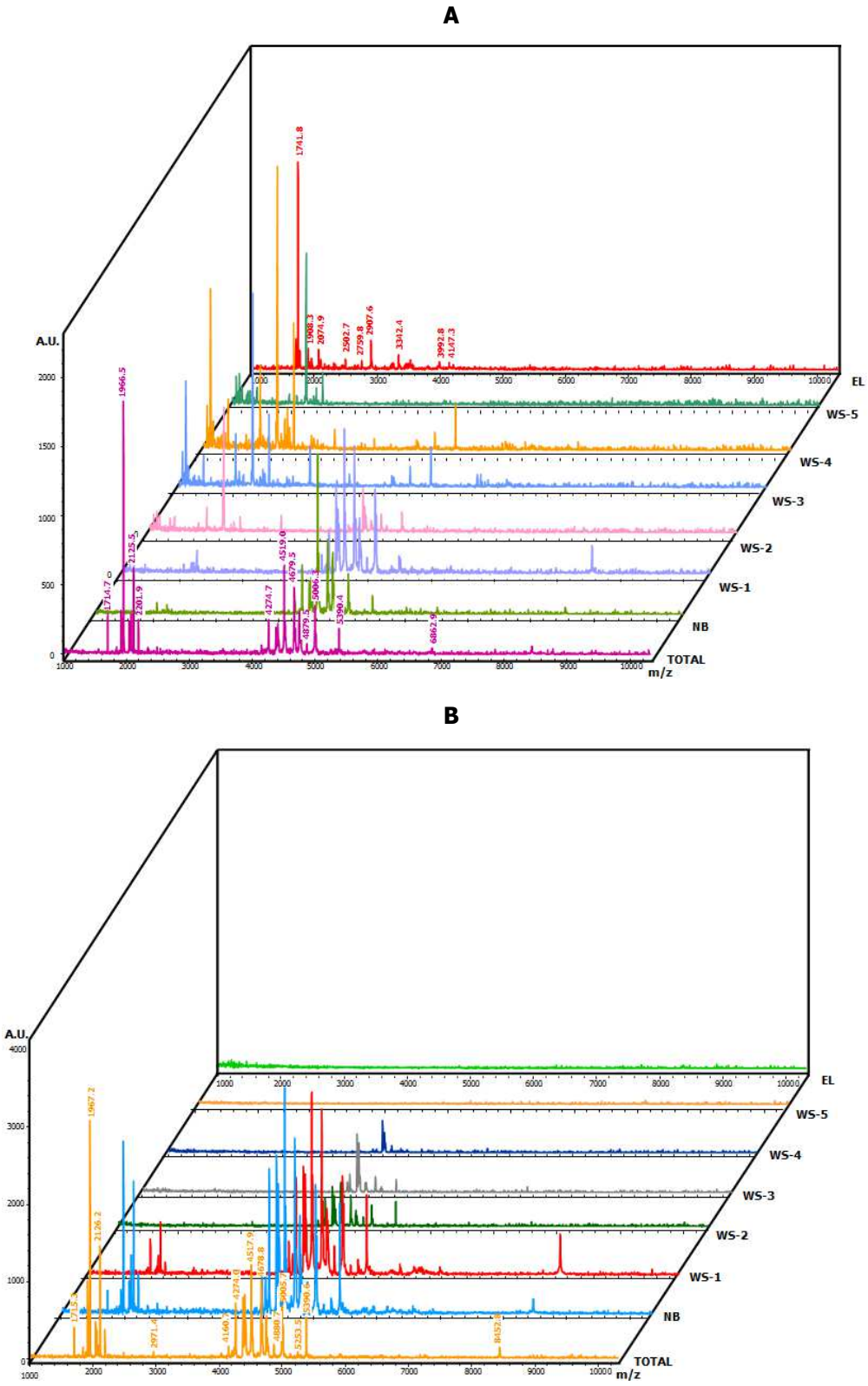


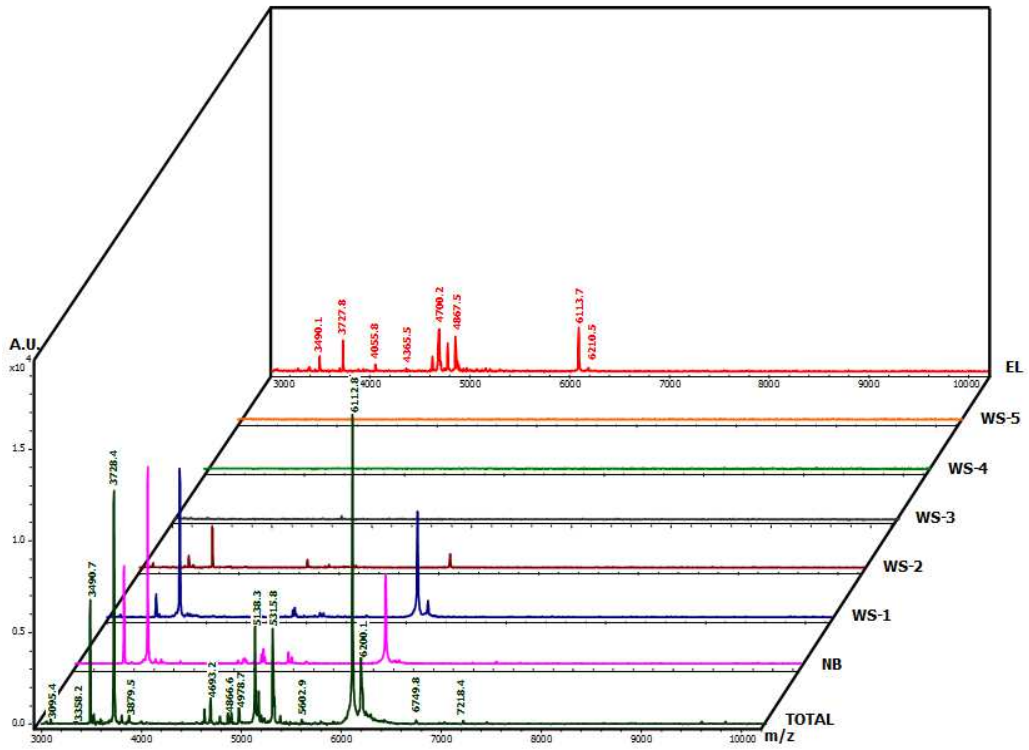
Figure 64. Intensity Fading MALDI-TOF MS of *H. carunculata* crude extract against pepsin- NHS activated Sepharose™ 4 fast flow (A) and NHS activated Sepharose™ 4 fast flow (B). TOTAL: crude extract of *H. carunculata* in 100 mM sodium acetate buffer at pH 4.4. **NB:** non-binding molecules, interaction at pH 4.4, 10 min. **WS-1 to WS-5:** washing at pH 4.4. . **EL:** elution with 0.5% v/v TFA, 10 min. Interaction assays were carried out in duplicate at room temperature. The experimental conditions of MALDI-TOF MS are described in section III.7.2**.**

Results with *H. carunculata* crude extract against pepsin-Sepharose initially showed a moderately complex control mass spectrum (figure 64A). After interaction with immobilized pepsin, peaks around 2000 Da disappeared from the spectrum (non-retained fraction), as a result of affinity interaction or enzymatic digestion by pepsin. During the washing steps, release of peaks between 4000 and 5000 Da was observed as well as other peaks that were not initially detected in the control mass spectrum.

Once the enzyme-inhibitor complexes were dissociated, one major signal at 1741.8 Da was recovered (pepsin inhibitory activity assay of the elution fraction = 9.3 ± 0.6 mU/ml). Control experiment with Sepharose did not reveal any evidence of non-specific interactions of extract with Sepharose matrix (figure 64B). Therefore, the previously observed peak of 1741.8 Da corresponds to specific interactions with pepsin. It is important to note that this peak of 1743.8 Da was not detected in the MALDI spectrum (figure 64A) of the crude extract (TOTAL), probably due to signal suppression by other peaks present in the heterogeneous extract. However, it is not possible to discard that this elution peak could be a result of pepsin degradation maintaining pepsin inhibitory properties.

IV.3.1.6. IF MALDI-TOF MS of *S. helianthus* crude extract against papain-glyoxal Sepharose® CL-4B

A



B

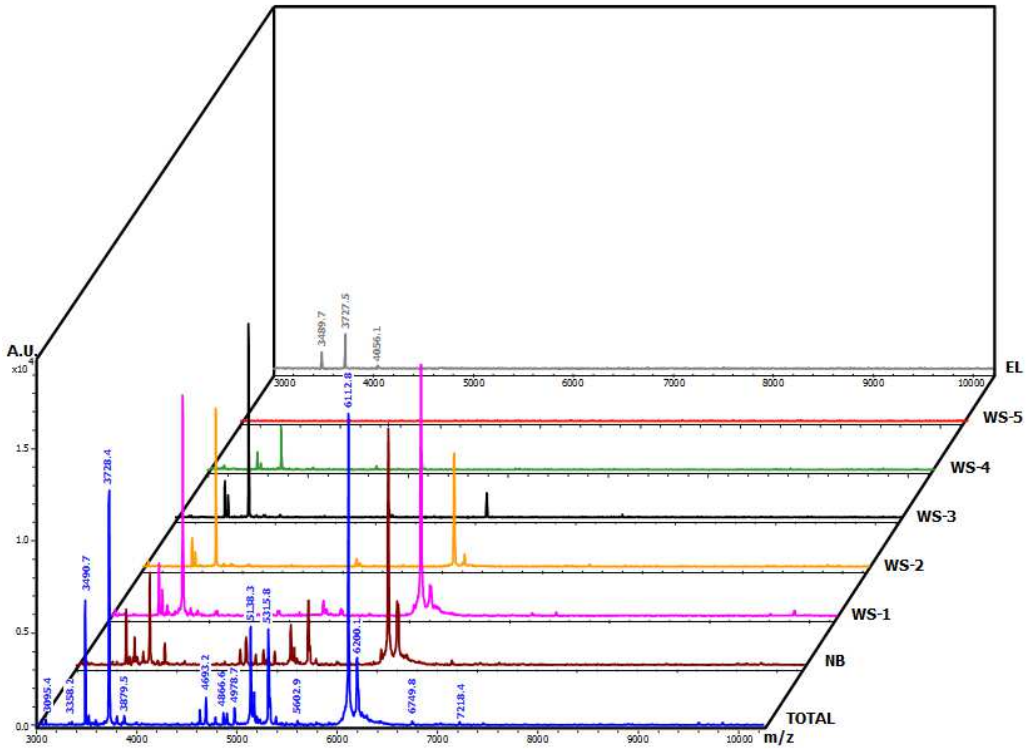


Figure 65. Intensity Fading MALDI-TOF MS of *S. helianthus* crude extract against papain-glyoxal Sepharose® CL-4B (A) and Sepharose® CL-4B (B). TOTAL: crude extract of *S. helianthus* in 20 mM Tris-HCl buffer pH 7.0, 100 mM KCl, 0.1 mM EDTA, 3 mM DTT. **NB:** non-binding molecules, interaction at pH 7.0, 10 min. **WS-1 to WS-5:** washing at pH 7.0. **EL:** elution with 0.5% v/v TFA, 10 min. Interaction assays were carried out in duplicate at room temperature. The experimental conditions of MALDI-TOF MS are described in section III.7.2**.**

Results obtained with *S. helianthus* crude extract and papain-Sepharose are shown in figure 65A. Control mass spectrum displayed peaks between 3000 and 7218.4 Da with a major peak of 6112.8 Da. After affinity interaction between the sample and the immobilized protease, a decay of a signal at m/z 6112.8 in the mass spectrum of the non-retained fraction was observed. Subsequent washings clarified mass spectra. After dissociation of papain-binding molecule complexes, a 6113.7 Da signal was recovered in the mass spectrum, as well as other peaks of 3490.1, 3727.8, 4055.8, 4700.2 and 4867.5 Da, where at least the first 3 ones must be discarded because they are product of non-specific interactions with Sepharose (figure 65B).

IF MALDI TOF MS indicates the presence of at least one molecule that selectively interacts with papain. Papain inhibitory activity assay of the elution fraction (34.7 ± 0.4 mU/ml) confirms the results obtained by MALDI analysis.

The presence of 6113.7 signal probably corresponds to ShPI-I, the major trypsin inhibitor in *S. helianthus* extract. Papain inhibitory activity has been demonstrated by kinetic studies (Delfín *et al.*, 1996) and more recently through Biacore™ with K_i value of 10^{-7} M (unpublished results). However, the occurrence of an additional papain inhibitor can not be discarded.

IV.3.1.7. IF MALDI-TOF MS of *N. peloronta* crude against trypsin-glyoxal Sepharose® CL-4B

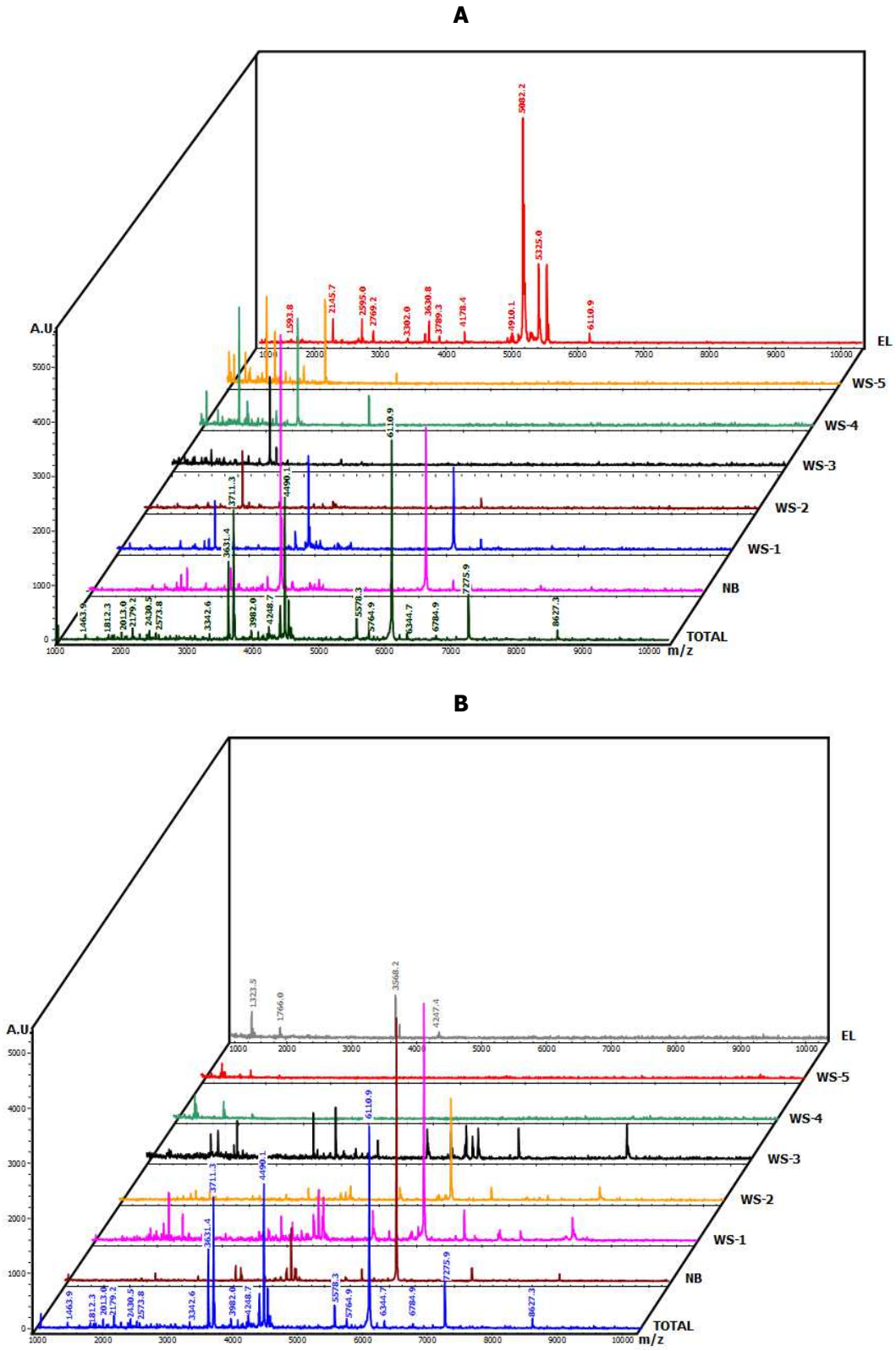


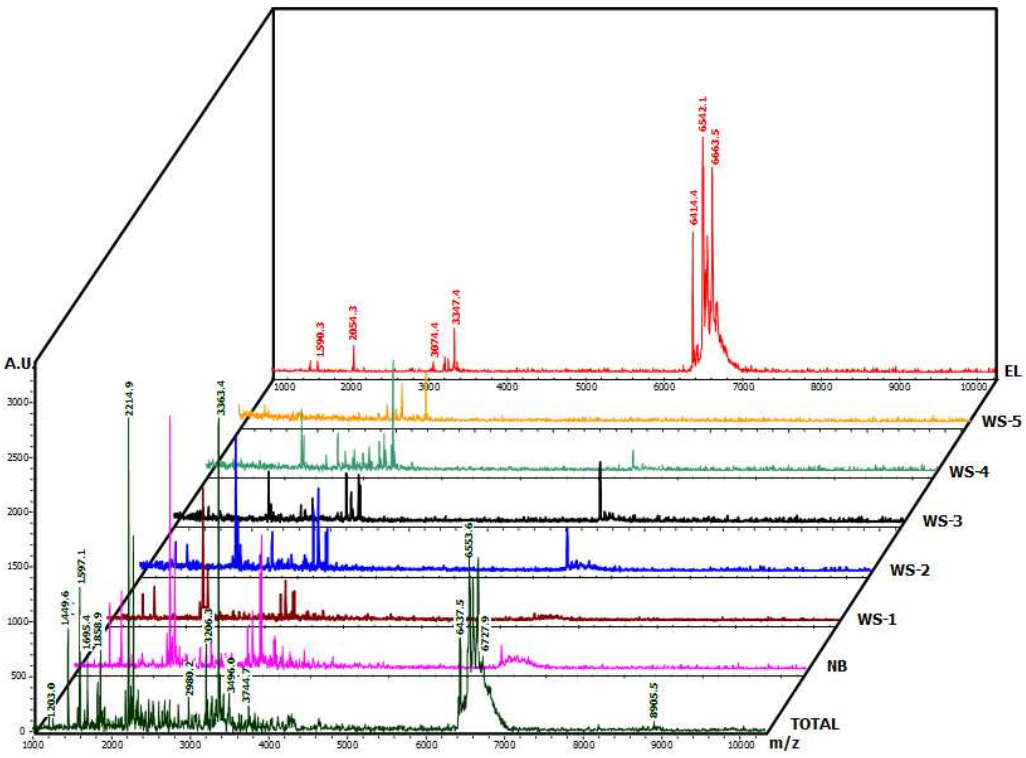
Figure 66. Intensity Fading MALDI-TOF MS of *N. peloronta* crude extract against trypsin-glyoxal Sepharose® CL-4B (A) and Sepharose® CL-4B (B). TOTAL: crude extract of *N. peloronta* in 20 mM Tris-HCl buffer pH 8.0, 150 mM NaCl, 20 mM $CaCl_2$. **NB:** non-binding molecules, interaction at pH 8.0, 10 min. **WS-1 to WS-5:** washing at pH 8.0. **EL:** elution with 0.5% v/v TFA, 10 min. Interaction assays were carried out in duplicate at room temperature. The experimental conditions of MALDI-TOF MS are described in section III.7.2**.**

Figure 66A shows IF MALDI-TOF MS of *N. peloronta* crude extract against trypsin-Sepharose. In this case, the elution fraction showed a major peak at 5082.2 Da, with minor peaks at m/z 5325.0, 3630.8, 2595.0, 2145.7, among others. Note that the presence of the peak at 5082.2 Da was not observed in the control mass spectrum or in the subsequent fractions in the analysis, which may be the result of signal suppression effect or partial tryptic digestion of a larger protein in the extract.

Minor peaks of 3568.2 and 3630.8 Da must be discarded as molecules that interact selectively with trypsin, since they also were observed in the elution fraction after the interaction assay with Sepharose (figure 66B). Trypsin inhibitory activity measured on the elution fraction confirmed the presence of at least one trypsin inhibitor in the *N. peloronta* extract (inhibitory activity = 710.6 ±19.5 mU/ml).

IV.3.1.8. IF MALDI-TOF MS of *L. isodyctialis* crude extract against subtilisin-glyoxal Sepharose® CL-4B

A



B

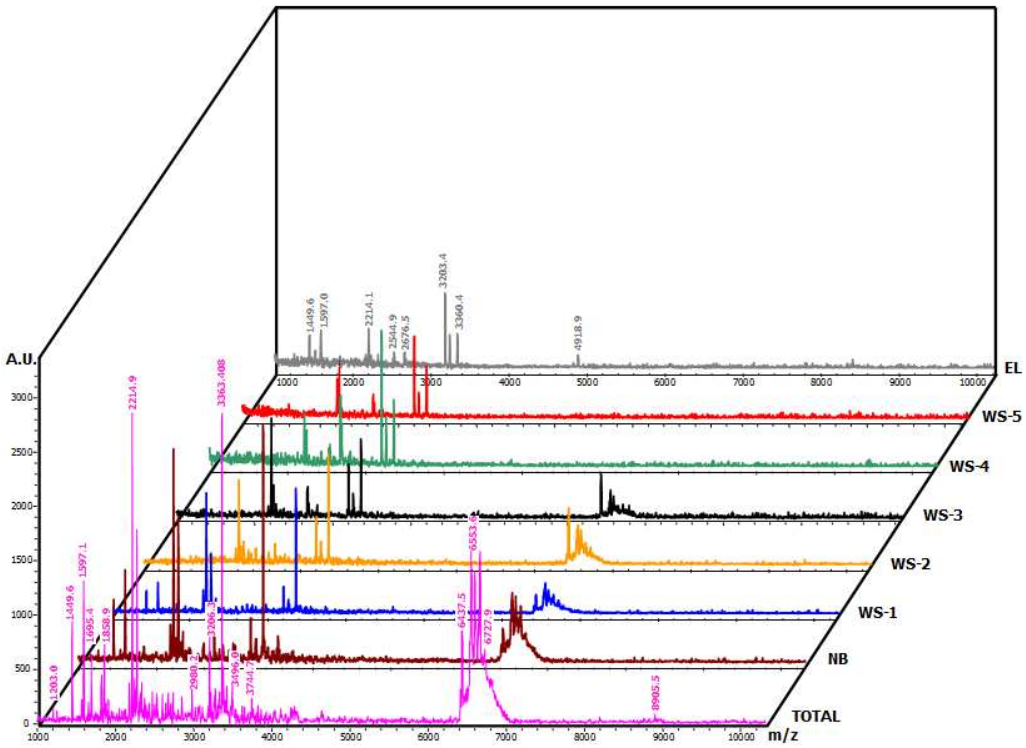


Figure 67. Intensity Fading MALDI-TOF MS of *L. isodyctialis* crude extract against subtilisin-glyoxal Sepharose® CL-4B (A) and Sepharose® CL-4B (B). TOTAL: crude extract of *L. isodyctialis* in 50 mM Tris-HCl buffer pH 8.5. **NB:** non-binding molecules, interaction at pH 8.5, 10 min. **WS-1 to WS-5:** washing at pH 8.5. **EL:** elution with 0.5% v/v TFA, 10 min. Interaction assays were carried out in duplicate at room temperature. The experimental conditions of MALDI-TOF MS are described in section III.7.2**.**

The *L. isodyctialis* crude extract against subtilisin-Sepharose showed several not completely resolved peaks around 6542 Da in the control mass spectrum, which were selectively retained and subsequently eluted under the experimental conditions used (figure 67A). The interaction assay between *L. isodytialis* crude extract and Sepharose revealed the presence of molecules in the elution fraction at 2214.1, 3203.4, 3360.4 Da, among other minor peaks (figure 67B). IF MALDI-TOF MS result as well as the subtilisin inhibitory activity found in the elution fraction (1945.2 ± 29.1 mU/ml) confirmed the presence of at least one inhibitor that selectively interacts with subtilisin.

IF MALDI-TOF MS approach has proven to be an important and effective strategy for the identification of protease inhibitors against proteases belonging to different mechanistic classes. Results confirm the validity of the method not only for CPA and trypsin inhibitors, as previously have been described, but also for inhibitors against other serine protease families, such as subtilisin and proteases from other mechanistic classes like cysteine and aspartic proteases.

IV.3.2. Analysis of protease inhibitors by MALDI Top-Down Sequencing (MALDI-TDS)

IF MALDI-TOF MS can be jointly used with MALDI Top-Down Sequencing (MALDI-TDS) to determine the sequence of resolved and/or single inhibitor peaks obtained in the elution fraction of IF MALDI-TOF MS. Two examples are described for this strategy: The first one is the major peak resulting from the elution of affinity interaction between *H. carunculata* crude extract and pepsin-NHS activated Sepharose. The major peak of this fraction was fragmented by CID MALDI-TOF/TOF and subsequently analyzed by *de novo* sequencing. Another example was ISD MALDI-TOF MS fragmentation and sequence analysis of the elution fraction of affinity interaction between *S. helianthus* crude extract and trypsin-glyoxal Sepharose.

IV.3.2.1. MALDI-Top-Down Sequencing using CID fragmentation

Taking into account that the molecular mass of the elution peak obtained in the IF MALDI-TOF MS of *H. carunculata* crude extract against immobilized pepsin is low, fragmentation by CID in the MALDI-TOF mass spectrometer was performed. For this purpose, the elution fraction of IF MALDI-TOF MS was initially reduced and S-carbamidomethylated in order to determine the possible existence of cysteine residues and disulfide bonds. This is an important aspect to be considered when sequencing, since it has been observed that disulfide-rich small molecules dramatically reduce MALDI fragmentation efficiency and hinder the subsequent *de novo* sequencing analysis.

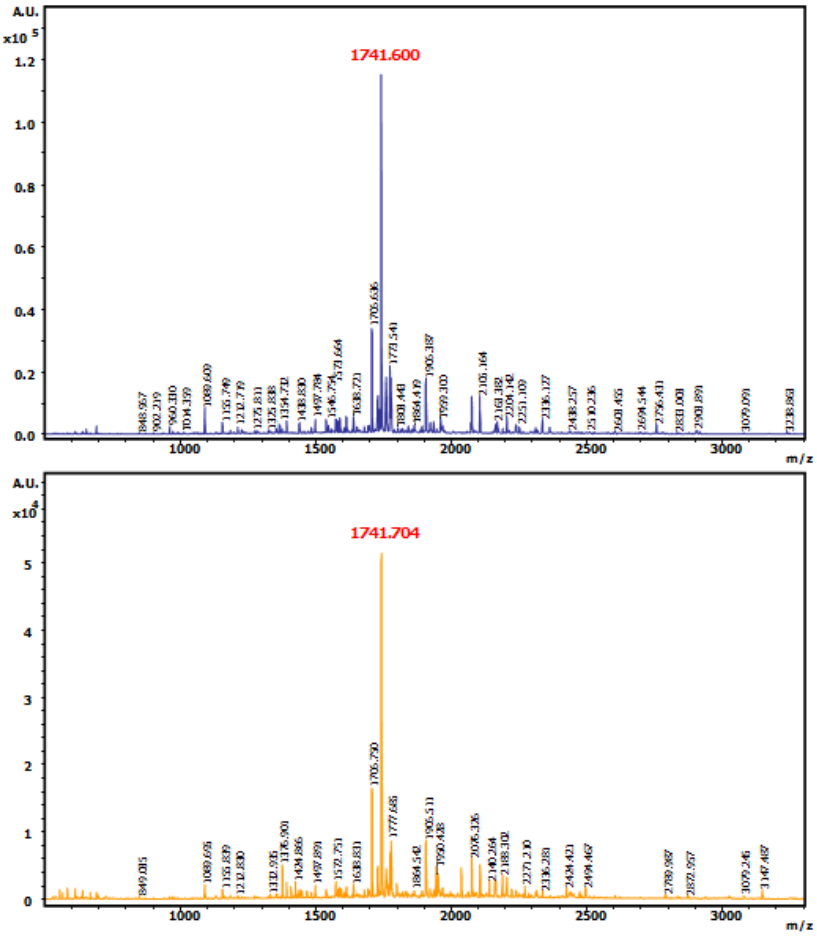


Figure 68. Reduction and S-carbamidomethylation of the IF MALDI-TOF MS elution fraction from *H. carunculata* crude extract against pepsin – NHS activated Sepharose™ 4 fast flow.

A. MALDI-TOF MS spectrum of the elution fraction, using α-CHCA as a matrix

B. MALDI-TOF MS spectrum after reduction and carbamidomethylation of the elution fraction, using α-CHCA as a matrix

Figure 68 shows the result of the reduction and S-carbamidomethylation of the elution fraction from IF MALDI-TOF MS. MALDI spectra clearly revealed that this peak of 1741.6 Da has no cysteines and/or disulfide bonds, since after treatment with DTT and iodoacetamide, shift of the signal at m/z 1741.6 in the mass spectrum was not visualized.

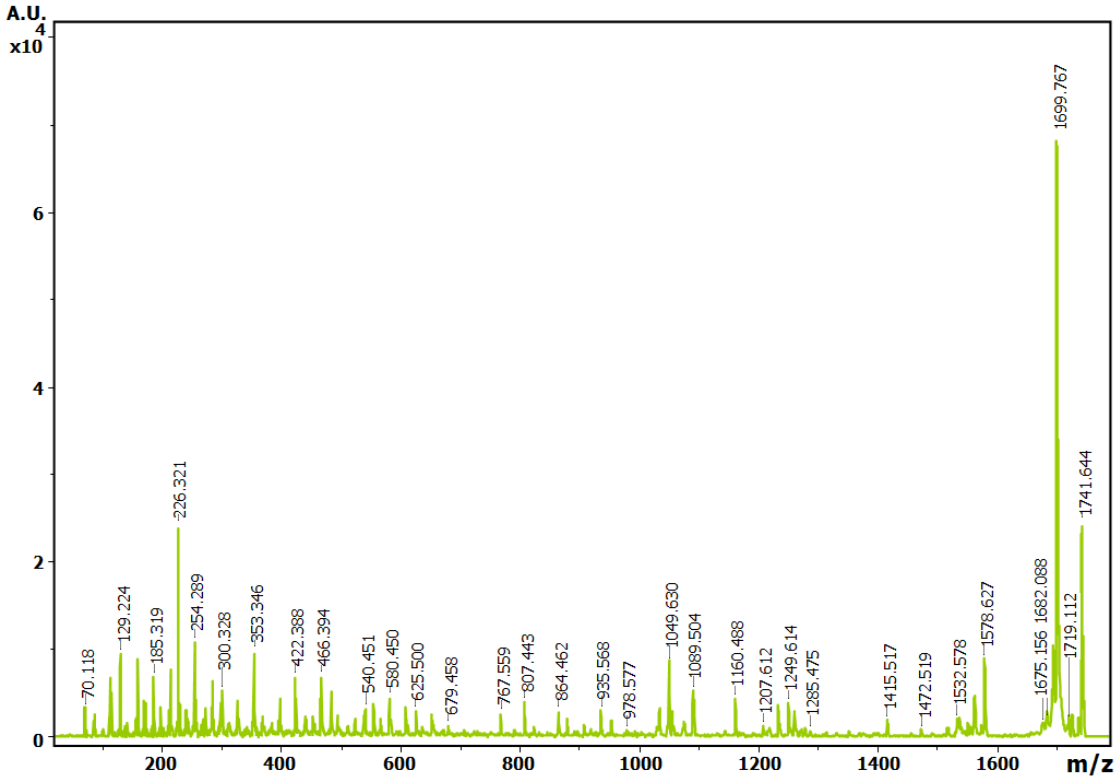


Figure 69. MALDI MS/MS spectrum of parent m/z=1741.6 from the IF MALDI-TOF MS of *H. carunculata* crude extract against pepsin – NHS activated Sepharose™ 4 fast flow. 1 µl of sample was deposited on a MTP 384 target plate polished steel T F (Bruker Daltonics), followed by deposition of 1 µl of α-CHCA as a matrix. The mixture was allowed to dry at room temperature.

CID MALDI-TOF/TOF fragmentation was directly performed using the elution fraction of IF MALDI-TOF MS of *H. carunculata* crude extract against pepsin – NHS activated Sepharose, according to the method described in section III.7.3.1. (figure 69). This mass spectrum, very rich in peptide fragmentation signals was analyzed using Biotools software (Bruker Daltonics) in which the "Rapid *de novo*" sequence tool was employed in order to elucidate the amino acid sequence of the peptide. For this experiment, the elution fraction was analyzed without reduction and S-carbamidomethylation, since as observed above, the peak of interest (parent m/z=1741.6) did not revealed the presence of cysteines and/or disulfide bonds.

The obtained sequences by *de novo* sequencing were subsequently submitted to the MS BLAST database in order to analyze the similarity of these sequences derived from the peptide at 1741.6 Da with other annotated proteins in this database. In addition, the first selected sequence was introduced to MASCOT database.

De novo sequencing of the 1743 Da molecule is shown in figure 71. The analysis produced 100 candidate sequences, from which the highest score sequence corresponds to the following amino acid sequence: YYRPARAGRAPRNGH. The other two candidate sequences equally scored showed differences in the last three amino acid residues at their C-terminal tails, compared to the sequence presented above: GNH and GHN, respectively (figure 70). Sequence was obtained by the *y*-series, which was confirmed by the *b*- and *a*-type series of ions.

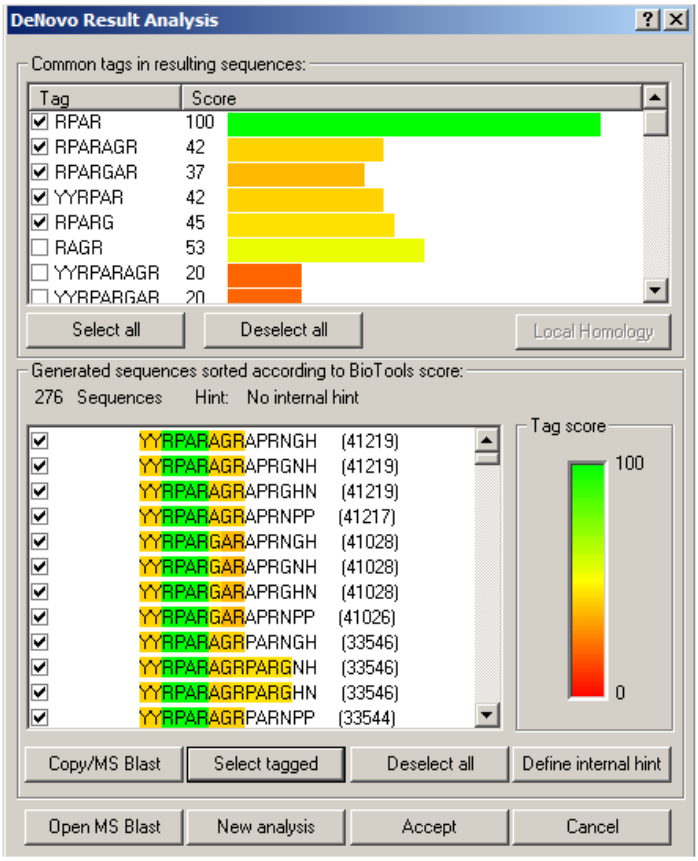


Figure 70. Analysis of the sequence by RapiDeNovo (Biootools 3.2) using MALDI MS/MS spectrum of the parent ion at m/z=1741.6 from the IF MALDI-TOF MS of *H. carunculata* crude extract against pepsin – NSH Sepharose™ 4 fast flow

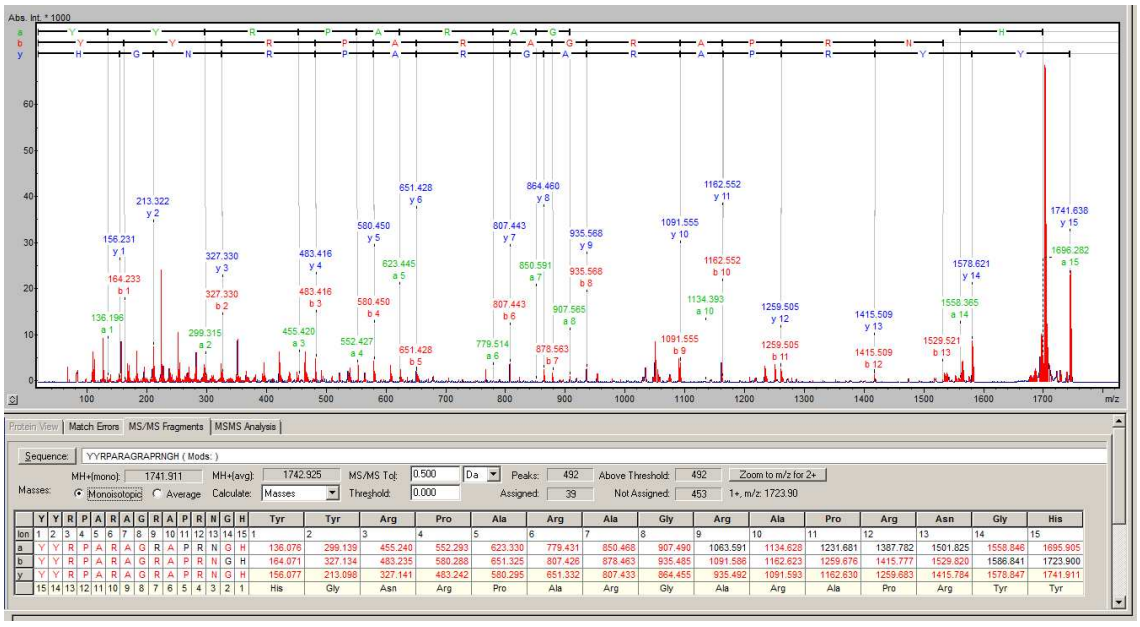


Figure 71. Sequence obtained from MS/MS spectrum by *de novo* sequencing of MALDI MS/MS spectrum of the parent ion at m/z=1741.6 from the IF MALDI-TOF of *H. carunculata* crude extract against pepsin – NSH Sepharose™ 4 fast flow

The analysis of MS/MS spectrum of the sample in the MS BLAST database produced hits corresponding to protease inhibitors or other unrelated proteins with the studied marine organism (figure 72). This result is in agreement with the fact that proteinaceous aspartic protease inhibitors are rare in nature and unevenly distributed among classes of organisms in contrast to proteinaceous inhibitors of serine and cysteine proteases (Laing and McManus, 2002).

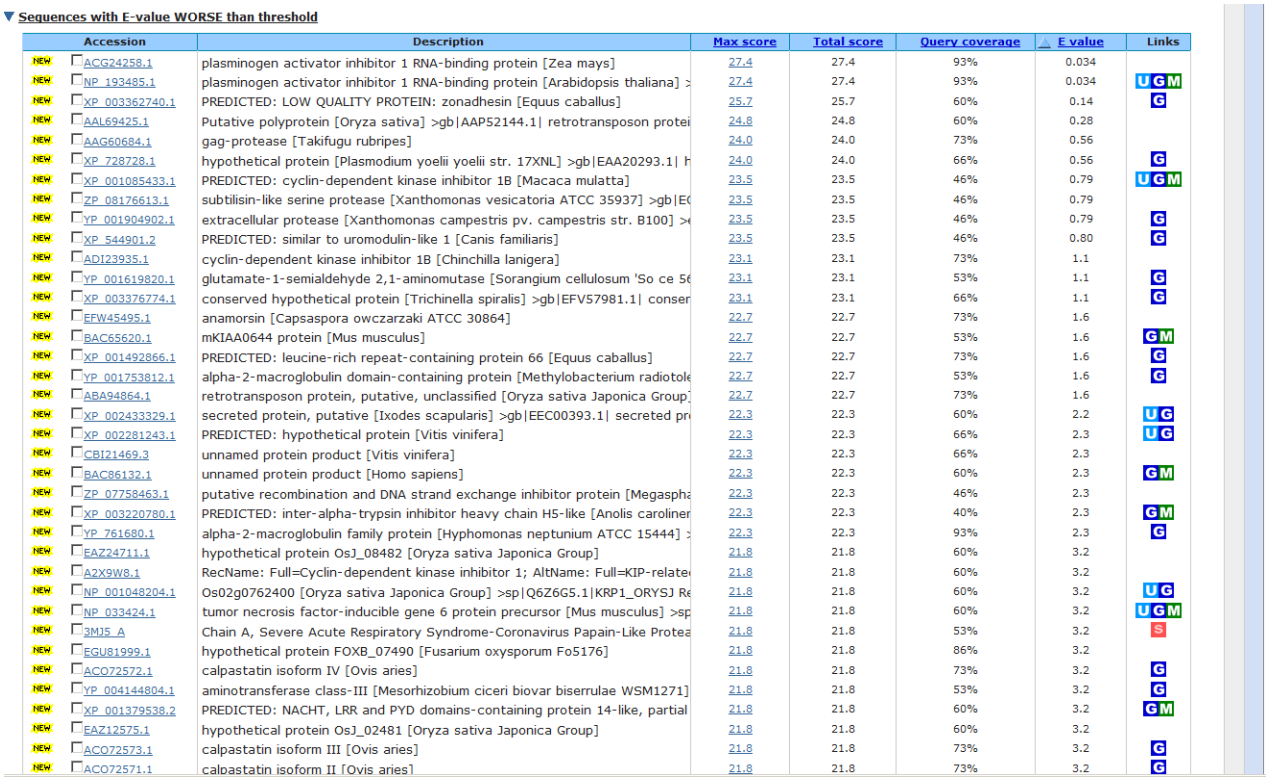
Homology searches by protein blast using NCBI/BLAST of the obtained sequence did not reveal any sequence homology with other aspartic protease inhibitors (figure 73). As the result obtained in MS BLAST search, identified proteins were neither related to aspartic acid protease inhibitors from other organisms nor proteins from the studied marine extract.

Color Key: red = positive hit; green = borderline hit; black = negative result

Summary: Get the selected sequences Restablecer

Sequences producing High-scoring Segment Pairs:	High Score	Total Score	
/:gi|283782395|ref|YP 003373150.1| pseudouridine synthase...	64	1246	☑
/:gi|323451402|gb|EGB07279.1| hypothetical protein AURAND...	56	1170	☑
/:gi|86156673|ref|YP 463458.1| LigA [Anaeromyxobacter deh...	63	1024	☑
/:gi|88854845|ref|ZP 01129511.1| ATP-dependent RNA helica...	57	810	☑
/:gi|170782512|ref|YP 001710845.1| putative ATP-dependent...	57	786	☑
/:gi|220926361|ref|YP 002501663.1| RNA-binding S4 domain-...	75	753	☑
/:gi|219847277|ref|YP 002461710.1| pseudouridine synthase...	62	697	☑
/:gi|308508467|ref|XP 003116417.1| hypothetical protein C...	53	658	☑
/:gi|76811844|ref|YP 332924.1| exodeoxyribonuclease V, be...	51	568	☑
/:gi|94970889|ref|YP 592937.1| pseudouridine synthase, Rs...	50	532	☑
/:gi|323451718|gb|EGB07594.1| hypothetical protein AURAND...	63	513	☑
/:gi|255531694|ref|YP 003092066.1| hypothetical protein P...	64	477	☑
/:gi|295840249|ref|ZP 06827182.1| conserved hypothetical ...	56	472	☑
/:gi|298293868|ref|YP 003695807.1| RNA-binding S4 domain ...	63	469	☑
/:gi|254384839|ref|ZP 05000176.1| hypothetical protein SS...	63	467	☑
/:gi|323454679|gb|EGB10549.1| hypothetical protein AURAND...	50	457	☑
/:gi|90102218|gb|ABD85297.1| antifreeze glycoprotein [Bor...	61	449	☑
/:gi|73541573|ref|YP 296093.1| pseudouridine synthase, Rs...	62	448	☑
/:gi|323445462|gb|EGB02058.1| expressed protein [Aureococ...	50	435	☑
/:gi|303282145|ref|XP 003060364.1| predicted protein [Mic...	49	435	☑
/:gi|323452022|gb|EGB07897.1| hypothetical protein AURAND...	45	432	☑
/:gi|302531681|ref|ZP 07284023.1| predicted protein [Stre...	47	399	☑
/:gi|66043725|ref|YP 233566.1| helicase, C-terminal:Type ...	60	391	☑
/:gi|254263783|ref|ZP 04954648.1| hypothetical protein BU...	56	388	☑
/:gi|323446801|gb|EGB02837.1| hypothetical protein AURAND...	58	385	☑
/:gi|323450142|gb|EGB06025.1| expressed protein [Aureococ...	51	377	☑
/:gi|323455940|gb|EGB11807.1| hypothetical protein AURAND...	53	373	☑
/:gi|86157987|ref|YP 464772.1| Fis family transcriptional...	58	367	☑
/:gi|323456001|gb|EGB11868.1| hypothetical protein AURAND...	61	360	☑
/:gi|255942327|ref|XP 002561932.1| Pc18g00860 [Penicilliu...	69	346	☑
/:gi|91784123|ref|YP 559329.1| pseudouridine synthase, Rs...	57	331	☑
/:gi|282860764|ref|ZP 06269830.1| NADH dehydrogenase (ubi...	67	329	☑

Figure 72. MALDI MS BLAST result for sequence obtained from *de novo* analysis



▼ Sequences with E-value WORSE than threshold

Accession	Description	Max score	Total score	Query coverage	E value	Links
NEW ACG24258.1	plasminogen activator inhibitor 1 RNA-binding protein [Zea mays]	27.4	27.4	93%	0.034	
NEW NP_193485.1	plasminogen activator inhibitor 1 RNA-binding protein [Arabidopsis thaliana] >	27.4	27.4	93%	0.034	U G M
NEW XP_003362740.1	PREDICTED: LOW QUALITY PROTEIN: zonadhesin [Equus caballus]	25.7	25.7	60%	0.14	G
NEW AAL69425.1	Putative polyprotein [Oryza sativa] >gb|AAP52144.1| retrotransposon protei	24.8	24.8	60%	0.28	
NEW AAG60684.1	gag-protease [Takifugu rubripes]	24.0	24.0	73%	0.56	
NEW XP_728728.1	hypothetical protein [Plasmodium yoelii yoelii str. 17XNL] >gb|EAA20293.1| h	24.0	24.0	66%	0.56	G
NEW XP_001085433.1	PREDICTED: cyclin-dependent kinase inhibitor 1B [Macaca mulatta]	23.5	23.5	46%	0.79	U G M
NEW ZP_08176613.1	subtilisin-like serine protease [Xanthomonas vesicatoria ATCC 35937] >gb|EC	23.5	23.5	46%	0.79	
NEW YP_001904902.1	extracellular protease [Xanthomonas campestris pv. campestris str. B100] ><	23.5	23.5	46%	0.79	G
NEW XP_544901.2	PREDICTED: similar to uromodulin-like 1 [Canis familiaris]	23.5	23.5	46%	0.80	G
NEW ADI23935.1	cyclin-dependent kinase inhibitor 1B [Chinchilla lanigera]	23.1	23.1	73%	1.1	
NEW YP_001619820.1	glutamate-1-semialdehyde 2,1-aminomutase [Sorangium cellulosum 'So ce 56	23.1	23.1	53%	1.1	G
NEW XP_003376774.1	conserved hypothetical protein [Trichinella spiralis] >gb|EFV57981.1| conser	23.1	23.1	66%	1.1	G
NEW EFW45495.1	anamorsin [Capsaspora owczarzaki ATCC 30864]	22.7	22.7	73%	1.6	
NEW BAC65620.1	mKIAA0644 protein [Mus musculus]	22.7	22.7	53%	1.6	G M
NEW XP_001492866.1	PREDICTED: leucine-rich repeat-containing protein 66 [Equus caballus]	22.7	22.7	73%	1.6	G
NEW YP_001753812.1	alpha-2-macroglobulin domain-containing protein [Methylobacterium radiotole	22.7	22.7	53%	1.6	G
NEW ABA94864.1	retrotransposon protein, putative, unclassified [Oryza sativa Japonica Group]	22.7	22.7	73%	1.6	
NEW XP_002433329.1	secreted protein, putative [Ixodes scapularis] >gb|EEC00393.1| secreted pr	22.3	22.3	60%	2.2	U G
NEW XP_002281243.1	PREDICTED: hypothetical protein [Vitis vinifera]	22.3	22.3	66%	2.3	U G
NEW CBI21469.3	unnamed protein product [Vitis vinifera]	22.3	22.3	66%	2.3	
NEW BAC86132.1	unnamed protein product [Homo sapiens]	22.3	22.3	60%	2.3	G M
NEW ZP_07758463.1	putative recombination and DNA strand exchange inhibitor protein [Megaspha	22.3	22.3	46%	2.3	
NEW XP_003220780.1	PREDICTED: inter-alpha-trypsin inhibitor heavy chain H5-like [Anolis caroliner	22.3	22.3	40%	2.3	G M
NEW YP_761680.1	alpha-2-macroglobulin family protein [Hyphomonas neptunium ATCC 15444] >	22.3	22.3	93%	2.3	G
NEW EAZ24711.1	hypothetical protein OsJ_08482 [Oryza sativa Japonica Group]	21.8	21.8	60%	3.2	
NEW A2X9W8.1	RecName: Full=Cyclin-dependent kinase inhibitor 1; AltName: Full=KIP-relate	21.8	21.8	60%	3.2	
NEW NP_001048204.1	Os02g0762400 [Oryza sativa Japonica Group] >sp|Q6Z6G5.1|KRP1_ORYSJ Re	21.8	21.8	60%	3.2	U G
NEW NP_033424.1	tumor necrosis factor-inducible gene 6 protein precursor [Mus musculus] >sp	21.8	21.8	60%	3.2	U G M
NEW 3MJ5_A	Chain A, Severe Acute Respiratory Syndrome-Coronavirus Papain-Like Protea	21.8	21.8	53%	3.2	S
NEW EGU81999.1	hypothetical protein FOXB_07490 [Fusarium oxysporum Fo5176]	21.8	21.8	86%	3.2	
NEW ACO72572.1	calpastatin isoform IV [Ovis aries]	21.8	21.8	73%	3.2	G
NEW YP_004144804.1	aminotransferase class-III [Mesorhizobium ciceri biovar biserrulae WSM1271]	21.8	21.8	53%	3.2	G
NEW XP_001379538.2	PREDICTED: NACHT, LRR and PYD domains-containing protein 14-like, partial	21.8	21.8	60%	3.2	G M
NEW EAZ12575.1	hypothetical protein OsJ_02481 [Oryza sativa Japonica Group]	21.8	21.8	60%	3.2	
NEW ACO72573.1	calpastatin isoform III [Ovis aries]	21.8	21.8	73%	3.2	G
NEW ACO72571.1	calpastatin isoform II [Ovis aries]	21.8	21.8	73%	3.2	G

Figure 73. Homology searches by protein blast using NCBI/BLAST of the obtained sequence by *de novo* sequencing

The results obtained by MALDI-TDS using CID fragmentation are very interesting given the infrequency of pepsin inhibitors in nature. Further purification of the affinity elution fraction will allow the structural characterization of this molecule in order to corroborate its amino acid sequence as well as the functional characterization as an aspartic protease inhibitor.

IV.3.2.2. MALDI-Top-Down Sequencing using ISD fragmentation

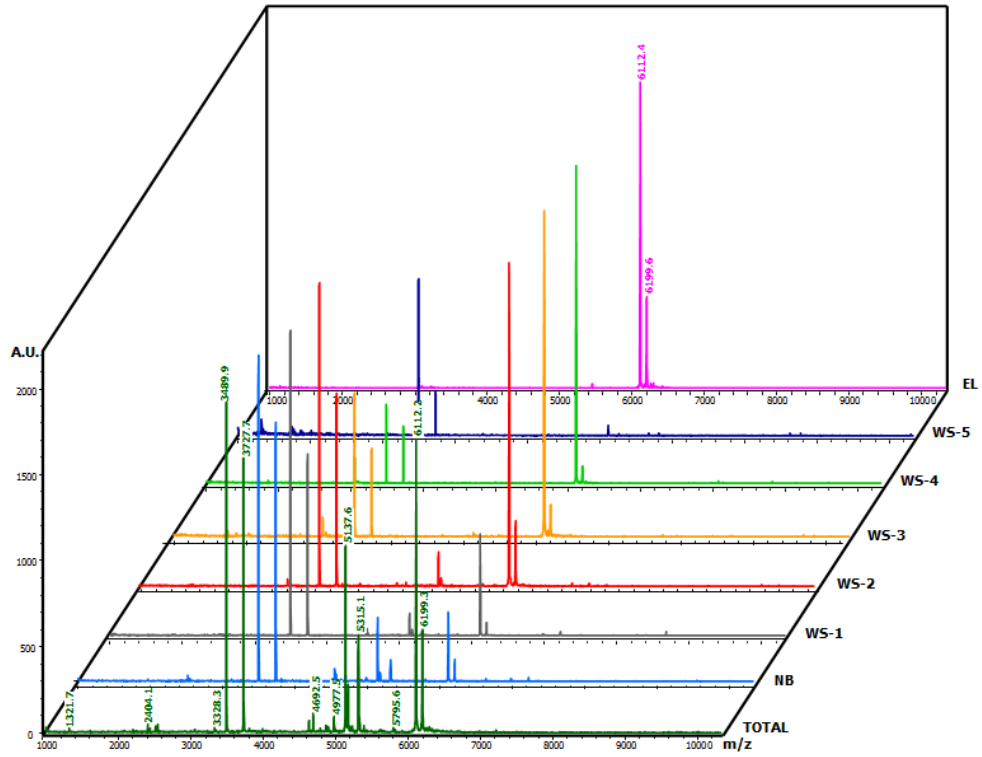
In order to analyze by MALDI-TDS previously identified protease inhibitors, partially purified by IF MALDI-TOF MS, ISD fragmentation was employed. The major advantage of this fragmentation technique is that it is not limited to low molecular weight peptides, in contrast to metastable fragmentation (Quinton *et al.*, 2007). Taking into account that ISD fragmentation needs to be performed on purified proteins (although mixture of few compounds is also possible), a homogeneous elution fraction from IF MALDI-TOF was required. Therefore, the interactions of several marine extracts with immobilized proteases of different mechanistic classes were analyzed and *S. helianthus* crude extract against trypsin-glyoxal Sepharose was selected.

Figure 74 displays IF MALDI-TOF MS results for *S. helianthus* crude extract against trypsin. Control mass spectrum revealed several peaks over the mass range analyzed, being its profile very similar to that shown above with papain. After interaction with the immobilized enzyme, fading of peaks at m/z 5137.6, 5315.1, 6112.2 and 6199.3 were visualized. Subsequently, during the washing steps, peaks at 6112.2 and 6199.3 Da were progressively removed from the affinity matrix, as a result of weak interaction under the experimental conditions used or due to an excess of sample applied. Finally, after obtaining a simplified mass spectrum, the elution step was performed. Elution fraction revealed the presence of two peaks at 6112.4 and 6199.6 Da.

These peaks detected in the MALDI spectrum correspond to the molecular masses of protease inhibitors ShPI-I (Delfín *et al.*, 1996) and ShPI-II (Díaz *et al.*, 1998) isolated and characterized from the sea anemone *S. helianthus*, earlier mentioned. Positive result for trypsin inhibitory activity (4866.6 ± 1229.7 mU/ml) was obtained from the elution fraction, confirming the presence of these inhibitors.

Signals at m/z 3489.5 and 3727.3 obtained from the interaction of *S. helianthus* crude extract on Sepharose must be discarded from the mass spectrum of the elution fraction obtained with trypsin-Sepharose, since they represent molecules that interacts in a non-specific mode with the support (figure 74B).

A



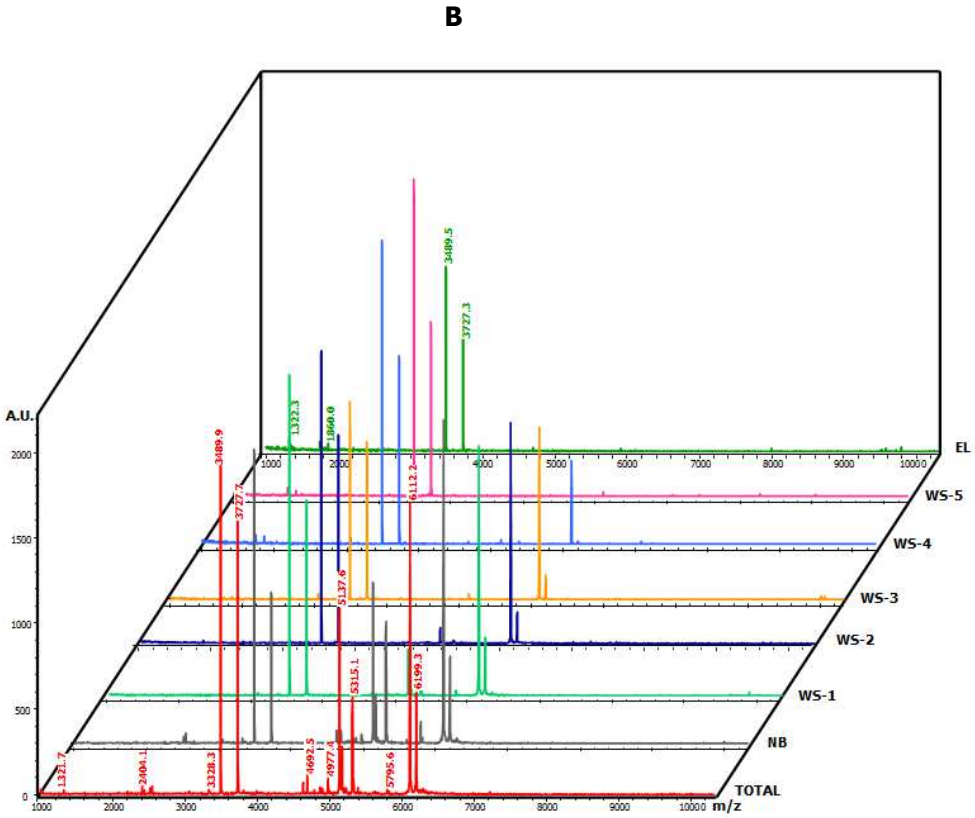


Figure 74. Intensity Fading MALDI-TOF MS of *S. helianthus* crude extract against trypsin-glyoxal Sepharose® CL-4B (A) and Sepharose® CL-4B (B). TOTAL: crude extract of *S. helianthus* in 20 mM Tris-HCl buffer pH 8.0, 150 mM NaCl, 20 mM $CaCl_2$. **NB:** non-binding molecules, interaction at pH 8.0, 10 min. **WS-1 to WS-5:** washing at pH 8.0. **EL:** elution with 0.5% v/v TFA, 10 min. Interaction assays were carried out in duplicate at room temperature. The experimental conditions of MALDI-TOF MS are described in section III.7.2.

ISD fragmentation assays were performed by using the following matrices: DHB, superDHB and 1,5-DAN. ISD spectra with the first two matrices produces C-terminal fragments (*y*- and *z*-ions) with reasonable intensities (Bruker Daltonics GmbH, 2008) whilst 1-5 DAN matrix partially reduces disulfide bonds (Serrano *et al.*, 2005).

ISD spectra were acquired through an automatic method, using reflectron mode over the mass range of 1000-5000 Da and calibrated with standard peptide mix prepared in the same matrix used for the sample. The most promising results in terms of quantity, intensity and resolution of fragments were achieved using 1,5-DAN matrix. Therefore, in this work only MALDI-TDS analyses obtained with this matrix are shown.

ISD spectrum obtained with 1,5 DAN matrix was processed by FlexAnalysis software (Bruker Daltonics, Germany), running a default method, where masses were annotated using a peak picking algorithm. In the ISD spectrum several fragments with remarkable intensities were mainly displayed over the mass range of 1000-3000 Da as it is shown in figure 75.

Subsequently, this spectrum was loaded on Biotools software (Bruker Daltonics, Germany) and the analysis strategy described by Bruker Daltonics GmbH, 2008 was followed. Initially, the spectrum with annotated masses was assigned as ISD type and then, Top-down analysis started by searching for sequence tags in the spectrum. At this point, 18 tags sequences mostly composed by 8 to 10 amino acid residues were found (figure 76). The ranking of the tags is made by a scoring factor (∑peak area/number of peak) (Bruker Daltonics GmbH, 2008). The 6 best scored sequences were selected and later submitted to MS BLAST database.

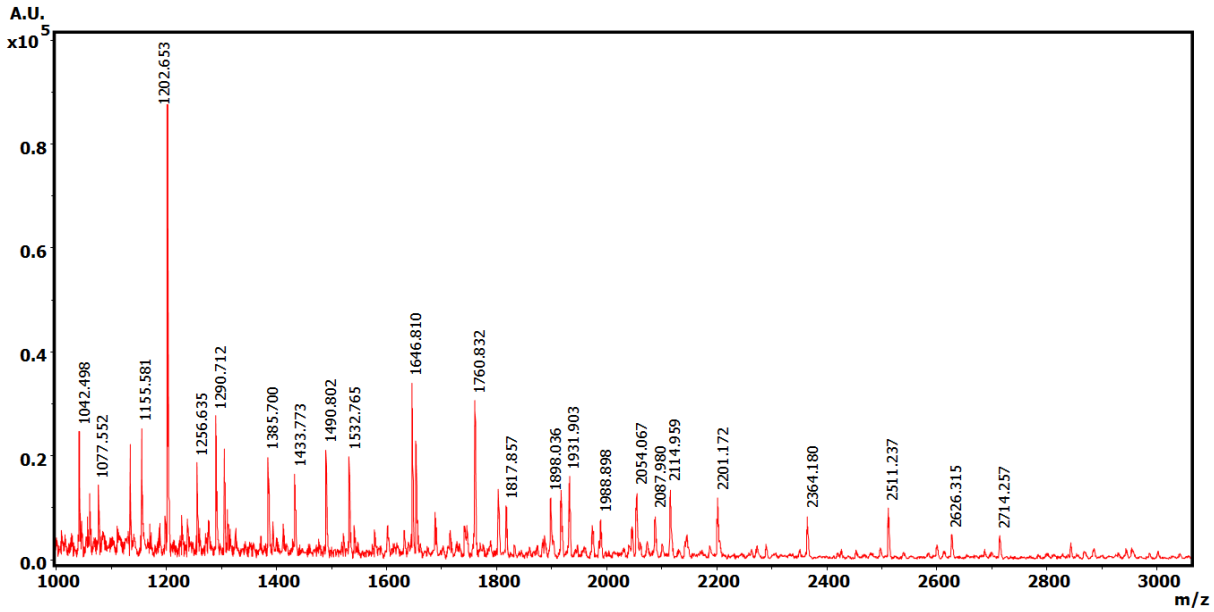


Figure 75. ReISD spectrum of IF MALDI-TOF MS elution fraction from *S. helianthus* crude extract against trypsin-glyoxal Sepharose® CL-4B. MALDI spectra were acquired automatically with five replicates and using 1,5-DAN matrix.

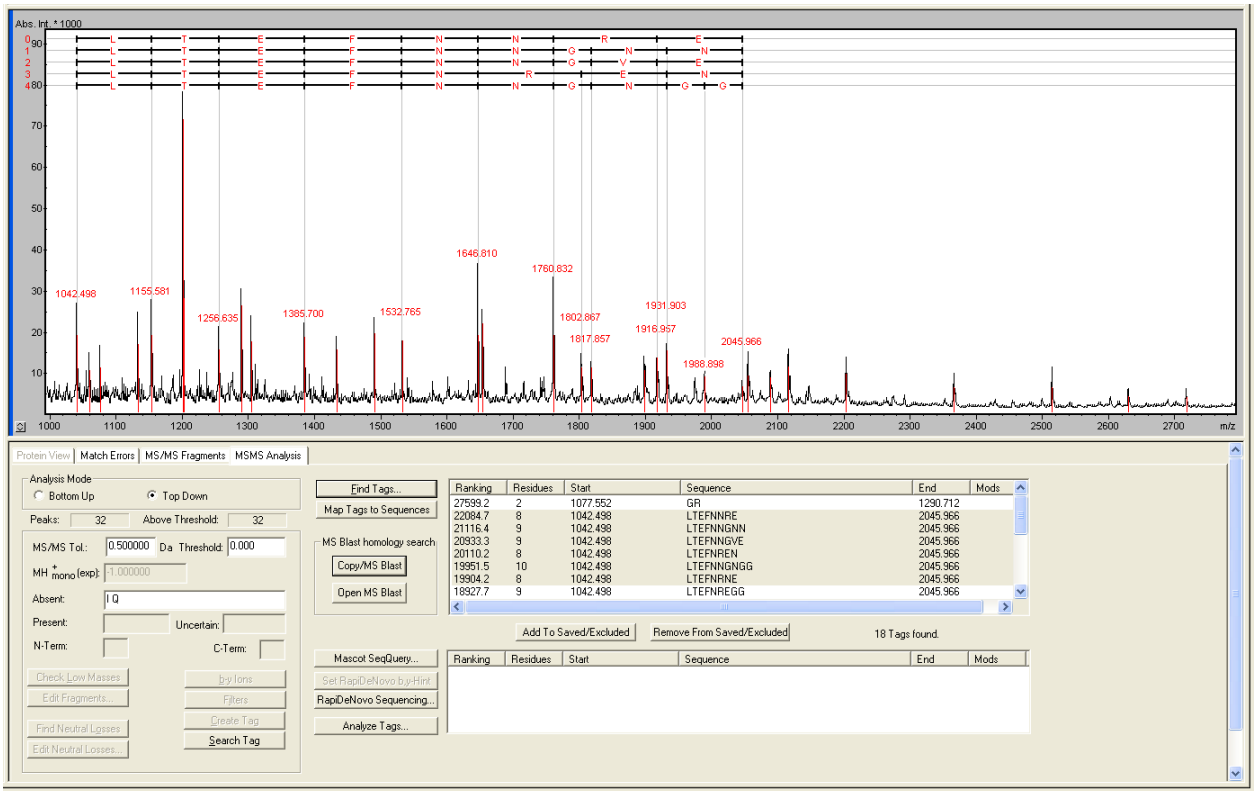


Figure 76. Sequence tags obtained from ReISD spectrum of IF MALDI-TOF MS elution fraction from *S. helianthus* crude extract against trypsin-glyoxal Sepharose® CL-4B. MALDI spectra were acquired automatically with five replicates and using 1,5-DAN matrix. Sequence tags were obtained with default settings

ShPI-I and ShPI-II were identified in the MS BLAST search result (figure 77) as positive hits (marked in red colour) with a score of 71 for both proteins. As shown in figure 78, 10 amino acid residues corresponding to these serine protease inhibitors were detected. Sequence analysis of the two molecules in Sequence Editor Software (Bruker Daltonics, Germany) revealed that the identified residues were common for both inhibitors and located close to the C-terminal tail (figure 79). This result is in agreement with the requirement mentioned above regarding to the sample purity, since more than one ISD fragment series could be detected in samples containing more than one main component. Thus, sequence differences established between ShPI I and II hindered the subsequent sequence analysis.

Figure 77. MS BLAST result for the sequence tags obtained from ReISD analysis

```
^ = gi|400070|sp|P31713.1|ISH1_STOHE RecName: Full=Kunitz-type proteinase
            inhibitor SHPI-1//:gi|157833813|pdb|1SHP|A Chain A, The Nmr
            Solution Structure Of A Kunitz-Type Proteinase Inhibitor From The
            Sea Anemone Stichodactyla Helianthus/
          Length = 55

Total Score: 71

                          0                 20                40
                          |                 |                 |              |55
gi|400070|sp|P31713.1   | ________________________________________________
Local hits (HSPs)        |                                   ________

 Score = 71 (36.1 bits)
 Identities = 10/10 (100%), Positives = 10/10 (100%)

Query:     94 GGNGNNFETL 103
              GGNGNNFETL
Sbjct:     37 GGNGNNFETL 46


^ = gi|3024045|sp|P81129.1|ISH2_STOHE RecName: Full=Kunitz-type proteinase
            inhibitor SHPI-2/
          Length = 55

Total Score: 71

                          0                 20                40
                          |                 |                 |              |55
gi|3024045|sp|P81129.   | ________________________________________________
Local hits (HSPs)        |                                   ________

 Score = 71 (36.1 bits)
 Identities = 10/10 (100%), Positives = 10/10 (100%)

Query:     94 GGNGNNFETL 103
              GGNGNNFETL
Sbjct:     37 GGNGNNFETL 46
```

Figure 78. MS BLAST result for the sequence tags obtained from ReISD analysis

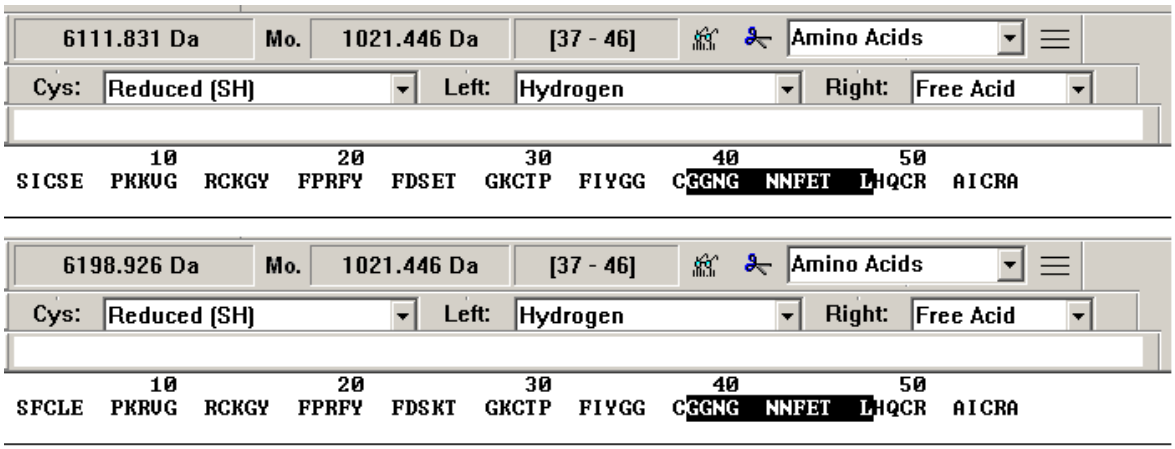


Figure 79. Sequence coverage for ShPI-1 (above) and ShPI-2 (below) by ReISD fragmentation

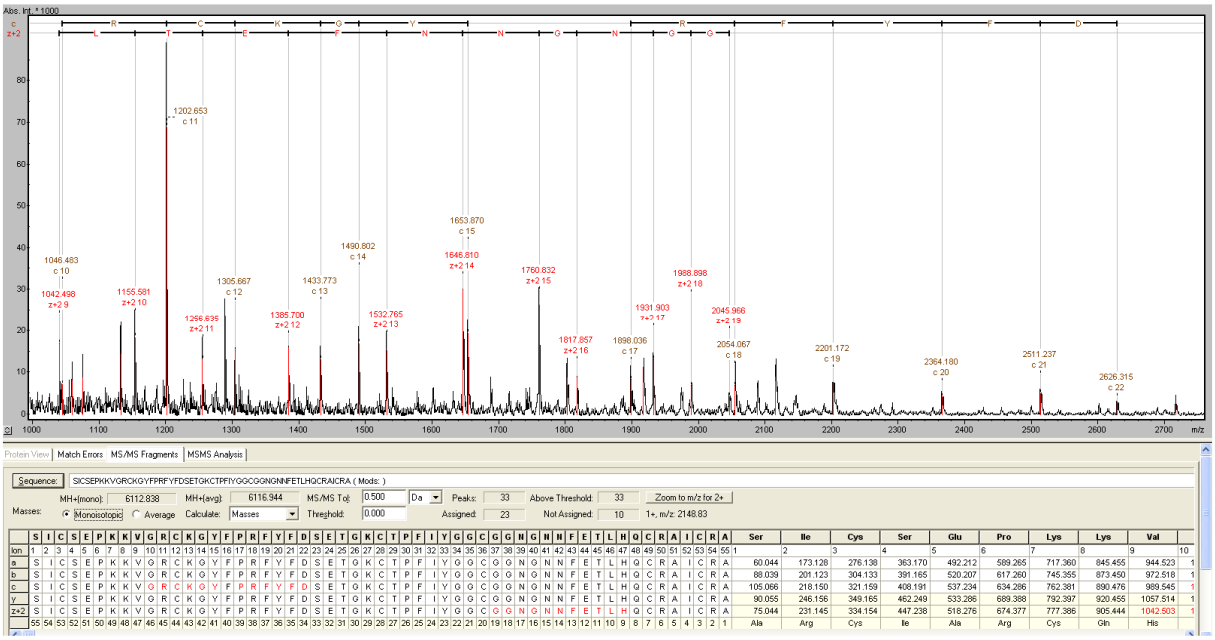


Figure 80. Tags mapping to sequences of ShPI-I in tree view

In order to check mapping of theoretical fragment masses and experimental masses in the spectrum, ShPI-I sequence was sent to Biotools. This MS/MS fragment analysis showed that the portion of sequence previously identified in the tag sequences corresponds to the *z+2*-ions series. In addition, other sequence fragment was identified: RCKGY— —RFYFD, which was located close to the N-terminal tail of the protein and corresponding to *C*-ions series (figure 80). Because of the special structure of proline, c- and z-ions do not occur at the N-terminal of proline residues. These fragments require a cross-ring cleavage of the cyclic proline residue (Bruker Daltonics GmbH, 2008). For this reason, the presence of a proline residue in the ShPI-I sequence caused a gap of two residues in the c-ions series. The same procedure was carried out with ShPI-II and results are presented in annex I. However, in this case, only two residues (GR) beside its N-terminal tail as well as an internal aspartic acid residue were identified.

The identification of a cysteine residue within the c-ions series obtained for ShPI-I confirmed the effectiveness of the 1,5 DAN matrix in terms of its disulfide bonds reducing power. This is an important experimental condition to be taken into account because ISD fragmentation stops when a disulfide bridge is found, leading to gaps in the sequence readout across the cyclic structure (Suckau and Resemann, 2003).

In general, ISD fragmentation was found to be a suitable approach for the sequence analysis of serine protease inhibitors known as ShPI-I and ShPI-II obtained from the elution fraction of IF MALDI-TOF MS of the crude extract of *S. helianthus* with immobilized trypsin, since 10 amino acid residues were identified for both molecules near their C-terminal tail as well as 9 amino acid residues were detected for ShPI-I close to its N-terminal tail, representing a total sequence coverage of 36.4%. In the case of ShPI-II, the coverage was 23.6%.

IV.4. Purification and characterization of a carboxypeptidase inhibitor from the marine snail *Nerita versicolor* (NvCI)

In order to purify the CPA inhibitor(s) previously detected in the crude extract of *N. versicolor* (by measuring CPA inhibitory activity and IF MALDI-TOF MS), an initial clarification step by heat treatment at 60°C for 30 min was performed. This clarification procedure is an effective and simple method, which eliminates thermo labile contaminating proteins, considering that the molecule of interest is stable at the tested temperature. It has also the advantage of not requiring the use and subsequent disposal of chemicals (Chávez *et al.*, 1990). This step allowed to increase specific inhibitory activity (1.7-fold) and to recover 98.7% of inhibitory activity (table 36).

After obtained the clarified extract, the strategy used to purify and characterize the responsible molecule(s) of CPA inhibitory activity in *N. versicolor* was the combination of two chromatographic methods such as affinity chromatography and reversed-phase high-performance liquid chromatography (RP-HPLC). It is important to be noted that in each chromatographic step, CPA inhibitory activity was determined in all the peaks.

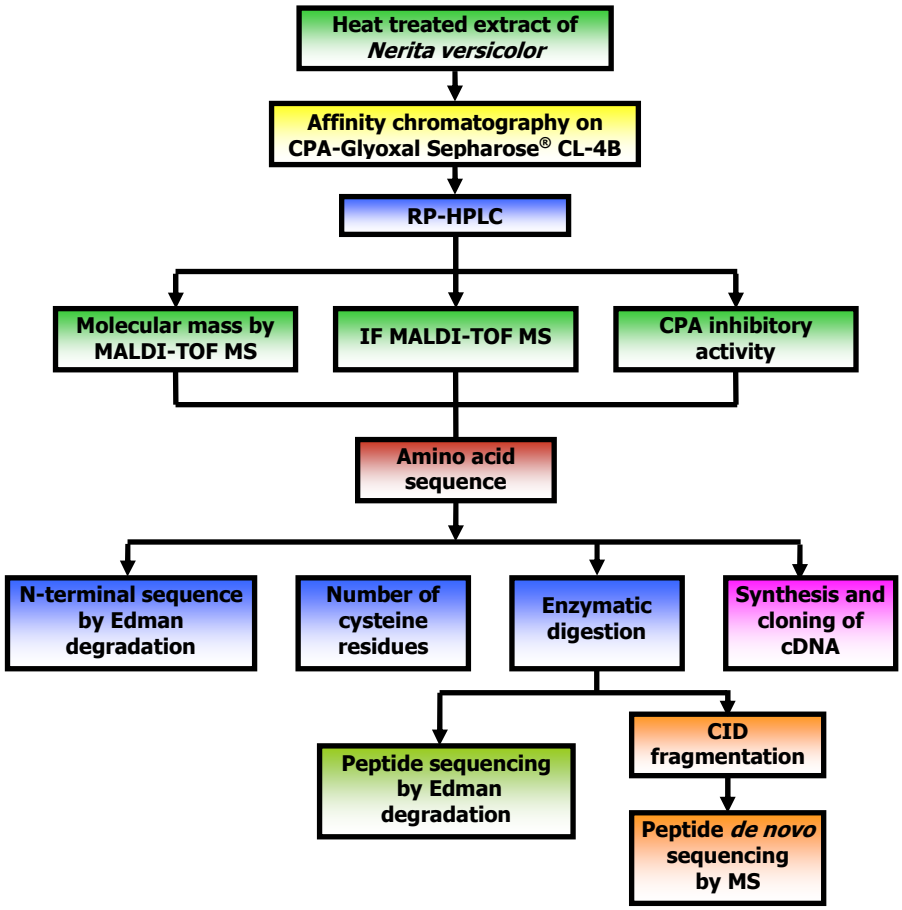


Figure 81. Flowchart process for purification and preliminary characterization of natural NvCI

Once purified the molecule responsible for CPA inhibitory activity, determination of its amino acid sequence was carried out. Reduction and S-carbamidomethylation of the inhibitor allowed the identification of cysteine residues within this molecule. Reduced and S-carbamidopethylated molecule was submitted to a combination of Edman degradation and proteolytic digestion of the protein, followed by *de novo* sequencing of fragments generated by CID MALDI-TOF/TOF MS from the digestion peptides. This strategy led to the elucidation of the amino acid sequence of the *N. versicolor* carboxypeptidase inhibitor called NvCI. The strategy of purification and characterization of NvCI is shown in figure 81.

IV.4.1. Purification of natural NvCI

The selected chromatographic method to start the purification process was an affinity chromatography using CPA-glyoxal Sepharose® CL-4B. Firstly, several affinity matrices with different degrees of immobilization (1.7, 3.5 and 4.4 mg CPA/ml drained gel) were assayed. The best result based on recovery of CPA inhibitory activity was obtained by using a CPA-Sepharose column with a degree of immobilization of 3.5 mg/ml gel.

This result is in agreement with the behavior obtained for the immobilization parameters in CPA-Sepharose derivatives, where the immobilized derivative with 3.5 mg CPA/ml drained gel and 21.4 U/ml gel in terms of CPA activity showed a higher capacity of ligand binding. This is based not only in the lower immobilization degree of the immobilized derivative with 1.7 mg CPA/ml drained gel, and consequently its lower binding capacity, but also in comparison with the immobilized derivative with 4.4 mg CPA/ml drained gel. The later has a higher immobilization degree in terms of protein and enzyme activity but not in terms of retention of functional activity (20.5%), due to mass transfer limitations (table 24). In addition, very high degree of immobilization determines strong binding interactions requiring drastic conditions of elution for tight-binding inhibitors, such as NvCI. Therefore, a derivative with 3.5 mg CPA/ml gel guaranteed a moderate binding and elution.

In order to perform affinity chromatography, heat-treated extract of *N. versicolor* (129 mg total protein) was dissolved in one equivalent volume of Tris-HCl buffer (20 mM, pH 7.5 containing 500 mM NaCl) and applied to a CPA-glyoxal Sepharose column. The presence of moderate salt contents contributes to eliminate non-specific electrostatic interactions (Bertonati *et al*., 2007).

Under these experimental conditions, a chromatographic profile of a typical affinity chromatography was reached, showing a non-retained heterogeneous peak before elution. During the elution stage a major peak with CPA inhibitory activity, called NvCI was displayed (figure 82).

Affinity chromatography allowed to recover 152.5% of the CPA inhibitory activity obtained by heat treatment of the extract (equivalent to 150.6% of the initial CPA inhibitory activity detected in the crude extract), which represents an excellent chromatographic performance.

Yields over 100% in affinity chromatography have been described during the purification of protease inhibitors from marine invertebrates (Delfín *et* al., 1996; González *et al*., 2004; González *et al*., 2007a; Alonso del Rivero *et al*., 2012). This behaviour has been attributed to the dissociation of endogenous inhibitor-protease complexes which were not dissociated during the preparation of the crude extract or by heat treatment. This dissociation is induced by the presence of high concentrations of immobilized CPA on the Sepharose matrix.

On the other hand, 50.9-fold increase of the specific inhibitory activity demonstrates the high selectivity of the designed affinity chromatography (degree of immobilization, experimental conditions of interaction and elution, among others). Experimental factors such as flow velocity (15 cm/h) were crucial to achieve these results, ensuring moderate binding of the inhibitor(s) to its immobilized molecular counterpart (CPA). As a consequence, an elution under moderate conditions was guaranteed, so that both, purified inhibitor(s) and immobilized enzyme retain their functional activity. This is a prerequisite for inhibitors such as NvCI, which as suggested in its identification and demonstrated below (paragraph IV.8.1.4.), is a tight-binding inhibitor against CPA.

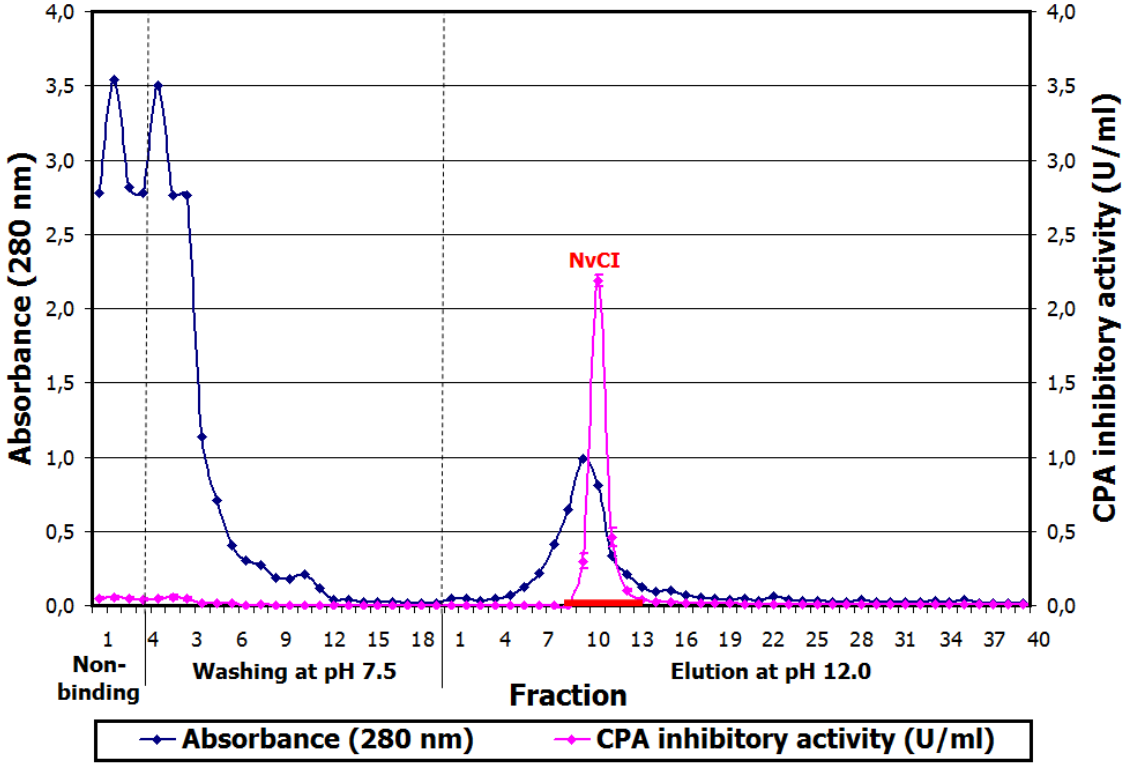


Figure 82. Affinity chromatography of the heat treated extract of *N. versicolor* on a CPA-glyoxal Sepharose® CL-4B column.

I.D. Prot: 3.5 mg CPA/ml gel. Column: 1.6 cm x 9.9 cm. The column was equilibrated at 15 cm/h with 20 mM Tris-HCl buffer, pH 7.5, containing 500 mM NaCl. Sample loading was performed at the same flow rate. Non-retained molecules were removed by washing the column at 30 cm/h with 20 mM Tris-HCl buffer, pH 7.5, containing 500 mM NaCl. Elution was performed at 30 cm/h with 10 mM NaOH, pH 12.0. The whole process was carried out at room temperature.

--◆-- Absorbance at 280 nm, --◆--CPA inhibitory activity (U/ml), ▬ NvCI peak of CPA inhibitory activity

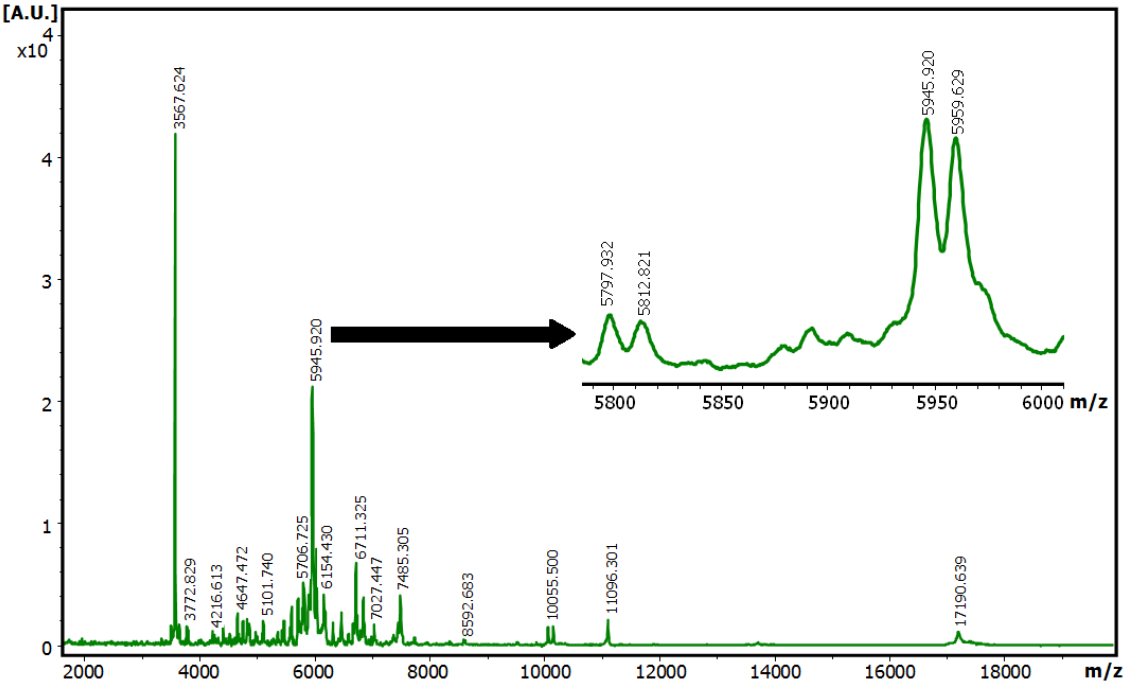


Figure 83. MALDI-TOF MS profile of the elution peak from affinity chromatography.

1 µl of each sample was deposited on a MTP 384 target plate polished steel T F (Bruker Daltonics), followed by deposition of 1 µl of sinapic acid (SA) as a matrix. The mixture was allowed to dry at room temperature.

It is important to note that the elution peak obtained from affinity chromatography is a heterogeneous peak which contains more than one inhibitor of CPA, as demonstrated indirectly by IF MALDI-TOF MS with *N. versicolor* crude extract (figures 60 and 61) and MALDI spectrum of affinity chromatography (figure 83). In this MALDI spectrum three main peaks are displayed: a peak of 3567.8 Da, and two peaks of 5945.920 and 5959.629 Da. Other two minor peaks were visualized at 5797.932 and 5812.821 Da.

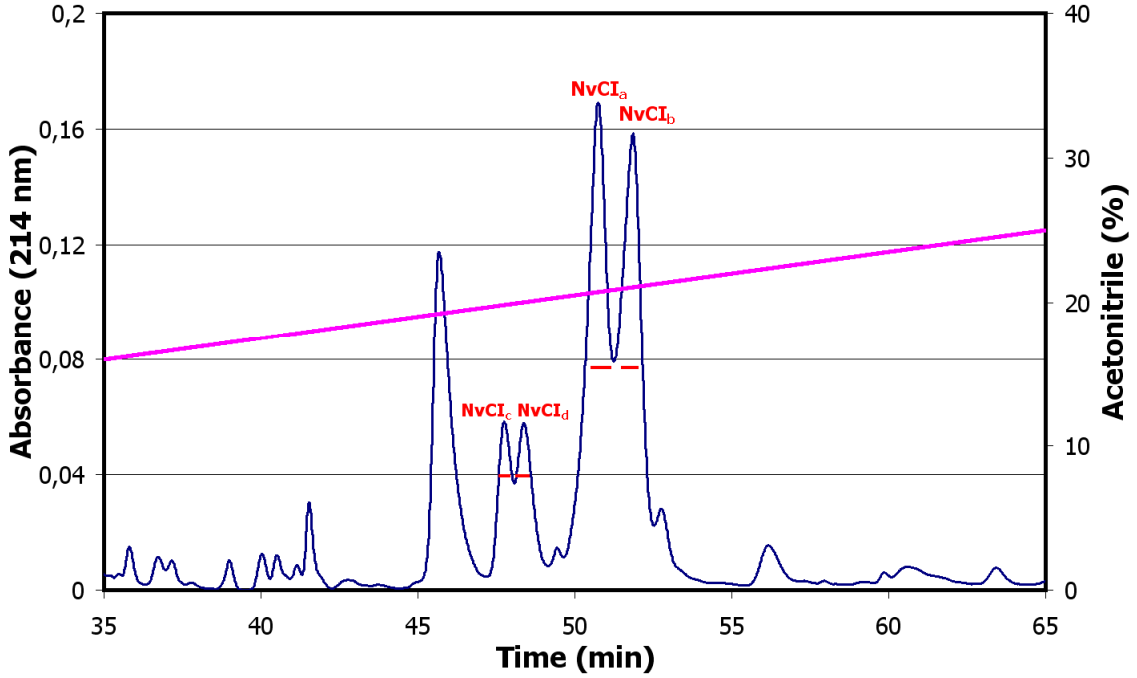


Figure 84. RP-HPLC of the elution peak from affinity chromatography.
C4 (3,9 mm x 150 mm) column. Solution A was 0.1% v/v TFA in water, while solution B was 0.1% v/v TFA in acetonitrile. Elution was performed under the following experimental conditions: At the start, 10% of solution B during 15 min was used, followed by a linear gradient from 10 to 40% over 100 min and a linear gradient from 40 to 98% over 1 min. The flow rate was 0.5 ml/min and room temperature. — Absorbance 214 nm, — NvCIa and b peaks of CPA inhibitory activity, — acetonitrile gradient (%volume)

Figure 84 shows RP-HPLC chromatographic profile obtained of the elution peak from affinity chromatography. This profile confirms the presence of several peaks but only the two major peaks called NvCIa and NvCIb along with other two minor peaks called NvCIc and NvCId displayed CPA inhibitory.

During this purification stage 25.7% recovery of CPA inhibitory activity was obtained for NvCIa, which is a poor performance relative to the initial activity of the extract. These yields are usual in this type of chromatography (Janson and Ryden, 1998). However, it is necessary to take into account that this yield is based on the CPA inhibitory activity of only one component (one NvCI isoform, in this case) with respect to the inhibitory activity of the whole components in the previous steps. On the other hand, a high purification degree is achieved: 4444.2-fold regarding to the initial extract and 87.6-fold compared to affinity chromatography.

Table 36. Summary of the purification procedure of NvCI from the marine snail *N. versicolor*

Step	Total protein (mg)	Inhibitory activity (U)	Specific inhibitory activity (U/mg)	Yield (%)	Purification (-fold)
Crude extract	218.8 ± 9.0	6.5 ± 1.4	0.030 ± 0.007	100.0 ± 0.0	1.0 ± 0.0
Heat treatment	129.0 ± 3.2	6.4 ± 0.3	0.050 ± 0.002	98.7 ± 22.0	1.7 ± 0.4
Peak from affinity chromatography	6.5 ± 0.4	9.8 ± 0.3	1.51 ± 0.04	150.6 ± 33.3	50.9 ± 11.2
NvCIa peak from RP-HPLC	0.013 ± 0.001	1.7 ± 0.1	132.3 ± 7.7	25.7 ± 5.8	4444 ± 1008

Data are means (n=3) ± S.D.

Based on the results, it should be noted that the procedure comprising a heat treatment of the extract at 60°C for 30 min, affinity chromatography and reversed-phase HPLC under the experimental conditions employed, allows the purification of NvCI (NvCIa, NvCIb, NvCIc and NvCId isoforms) with a high degree of purity from the marine snail *N. versicolor*.

IV.4.2. Molecular characterization of natural NvCI

IV.4.2.1. Molecular mass determination of NvCI by MALDI-TOF MS

MALDI spectra of the two major peaks obtained from RP-HPLC confirm the purity of each sample. Single peaks at molecular masses of 5945.997 Da, 5959.853 Da, 5798.289 Da and 5812.535 Da, corresponding to NvCIa, NvCIb, NvCIc and NvCId, respectively, were visualized in the spectra (figures 85 to 88).

All purified peaks, NvCIa, NvCIb, NvCIc and NvCId were analyzed by automatic Edman degradation, resulting in the same sequence at the first 8 amino acid residues analyzed for NvCIa and NvCIb whilst NvCIc and NvCId displayed the same protein sequence of NvCIa and NvCIb with the absence of the first amino acid residue in these two later isoforms. In addition, the determination of the apparent inhibition constant (K_{iapp}) against bCPA and pCPB demonstrated that all these four NvCI isoforms displayed the same functional behaviour (Annex II).

The primary structure of the four isoforms of NvCI was analyzed, but during the course of this research work, the peak corresponding to NvCIa was called NvCI and it was the selected molecule used in the further structural and functional characterization. In addition, the amino acid sequence of this isoform was used in the subsequent gene design and recombinant production of this MCPs inhibitor.

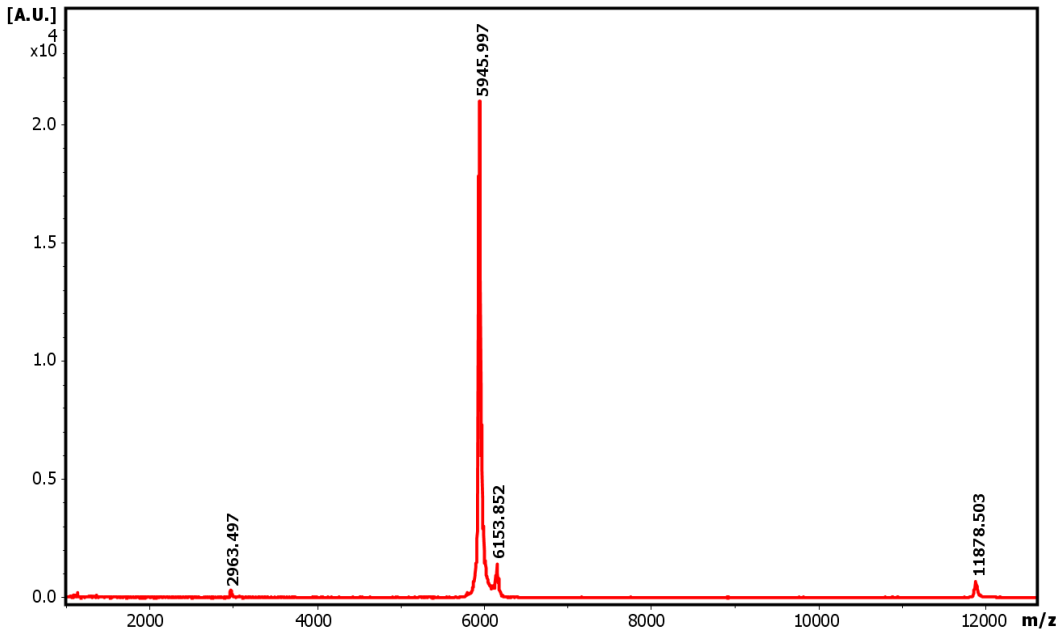


Figure 85. MALDI-TOF spectrum of a Carboxypeptidase A inhibitor purified from the marine snail *Nerita versicolor* (NvCIa). 1 µl of sample was deposited on a MTP 384 target plate polished steel T F (Bruker Daltonics), followed by deposition of 1 µl of DHAP as a matrix. The mixture was allowed to dry at room temperature.

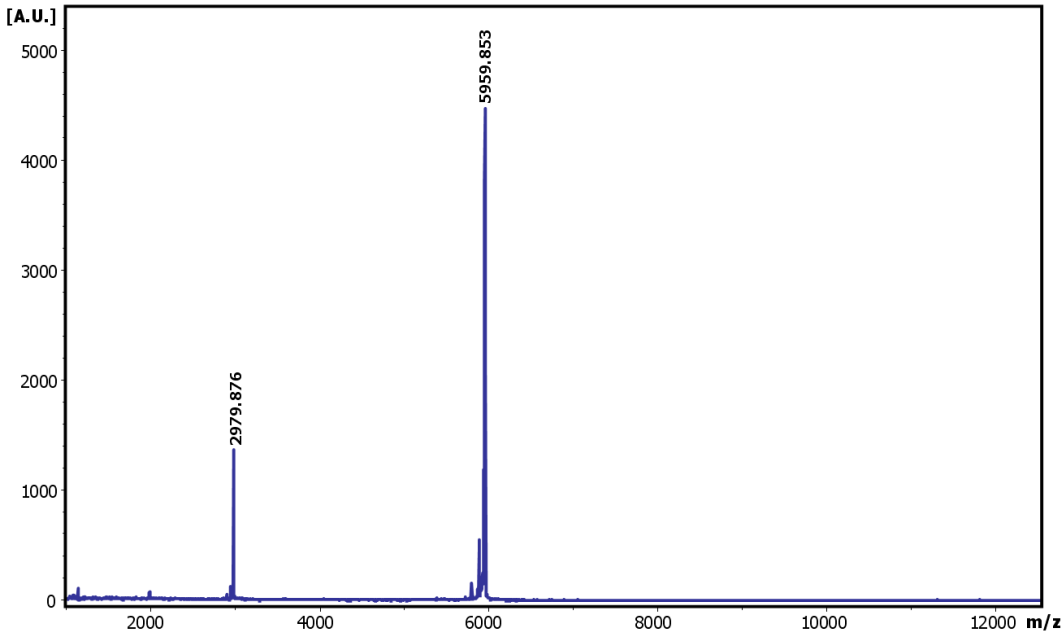


Figure 86. MALDI-TOF spectrum of a Carboxypeptidase A inhibitor purified from the marine snail *Nerita versicolor* (NvCIb). 1 µl of sample was deposited on a MTP 384 target plate polished steel T F (Bruker Daltonics), followed by deposition of 1 µl of DHAP as a matrix. The mixture was allowed to dry at room temperature.

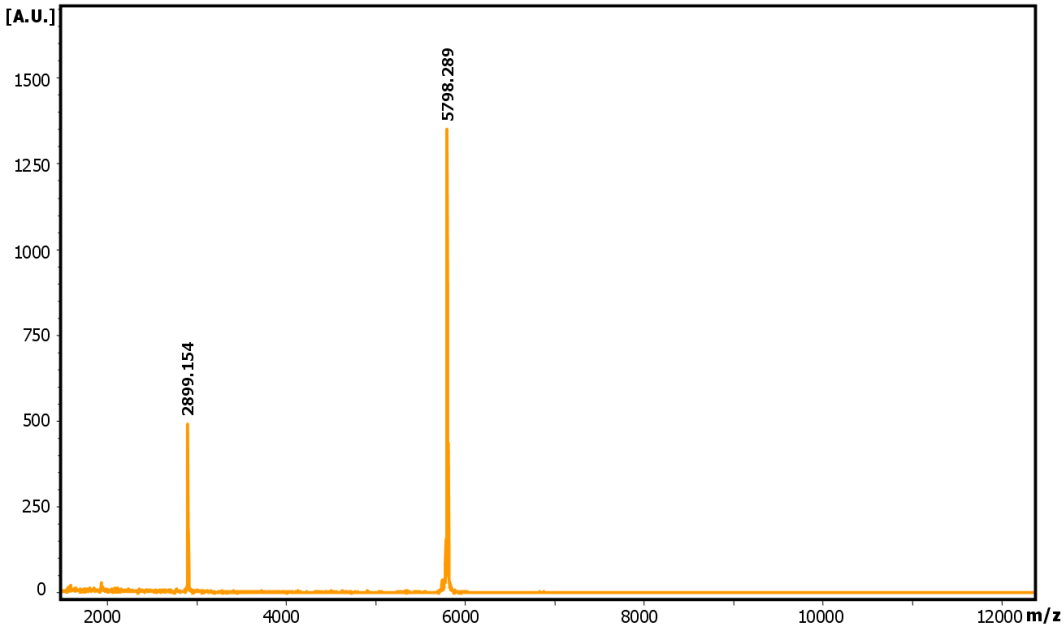


Figure 87. MALDI-TOF spectrum of a Carboxypeptidase A inhibitor purified from the marine snail *Nerita versicolor* (NvCIc). 1 µl of sample was deposited on a MTP 384 target plate polished steel T F (Bruker Daltonics), followed by deposition of 1 µl of DHAP as a matrix. The mixture was allowed to dry at room temperature.

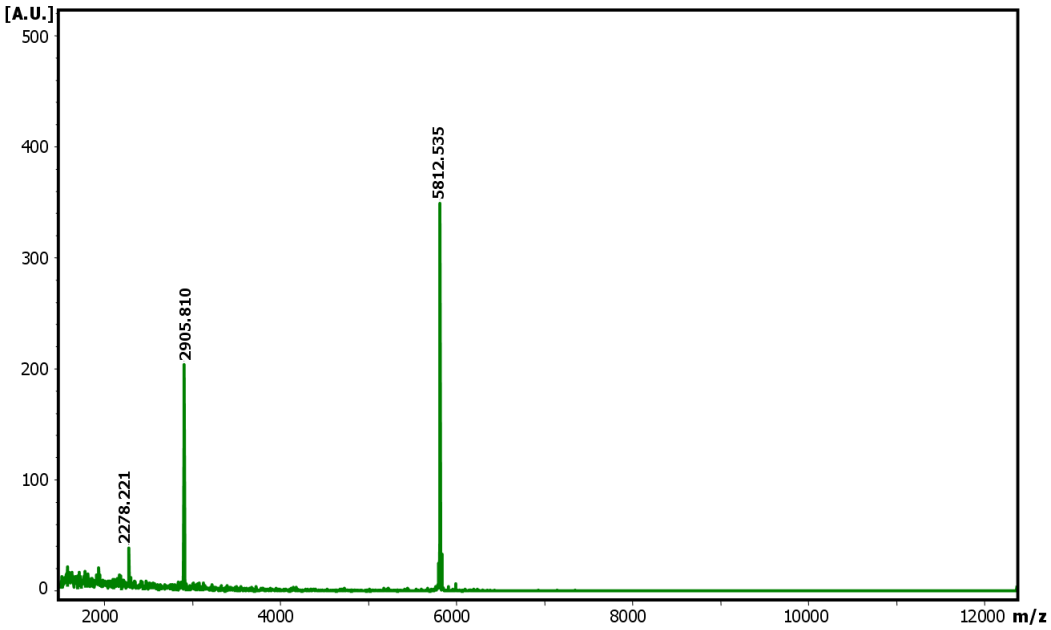
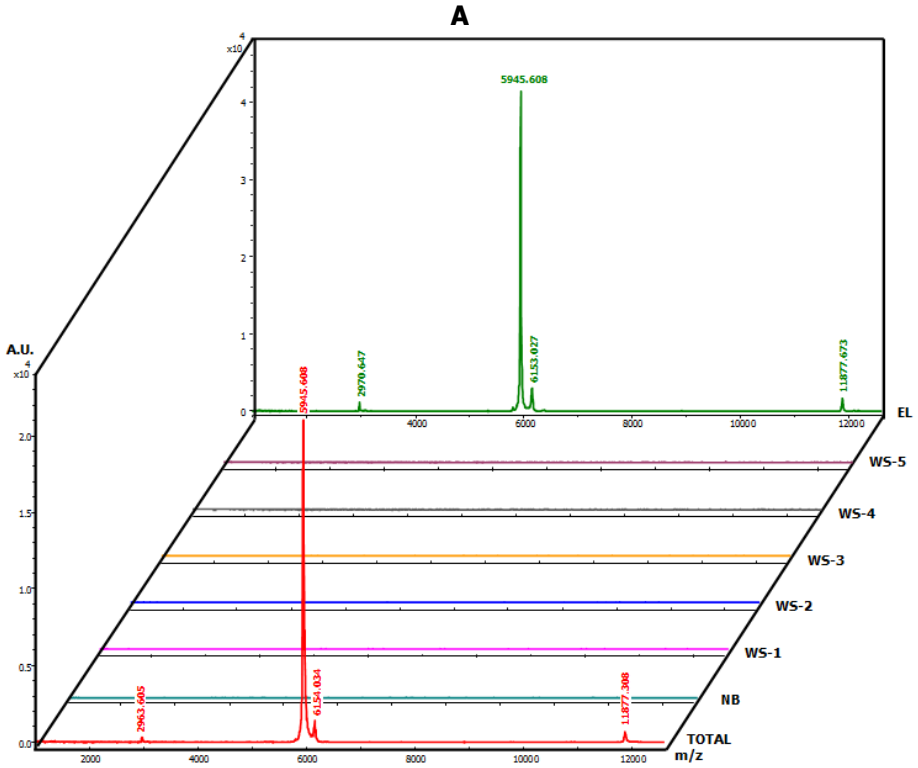


Figure 88. MALDI-TOF spectrum of a Carboxypeptidase A inhibitor purified from the marine snail *Nerita versicolor* (NvCId). 1 µl of sample was deposited on a MTP 384 target plate polished steel T F (Bruker Daltonics), followed by deposition of 1 µl of DHAP as a matrix. The mixture was allowed to dry at room temperature.

Affinity interaction assays by IF MALDI-TOF MS between purified NvCI and immobilized CPA demonstrated not only sample purity, but also functional activity of the inhibitor through its molecular recognition of the immobilized protease. As shown in figure 88A, once the sample was allowed to interact with the affinity matrix, peak corresponding to NvCI completely disappeared from the mass spectrum, indicating its total retention. During the subsequent washing steps any released molecule was observed, and finally, after dissociation of enzyme-inhibitor complex, signal at m/z 5945.6 was recovered in this elution fraction. Control assay using Sepharose did not evidence non-specific interactions between NvCI and the support (figure 88B).



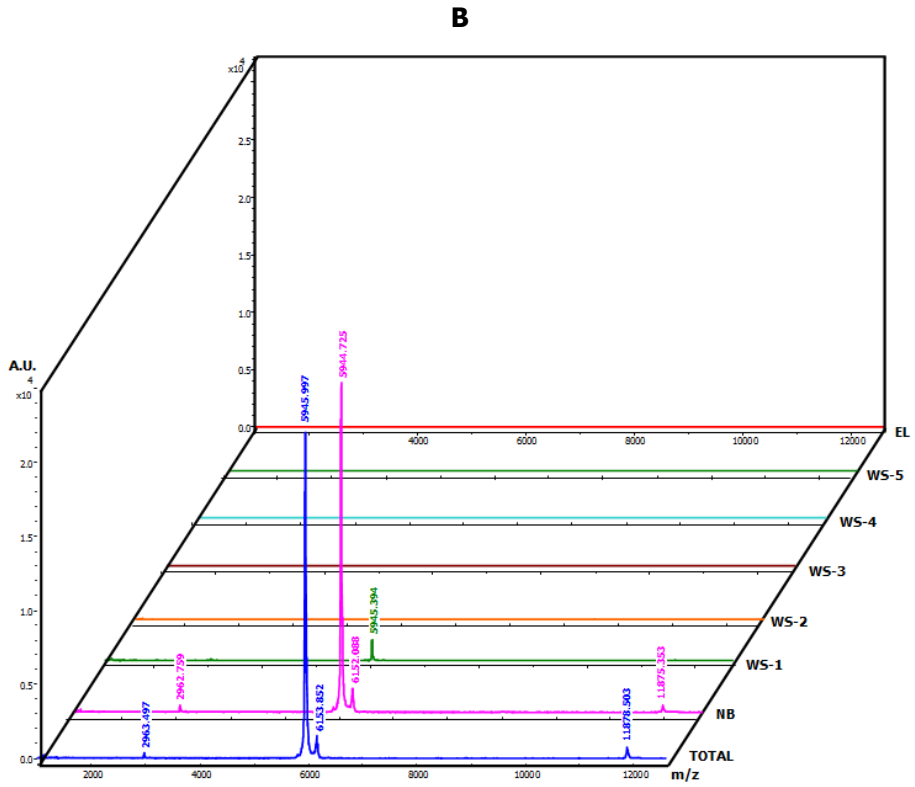


Figure 89. Intensity Fading MALDI-TOF MS of natural NvCI on CPA-glyoxal Sepharose® CL-4B (A) and Sepharose® CL-4B (B). TOTAL: NvCI in 20 mM Tris-HCl buffer pH 7.5, 500 mM NaCl. **NB:** non-binding molecules, interaction at pH 7.5, 10 min. **WS-1 to WS-5:** washing at pH 7.5. **EL:** elution with 0.5% v/v TFA, 10 min. Interaction assays were carried out in duplicate at room temperature. The experimental conditions of MALDI-TOF MS are described in section III.7.2**.**

IV.4.2.2. Determination of the amino acid sequence of NvCI

The strategy used to determine the amino acid sequence of NvCI involved an initial analysis of the whole protein by automatic Edman degradation in order to obtain its N-terminal sequence. Denaturation, reduction and S-carbamidomethylation of cysteine residues were subsequently performed, allowing the identification of cysteine residues. Disulfide bonds were estimated from S-carbamidomethylation of the non reduced protein. Once NvCI was denatured, reduced and alkylated, enzymatic digestions of the protein with different proteases were carried out. Digestion products were separated by RP-HPLC and sequenced by automatic Edman degradation. These peptides were further fragmented by CID MALDI-TOF and sequenced *de novo*. Overlapping of the determined sequences belonging to digestion peptides allowed the establishment of the complete sequence of NvCI. On the other hand, the synthesis and cloning of cDNA encoding NvCI allowed to coroborate the amino acid sequence of this molecule.

IV.4.2.2.1. Determination of the N-terminus amino acid sequence

Table 37. N-terminus sequence of NvCI by Edman degradation

F	H	V	P	D	D	R	P	C	I	N	P	G	R	C	P	L	V
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18

Automatic Edman degradation allowed the identification of the first 18 amino acid residues of NvCI (table 37). The presence of 2 cysteine residues in this segment led to determine the total number of cysteine residues as well as disulfide bonds into the molecule.

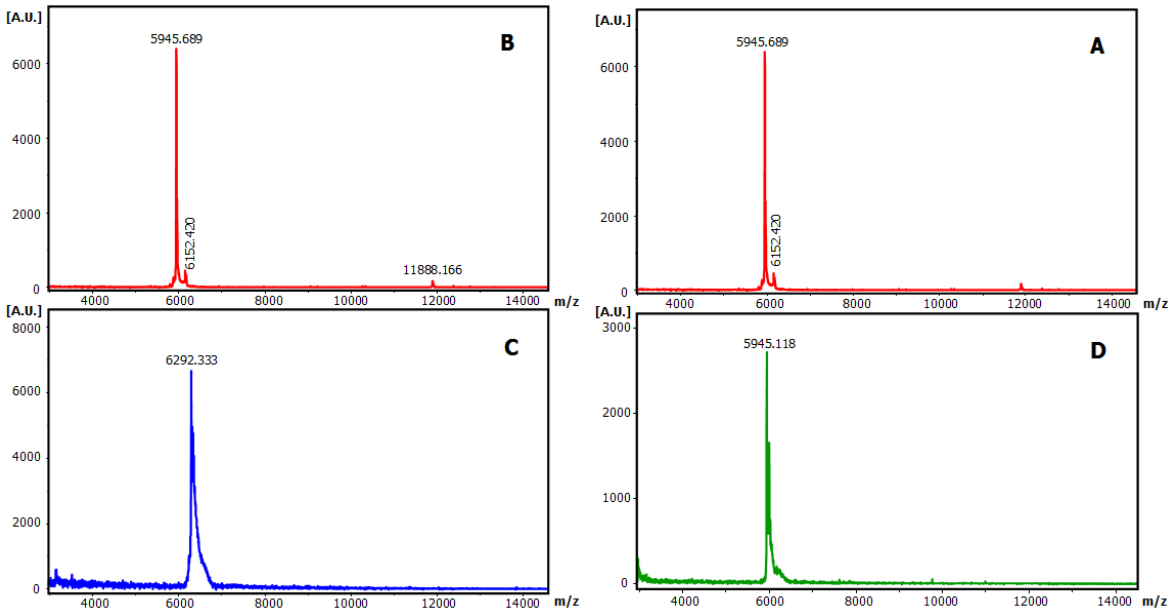


Figure 90. Determination of total and free cysteine residues in NvCI.
A. and **B.** MALDI-TOF spectrum of NvCI. **C.** MALDI-TOF spectrum after reduced and S-carbamidomethylated NvCI. **D.** MALDI-TOF spectrum of S-carbamidomethylated NvCI. 1 µl of each sample was deposited on a MTP 384 target plate polished steel T F (Bruker Daltonics), followed by deposition of 1 µl of DHAP as a matrix. The mixture was allowed to dry at room temperature.

IV.4.2.2.2. Determination of the number of cysteine residues

Reduced and alkylated molecule was analyzed by MALDI-TOF MS and compared to the native molecule (figure 90). MALDI spectrum indicated that reduced and S-carbamidomethylated NvCI displayed a molecular mass of 6292.3 Da, which is higher than that obtained for the native protein (5945.7) Da. This molecular weight difference divided by 57 Da, corresponding to the alkylation showed the presence of 6 cysteine residues in NvCI. MALDI spectrum of the non reduced molecule after S-carbamidomethylation showed a signal with the same molecular mass of the native NvCI. This result suggests that all 6 cysteine residues in NvCI are involved in disulfide bonds.

IV.4.2.2.3. Amino acid sequences of the products of NvCI hydrolysis with different proteases determined by automated Edman degradation and *de novo* sequencing

Enzymatic hydrolysis of NvCI was performed by using the endoproteinases Lys-C, trypsin and Glu-C. MALDI-TOF spectrum of Lys-C digestion products is shown in figure 91, where 5 major peaks at m/z 2980.5 and 3037.6 along with other minor peaks at m/z 2909.5, 3094.7, 3328.9 and 3387.0 were identified. It is important to note that the molecular weight difference between these two major peaks is 57 Da, which could be related to overalkylation of the molecule.

Digestion peptides were subsequently purified by RP-HPLC and the major peak of 2980 Da was analyzed by automatic Edman degradation. This peptide was also fragmented by CID MALDI-TOF MS and analyzed by *de novo* sequencing (figure 92). This strategy was also performed with trypsin and Glu-C and their results are shown in figures 93 to 96, as well as in annex III.

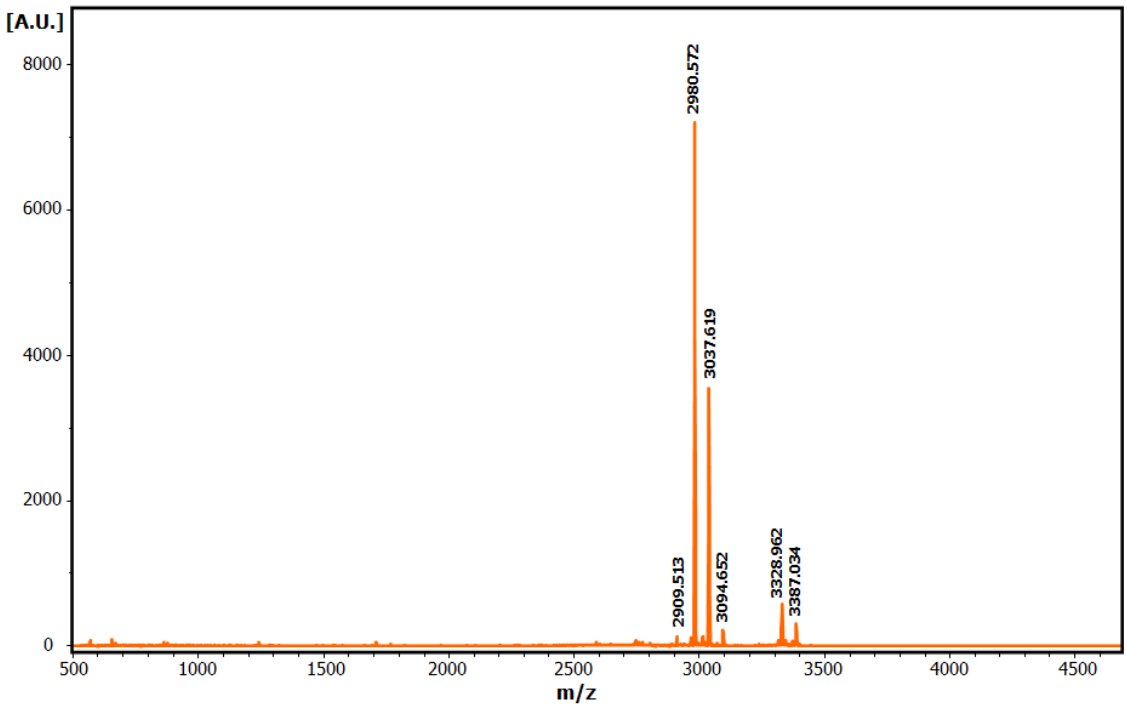


Figure 91. MALDI-TOF spectrum of NvCI enzymatic digestion with Lys-C.
Enzymatic digest was desalted using ZipTip®C_4 pipette tips (millipore). Then, 1 µl of desalted sample was deposited on a MTP 384 target plate polished steel T F (Bruker Daltonics), followed by deposition of 1 µl of α-CHCA as a matrix. The mixture was allowed to dry at room temperature.

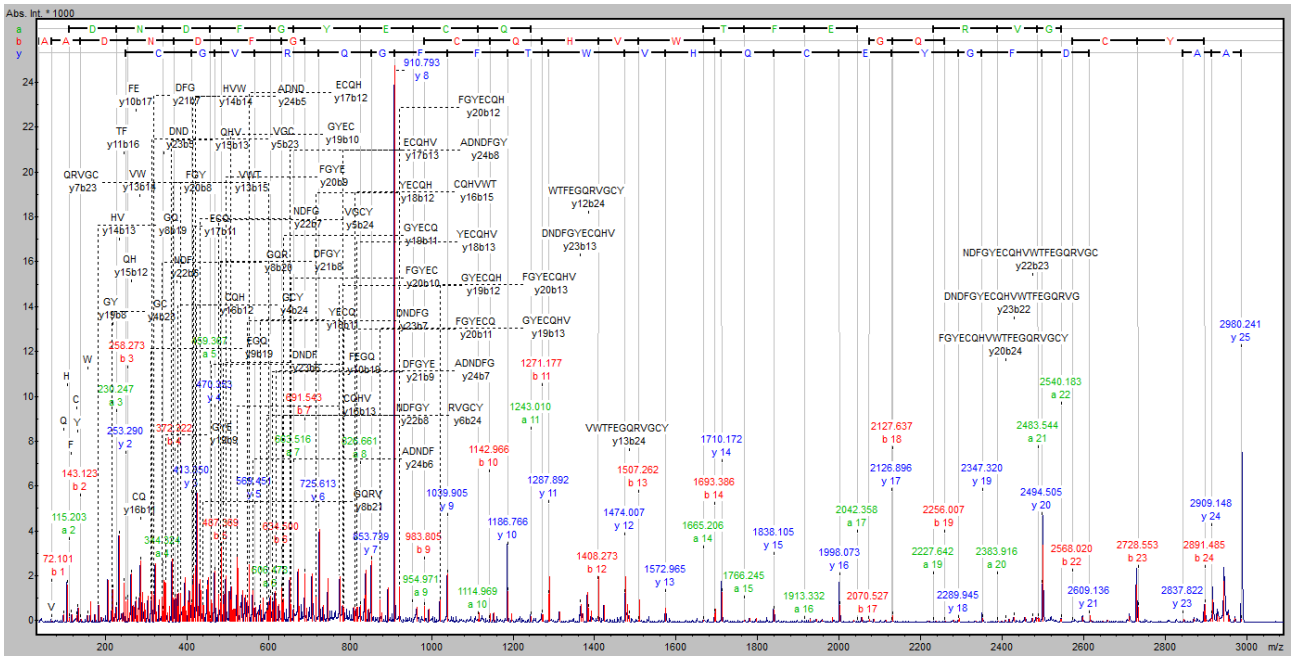


Figure 92. MALDI MS/MS spectrum of parent m/z=2980 from NvCI enzymatic digestion with Lys-C.
Sequence derived from MS/MS spectrum by *de novo* sequencing: AADNDFGYECQHVWTFEGQRVGCYA

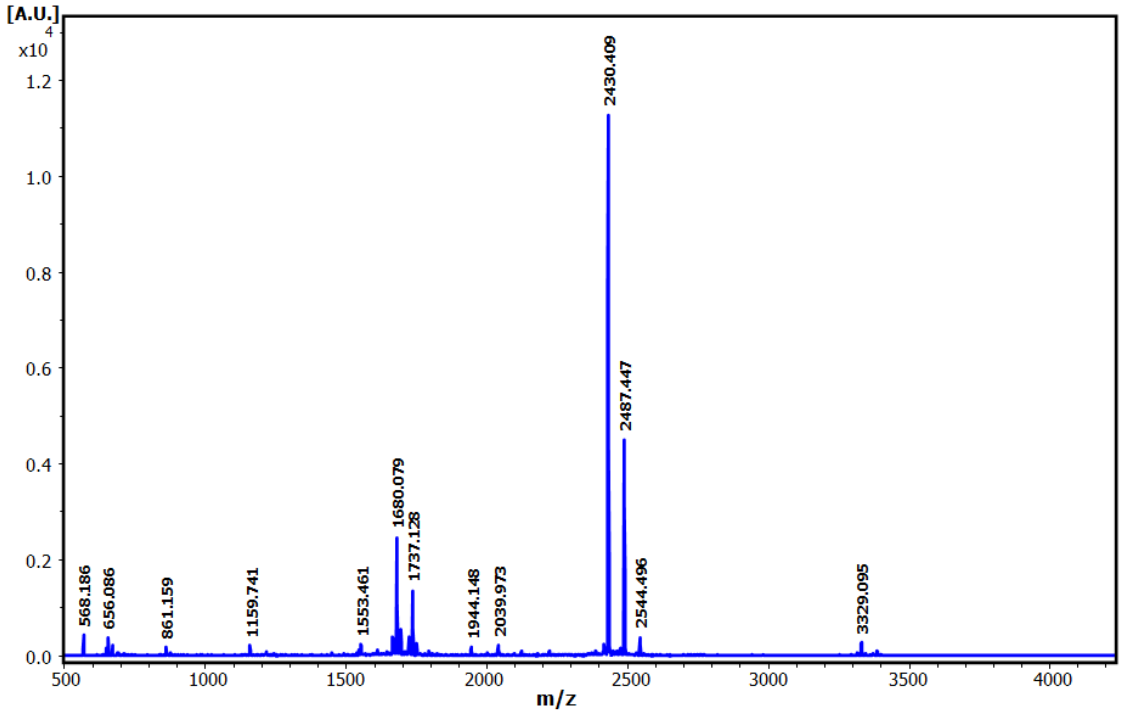


Figure 93. MALDI-TOF spectrum of NvCI enzymatic digestion with trypsin.
Enzymatic digest was desalted using ZipTip®C_4 pipette tips (millipore). Then, 1 µl of desalted sample was deposited on a MTP 384 target plate polished steel T F (Bruker Daltonics), followed by deposition of 1 µl of α-CHCA as a matrix. The mixture was allowed to dry at room temperature.

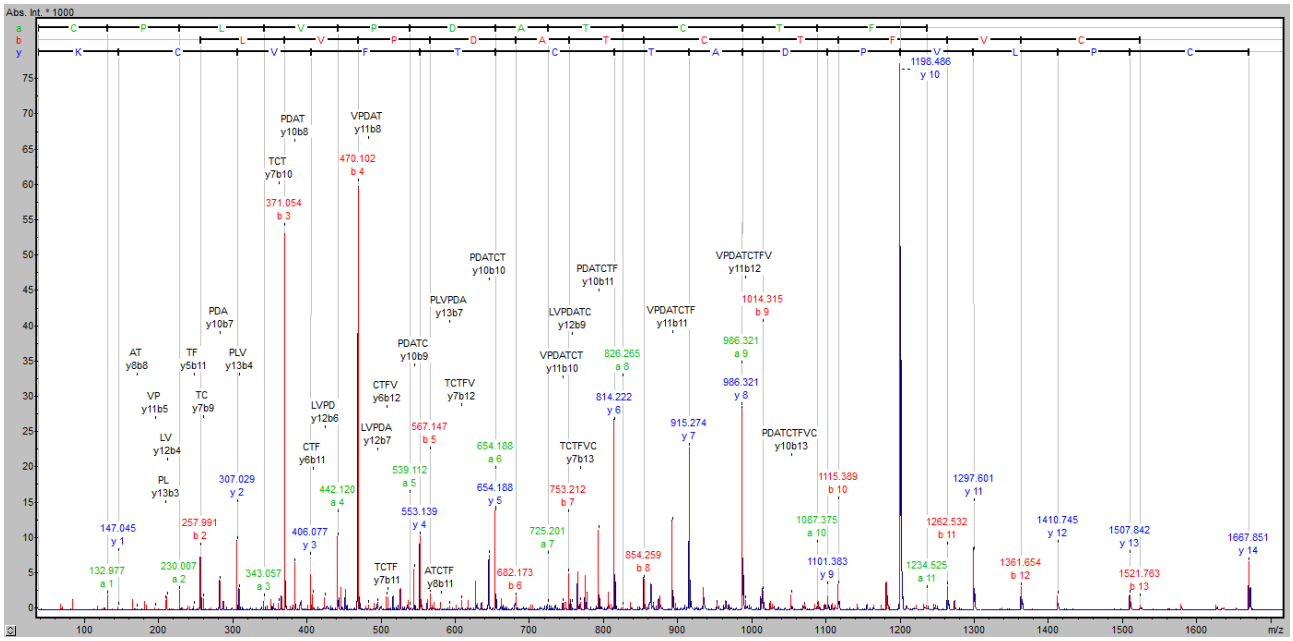


Figure 94. MALDI MS/MS spectrum of parent m/z=1668 from NvCI enzymatic digestion with trypsin.
Sequence derived from MS/MS spectrum by *de novo* sequencing: CPLVPDATCTFVCK

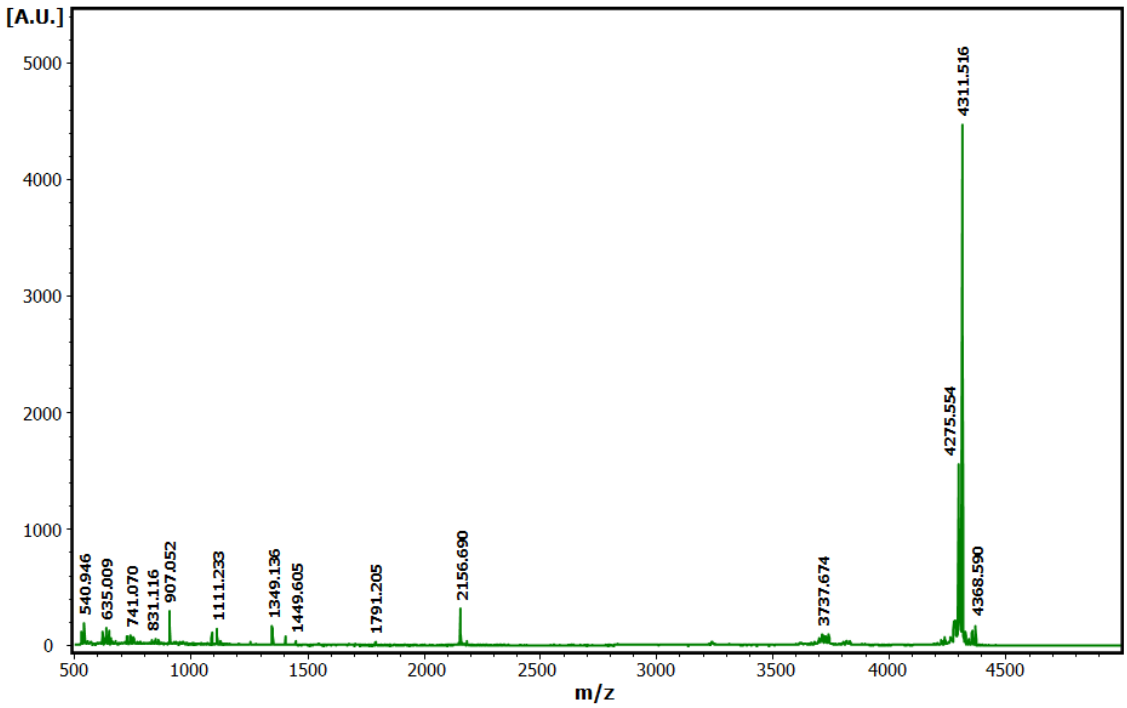


Figure 95. MALDI-TOF spectrum of NvCI enzymatic digestion with Glu-C.
Enzymatic digest was desalted using ZipTip®C_4 pipette tips (millipore).). Then, 1 µl of desalted sample was deposited on a MTP 384 target plate polished steel T F (Bruker Daltonics), followed by deposition of 1 µl of α-CHCA as a matrix. The mixture was allowed to dry at room temperature.

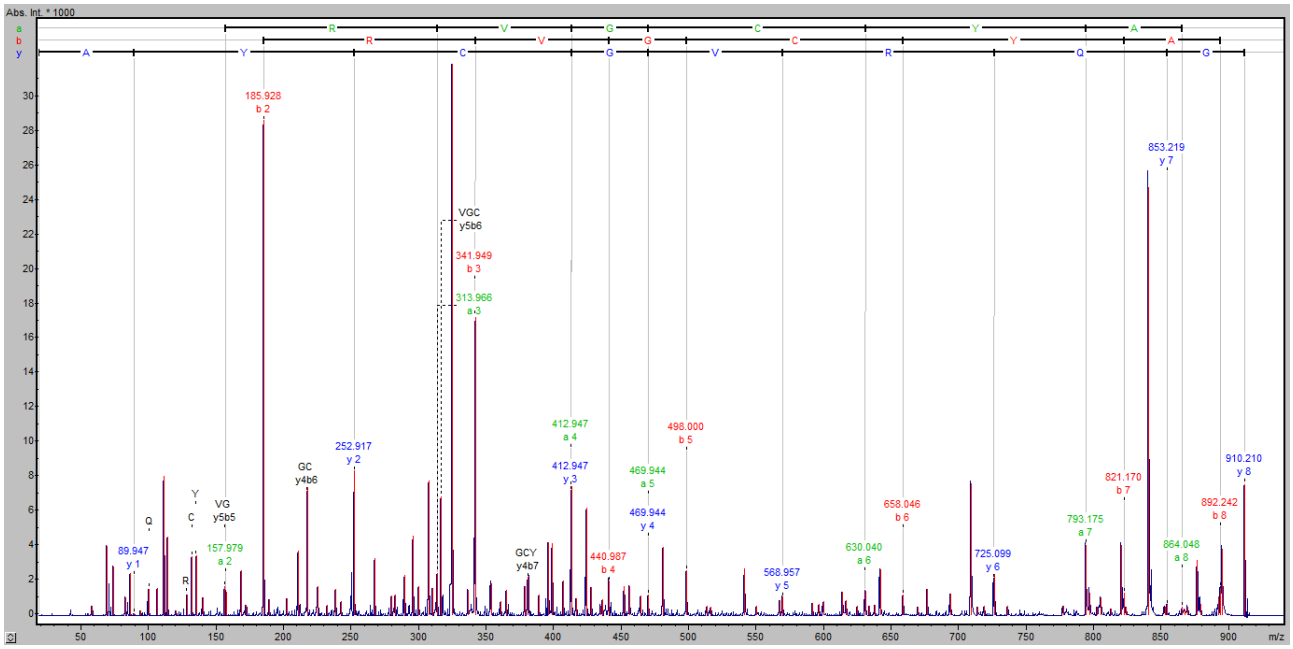


Figure 96. MALDI MS/MS spectrum of parent m/z=910 from NvCI enzymatic digestion with Glu-C.
Sequence derived from MS/MS spectrum by *de novo* sequencing: GQRVGCYA

IV.4.2.2.4. Synthesis and cloning of cDNA encoding NvCI

Starting from the total RNA extracted from the body of *N. versicolor* (1351 ng RNA/µl), the cDNA encoding NvCI by 3'-RACE was obtained as described in paragraph III.9.2.4.3. Initially, the first chain of cDNA was synthesized by RT-PCR using the R_0R_1polydT primer. Then, the cDNA chain obtained in the previous step was used as template DNA in a PCR using the P2_NvCI$_{1\text{-}8}$ and R_0 primers. PCR products are shown in figure 97A, where in lane 1 are displayed two clear bands of about 170-bp and another at approximately 370-bp as well as two thin bands around 500 and 600-bp. All these bands were cut, purified from agarose gel and sequenced using the P2_NvCI$_{1\text{-}8}$ primer. Sequences obtained showed that the band of approximately 370-bp contained the cDNA encoding NvCI (marked with red box in figure 97A).

In figure 97B are shown the results obtained for RT-PCR controls. In lane 1, a DNA band at approximately 396 bp was visualized for the positive control of RT-PCR. This result is in agreement with that described by the manufacturers of RT-PCR kit, where the amplification by PCR of a 382-bp fragment has been determined. Result for negative control of RT-PCR is displayed in lane 2, in which a very diffuse band less than 75-bp was observed, possibly corresponding to the primers used in the PCR.

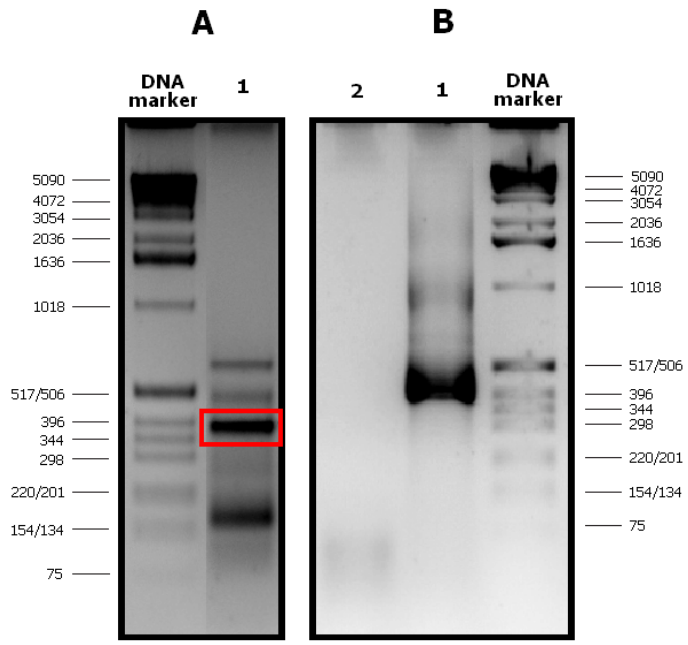


Figure 97. Analysis of PCR products by 2% w/v agarose gel.
(A) Lane 1: PCR products obtained using the template cDNA from the body of *N. versicolor*. The band marked with red box represents the cDNA encoding NvCI. **(B)** Lane 1: Positive control of RT-PCR. Lane 2: Negative control of the RT-PCR

cDNA band encoding NvCI was cloned into *E.coli* XL1-Blue cells using the pBE vector, as shown in figure 98. Insertion of the cDNA of interest in clones of white colonies was verified by PCR using R_0 and P2_NvCI$_{1\text{-}8}$ primers, followed by analysis in 2% w/v agarose gel to visualize the reaction products.

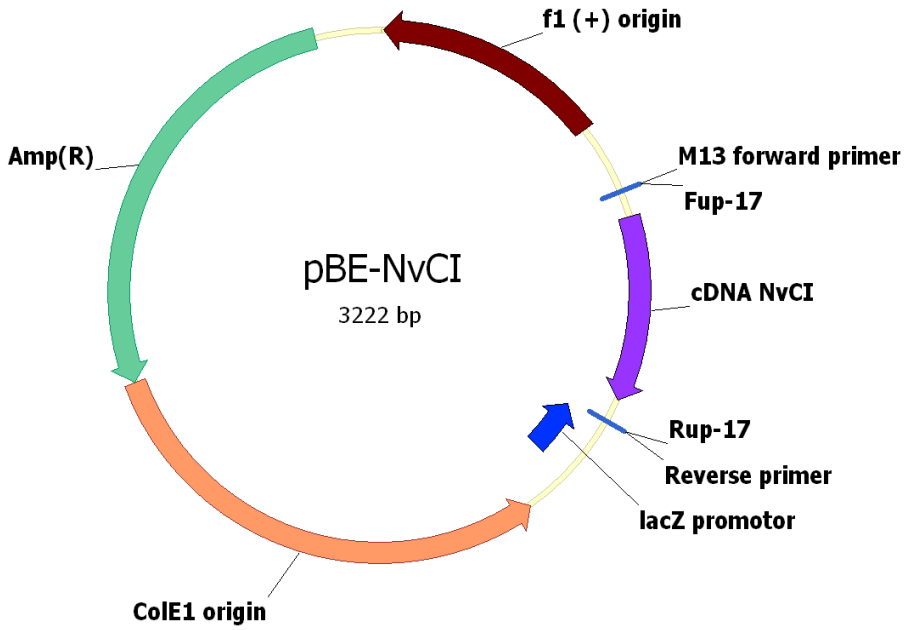


Figure 98. pBE vector used in the cloning of cDNA encoding NvCI

Eight clones that amplified by PCR one band at approximately 370-bp were selected to be sequenced. DNA sequencing showed that 3 clones incorporated the cDNA encoding NvCI. The analysis of chromatograms using the DNA Baser software (Heracle BioSoft S.R.L., Romania) allowed to obtain a consensus sequence. Subsequently, a multiple sequence alignment in the ClustalX version 2.0 software was performed, as shown in figure 99.

This consensus sequence of the gene encoding NvCI consists of 372 nucleotides. The first 180 nucleotides, just before the location of the stop codon "tga", encoding a polypeptide chain revealed the amino acid sequence of NvCI from the first amino acid residue until the Alanine-53 at the C-terminal tail of the protein (printed in magenta, figure 100). This sequence coincided with that previously obtained by Edman degradation and de *novo sequencing* by MALDI-TOF MS, with the exception of the 11th amino acid residue, where a change of asparagine (N) by a lysine (K) residue at this position was detected, impliying a difference of 14 Da over NvCI, corresponding to the isoform of this molecule previously called NvCIb. At genomic level, it can be observed in the triplet encoding a lysine residue (aag) at the 11th position (figure 100), found in all the analyzed clones, one mutation in the last nucleotide leading to the translation of an asparagine residue, which is in agreement with the isoform called NvCI.

This point mutation has been described by Collins and Jukes, 1994, in which the most common one is the transition that exchanges a purine for another purine (A↔G) or a pyrimidine for a pyrimidine (C↔T) whilst a less common one is the transversion that exchanges a purine for a pyrimidine or pyrimidine for a purine (C/T ↔ A/G). Thus, a point mutation in the fundamental DNA encoding NvCI or errors during transcription leads to the production of two isoforms of this molecule: NvCI (NvCIa) and NvCIb.

Into the consensus sequence the canonical polyadenylation signal "aataaa" (underlined in figure 100), and located 13 nucleotides upstream from the poly(A) tail itself, consisting of 23 adenine nucleotides was also identified.

```
Clone_A_Rup      TAGGGCGAATTGGGTACGGAT----ACGTCCCGGATGACAGGCCGTGCATTAAGCCAGGC 56
Clone_A_Fup      TAGGGCGA-TTGGGTACGGAT----ACGTCCCGGATGACAGGCCGTGCATTAAGCCAGGC 55
Clone_B_Rup      TAGGGCGAATTGGGTACGGATTTTCACGTGCCAGACGATAGGCCGTGCATTAAGCCAGGC 60
Clone_B_Fup      TAGGGCGA-TTGGGTACGGATTTTCACGTGCCAGACGATAGGCCGTGCATTAAGCCAGGC 59
Clone_C_Fup      TAGGGCGA-TTGGGTACGGATTC-CACGTGCCGGATGACAGGCCGTGCATTAAGCCAGGC 58
                 ******** ***********    **** ** ** ** ********************
Consensus        TAGGGCGAATTGGGTACGGATTTTCACGTGCCGGACGACAGGCCGTGCATTAAGCCAGGC 60

Clone_A_Rup      CGGTGCCCTCTAGTTCCTGATGCCACGTGCACATTCGTCTGCAAAGCAGCGGACAACGAC 116
Clone_A_Fup      CGGTGCCCTCTAGTTCCTGATGCCACGTGCACATTCGTCTGCAAAGCAGCGGACAACGAC 115
Clone_B_Rup      CGGTGCCCTCTAGTTCCTGATGCCACGTGCACATTCGTCTGCAAAGCAGCGGACAACGAC 120
Clone_B_Fup      CGGTGCCCTCTAGTTCCTGATGCCACGTGCACATTCGTCTGCAAAGCAGCGGACAACGAC 119
Clone_C_Fup      CGGTGCCCTCTAGTTCCTGATGCCACGTGCACATTCGTCTGCAAAGCAGCGGACAACGAC 118
                 ************************************************************
Consensus        CGGTGCCCTCTAGTTCCTGATGCCACGTGCACATTCGTCTGCAAAGCAGCGGACAACGAC 120

Clone_A_Rup      TTTGGCTATGAATGCCAGCACGTTTGGACCTTTGAGGGCCAACGGGTGGGCTGCTATGCC 176
Clone_A_Fup      TTTGGCTATGAATGCCAGCACGTTTGGACCTTTGAGGGCCAACGGGTGGGCTGCTATGCC 175
Clone_B_Rup      TTTGGCTATGAATGCCAGCACGTTTGGACCTTTGAGGGCCAACGGGTGGGCTGCTATGCC 180
Clone_B_Fup      TTTGGCTATGAATGCCAGCACGTTTGGACCTTTGAGGGCCAACGGGTGGGCTGCTATGCC 179
Clone_C_Fup      TTTGGCTATGAATGCCAGCACGTTTGGACCTTTGAGGGCCAACGGGTGGGCTGCTATGCC 178
                 ************************************************************
Consensus        TTTGGCTATGAATGCCAGCACGTTTGGACCTTTGAGGGCCAACGGGTGGGCTGCTATGCC 180

Clone_A_Rup      TGAGAGACAGACAGGGAGAGGGAGGGAGTCAGCTCCAAGCTAGCTGCCATGCCTCCCCCT 236
Clone_A_Fup      TGAGAGACAGACAGGGAGAGGGAGGGAGTCAGCTCCAAGCTAGCTGCCATGCCTCCCCCT 235
Clone_B_Rup      TGAGAGACAGACAGGGAGAGGGAGGGAGTCAGCTCCAAGCTAGCTGCCATGCCTCCCCCT 240
Clone_B_Fup      TGAGAGACAGACAGGGAGAGGGAGGGAGTCAGCTCCAAGCTAGCTGCCATGCCTCCCCCT 239
Clone_C_Fup      TGAGAGACAGACAGGGAGAGGGAGGGAGTCAGCTCCAAGCTAGCTGCCATGCCTCCCCCT 238
                 ************************************************************
Consensus        TGAGAGACAGACAGGGAGAGGGAGGGAGTCAGCTCCAAGCTAGCTGCCATGCCTCCCCCT 240

Clone_A_Rup      CTTGCTCCTGCCTCCATCACAGACAGCCTGGAGCCCAGTCCACAGGCGTCCTTCAAGTTG 296
Clone_A_Fup      CTTGCTCCTGCCTCCATCACAGACAGCCTGGAGCCCAGTCCACAGGCGTCCTTCAAGTTG 295
Clone_B_Rup      CTTGCTCCTGCCTCCATCACAGACAGCCTGGAGCCCAGTCCACAGGCGTCCTTCAAGTTG 300
Clone_B_Fup      CTTGCTCCTGCCTCCATCACAGACAGCCTGGAGCCCAGTCCACAGGCGTCCTTCAAGTTG 299
Clone_C_Fup      CTTGCTCCTGCCTCCATCACAGACAGCCTGGAGCCCAGTCCACAGGCGTCCTTCAAGTTG 298
                 ************************************************************
Consensus        CTTGCTCCTGCCTCCATCACAGACAGCCTGGAGCCCAGTCCACAGGCGTCCTTCAAGTTG 300

Clone_A_Rup      TGTTGGGAAGTCTGTCCTGAAGACCCCGTCAATAAAGATCCCTCTCCTGAAA---AAAAA 353
Clone_A_Fup      TGTTGGGAAGTCTGTCCTGAAGACCCCGTCAATAAAGATCCCTCTCCTGAAA---AAAAA 352
Clone_B_Rup      TGTTGGGAAGTCTGTCCTGAAGACCCCGTCAATAAAGATCCCTCTCCTGAAACGCAAAAA 360
Clone_B_Fup      TGTTGGGAAGTCTGTCCTGAAGACCCCGTCAATAAAGATCCCTCTCCTGAAACGCAAAAA 359
Clone_C_Fup      TGTTGGGAAGTCTGTCCTGAAGACCCCGTCAATAAAGATCCCTCTCCTGAAAAAAAAAAA 358
                 ************************************************   *****
Consensus        TGTTGGGAAGTCTGTCCTGAAGACCCCGTCAATAAAGATCCCTCTCCTGAAAAAAAAAAA 360

Clone_A_Rup      AAAAAAAAAA--- 363
Clone_A_Fup      AAAAAAAAAA--- 362
Clone_B_Rup      AAAAAAAAAAA-- 371
Clone_B_Fup      AAAAAAAAAAA-- 370
Clone_C_Fup      AAAAAAAAAAAAA 371
                 **********
Consensus        AAAAAAAAAAAA  372
```

Figure 99. Multiple DNA sequence alignment of different clones encoding NvCI.
These sequences were aligned using ClustalX version 2.0. Rup and Fup indicate the Fup-17 and Rup-17 primers used for DNA sequencing.

```
1    tagggcgaattgggtacggattttcacgtgccggacgacaggccgtgcattaagccaggc
      -  G  E  L  G  T  D  F  H  V  P  D  D  R  P  C  I  K  P  G    13

61   cggtgccctctagttcctgatgccacgtgcacattcgtctgcaaagcagcggacaacgac
      R  C  P  L  V  P  D  A  T  C  T  F  V  C  K  A  A  D  N  D    33

121  tttggctatgaatgccagcacgtttggacctttgagggccaacgggtgggctgctatgcc
      F  G  Y  E  C  Q  H  V  W  T  F  E  G  Q  R  V  G  C  Y  A    53

181  tgagagacagacagggagagggagggagtcagctccaagctagctgccatgcctcccccct
      *

241  cttgctcctgcctccatcacagacagcctggagcccagtccacaggcgtccttcaagttg

301  tgttgggaagtctgtcctgaagaccccgtcaataaagatccctctcctgaaaaaaaaaaa

361  aaaaaaaaaaaa
```

Figure 100. cDNA consensus sequence of the gene encoding NvCI.
This figure shows the amino acid sequence (in magenta color) and consensus nucleotide sequence encoding NvCI. The triplet "tga" marked with an asterisk corresponds to the stop codon for amino acid translation and underlined sequence "aataaa" corresponds to the canonical polyadenylation signal.

The overlapping of several sequence segments, including the N-terminus as well as the amino acid sequence translated from the cDNA encoding NVCI allowed to obtain the complete primary structure of the inhibitor (figure 101) indicating that NvCI is a polypeptide of 53 amino acid residues including 6 cysteine residues involved in disulfide bridges.

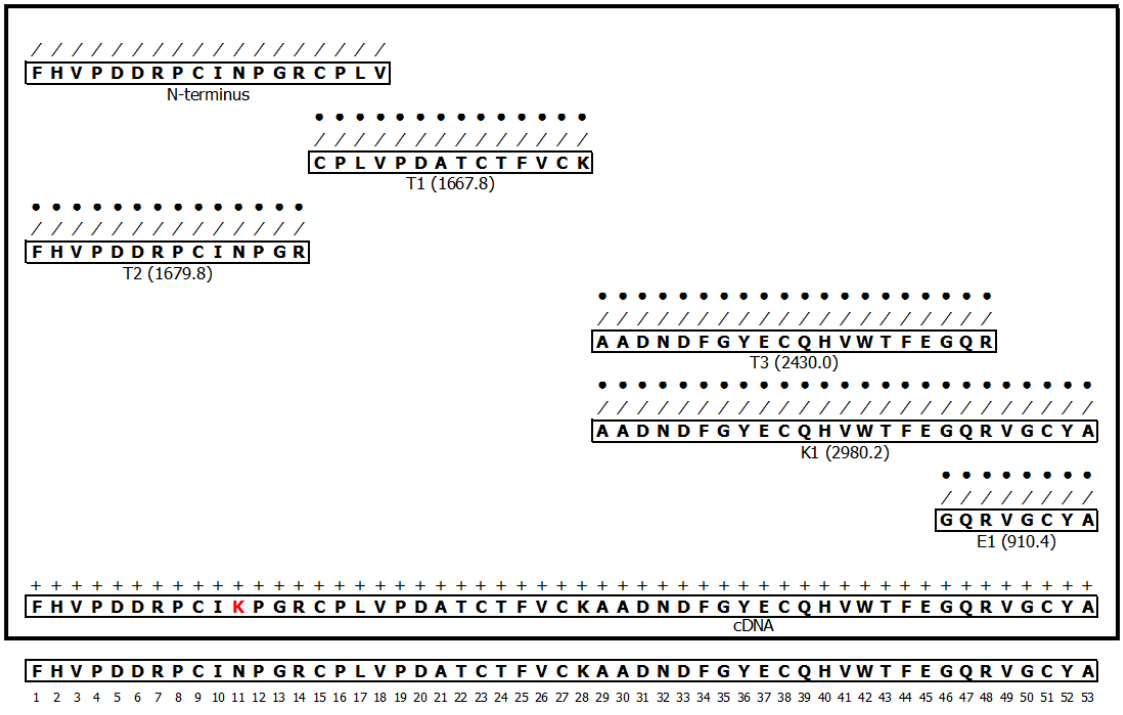
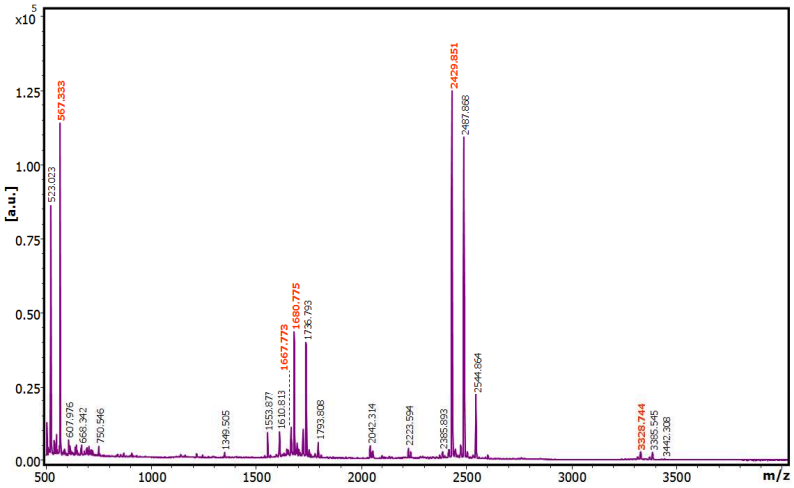


Figure 101. Amino acid sequence of NvCI.
Peptides are identified by letters indicating the enzyme used for each digestion: T indicates tryptic peptides, K endoproteinase-Lys-C peptides and E endoproteinase Glu-C peptides. Theoretical molecular mass of each peptide is shown in parentheses. N-terminus sequence was obtained by Edman degradation of the entire molecule. Bars indicate amino acid residues identified by Edman degradation. Dots indicate residues determined by MALDI-TOF MS. Crosses indicate residues obtained by translating the nucleotide cDNA sequence. Lysine residue printed in red represents the difference found in the NvCI isoform called NvCIb.

IV.4.2.3. Analysis of primary structure of NvCI isoforms by MALDI-TOF MS

Once identified the amino acid sequence of NvCI by Edman degradation, *de novo* sequencing by MALDI-TOF MS along with the identification of NvCIb isoform by synthesis and cloning of cDNA, the identification of each protein by peptide mass fingerprinting was performed. In the case of NvCIc and NvCId isoforms, as previously determined by Edman degradation, identification of an identical amino acid sequence to the first two NvCI isoforms with removal of the phenylalanine residue at the N-terminal tail was determined. In addition, taking into account that the mass difference of 14 Da remains between these NvCIc and NvCId isoforms, as observed for NvCI and NvCIb, the entire amino acid sequence of NvCI was taken as template for the establishment of the total sequence of NvCIc, which coincides with the molecular mass of NvCIc isoform determined by MALDI-TOF MS. In turn, the complete sequence of NvCIb was taken as template for the establishment of the entire amino acid sequence of NvCIc that also coincides with its molecular mass determined by MALDI-TOF MS.

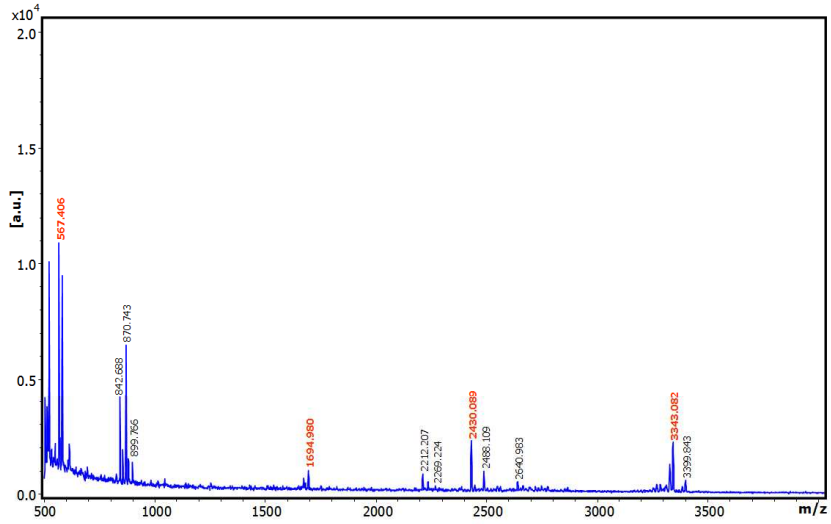
Figure 102 shows the peptide mass fingerprinting of the tryptic digestion of NvCI. In this experiment, the identification of 100% of the amino acid sequence was reached by comparing the theoretical digestion of the protein with experimentally generated peptides after trypsin digestion.



Peptide (m/z)	Aminoacid range	Sequence
567.333	49-53	FHVPDDRPCI NPGRCPLVPD ATCTFVCKAA DNDFGYECQH VWTFEGQRVG CYA
1667.773	15-28	FHVPDDRPCI NPGRCPLVPD ATCTFVCKAA DNDFGYECQH VWTFEGQRVG CYA
1680.775	1-14	FHVPDDRPCI NPGRCPLVPD ATCTFVCKAA DNDFGYECQH VWTFEGQRVG CYA
2429.851	29-48	FHVPDDRPCI NPGRCPLVPD ATCTFVCKAA DNDFGYECQH VWTFEGQRVG CYA
3328.744	1-28	FHVPDDRPCI NPGRCPLVPD ATCTFVCKAA DNDFGYECQH VWTFEGQRVG CYA

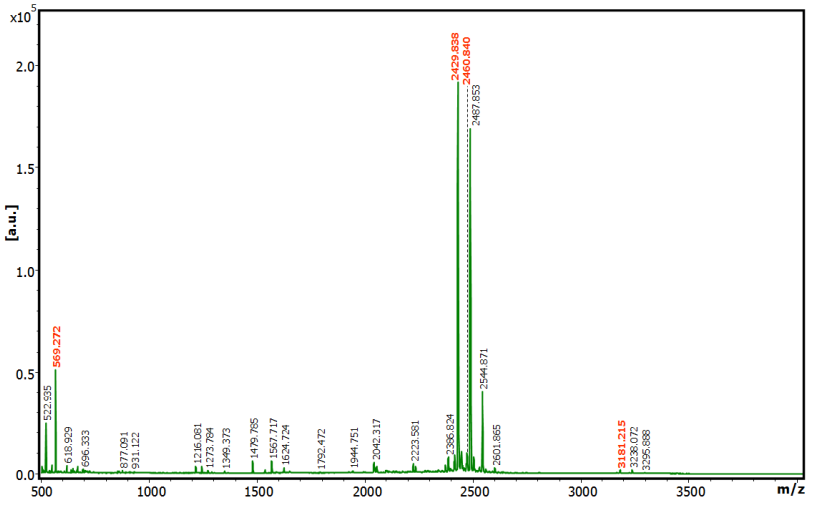
Figure 102. Peptide mass fingerprinting of the tryptic digestion of NvCI.
Analysis by MALDI-TOF MS using Biotools 3.2. 1 µl of sample was deposited on a MTP 384 target plate polished steel T F (Bruker Daltonics), followed by deposition of 1 µl of α-CHCA as a matrix. The mixture was allowed to dry at room temperature. The experimental conditions of tryptic digestion are described in section III.9.2.3.

Peptide mass fingerprinting for NvCIb isoform is presented in figure 103. In this case, identification of 4 digestion peptides allowed the coverage of 100% of NvCIb amino acid sequence. Interestingly, clear identification of lysine residue at 11th position in digestion peptides at 3342,100 and 1694,980 Da corroborates the presence of this amino acid residue previously identified in the synthesis and cloning of cDNA.



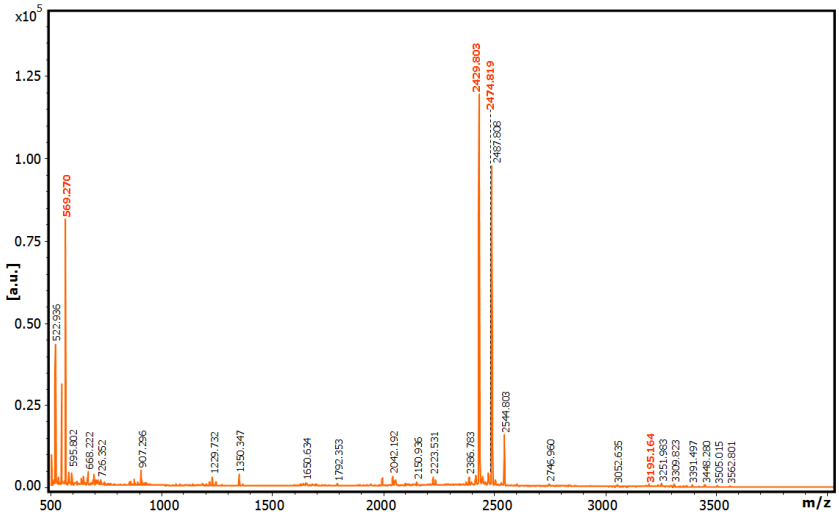
Peptide (m/z)	Aminoacid range	Sequence
567.406	49-53	10 20 30 40 50 FHVPDDRPCI KPGRCPLVPD ATCTFVCKAA DNDFGYECQH VWTFEGQRVG CYA
1694.980	1-14	10 20 30 40 50 FHVPDDRPCI KPGRCPLVPD ATCTFVCKAA DNDFGYECQH VWTFEGQRVG CYA
2430.089	29-48	10 20 30 40 50 FHVPDDRPCI KPGRCPLVPD ATCTFVCKAA DNDFGYECQH VWTFEGQRVG CYA
3342.100	1-28	10 20 30 40 50 FHVPDDRPCI KPGRCPLVPD ATCTFVCKAA DNDFGYECQH VWTFEGQRVG CYA

Figure 103. Peptide mass fingerprinting of the tryptic digestion of NvCIb.
Analysis by MALDI-TOF MS using Biotools 3.2. 1 µl of sample was deposited on a MTP 384 target plate polished steel T F (Bruker Daltonics), followed by deposition of 1 µl of α-CHCA as a matrix. The mixture was allowed to dry at room temperature. The experimental conditions of tryptic digestion are described in section III.9.2.3.



Peptide (m/z)	Aminoacid range	Sequence
569.272	48-52	10 20 30 40 50 HVPDDRPCIN PGRCPLVPDA TCTFVCKAAD NDFGYECQHV WTFEGQRVGC YA
2429.838	28-47	10 20 30 40 50 HVPDDRPCIN PGRCPLVPDA TCTFVCKAAD NDFGYECQHV WTFEGQRVGC YA
2460.840	7-27	10 20 30 40 50 HVPDDRPCIN PGRCPLVPDA TCTFVCKAAD NDFGYECQHV WTFEGQRVGC YA
3181.215	1-27	10 20 30 40 50 HVPDDRPCIN PGRCPLVPDA TCTFVCKAAD NDFGYECQHV WTFEGQRVGC YA

Figure 104. Peptide mass fingerprinting of the tryptic digestion of NvCIc.
Analysis by MALDI-TOF MS using Biotools 3.2. 1 µl of sample was deposited on a MTP 384 target plate polished steel T F (Bruker Daltonics), followed by deposition of 1 µl of α-CHCA as a matrix. The mixture was allowed to dry at room temperature. The experimental conditions of tryptic digestion are described in section III.9.2.3.



Peptide (m/z)	Aminoacid range	Sequence
569.270	48-52	10 20 30 40 50 HVPDDRPCIK PGRCPLVPDA TCTFVCKAAD NDFGYECQHV WTFEGQRVGC YA
2429.803	28-47	10 20 30 40 50 HVPDDRPCIK PGRCPLVPDA TCTFVCKAAD NDFGYECQHV WTFEGQRVGC YA
2474.819	7-27	10 20 30 40 50 HVPDDRPCIK PGRCPLVPDA TCTFVCKAAD NDFGYECQHV WTFEGQRVGC YA
3195.614	1-27	10 20 30 40 50 HVPDDRPCIK PGRCPLVPDA TCTFVCKAAD NDFGYECQHV WTFEGQRVGC YA

Figure 105. Peptide mass fingerprinting of the tryptic digestion of NvCId.
Analysis by MALDI-TOF MS using Biotools 3.2. 1 µl of sample was deposited on a MTP 384 target plate polished steel T F (Bruker Daltonics), followed by deposition of 1 µl of α-CHCA as a matrix. The mixture was allowed to dry at room temperature. The experimental conditions of tryptic digestion are described in section III.9.2.3.

Peptide mass fingerprinting obtained for NvCIc and NvCId, shown in figures 104 and 105, also allowed the identification of 100% sequence of each of these NvCI isoforms. For NvCIc isoform, identification of digestion peptide at 2460.8 Da led to detect the asparagine residue at 10th position, also presents in NvCI. In adition, the peak at 3181,2 Da containing the N-terminal peptide without phenylalanine residue at the N-terminus was also identified, as well as the other two peptides at 569,3 Da and 2429,8 Da, common in both proteins.

On the other hand, peptide mass fingerprinting of NvCId isoform allowed to identify the lysine residue at 10th position in the digestion peptide at 2474,8 Da. Besides, identification of the N-terminal peptide at 3195,6 Da without phenylalanine residue at the N-terminus was reached. Finally, the two common tryptic peptides in all NvCI isoforms at 569,3 Da and 2429,8 Da were also found.

It is important to highlight the identification of the digestion peptide matching to amino acids 7-27 in NvCIc and NvCId isoforms, which could not be identified in NvCI and NvCIb and corresponding to amino acids 8-28 the later two isoforms. This is probably due to the signal suppression effect in MALDI espectra of these isoforms. In general, peptide mass fiingerprinting results corroborate the proper establishment of the amino acid sequence of NvCIc and NvCId isoforms.

As shown in figure 106, comparison of amino acid sequence of the four isoforms of NvCI reveals differences in the N-terminus of the molecule. These differences as demonstrated later in the three-dimensional structure of the recombinant form of NvCI, do not play an important role in the inhibition of carboxypeptidases, explaining that the inhibitory activity remains unchanged between all isoforms.

In addition, as mentioned above, the change found in the asparagine by a lysine residue at 11th position for NvCI and NvCIb isoforms, which are also kept in NvCIc and NvCId isoforms, is adscribed to a point mutation, given at genomic level. However, NvCIc and NvCId isoforms were not identified in MALDI spectrum of *N. versicolor* crude extract, although these peaks were visualized after the interaction of this marine extract with the immobilized CPB, as shown in figure 63. Thus, the presence of these last NvCI isoforms could represent a hydrolysis product by an aminopeptidase present in *N. versicolor* extract, releasing the N-terminal phenylalanine. On the other hand, given the complexity of MALDI spectrum of *N. versicolor* crude extract, mass suppression effects can not be discarded. The later hypothesis suggests that these NvCIc and NvCId isoforms arise due to mutations at genomic level, given by a total or partial deletion in the codon encoding this N-terminal amino acid residue.

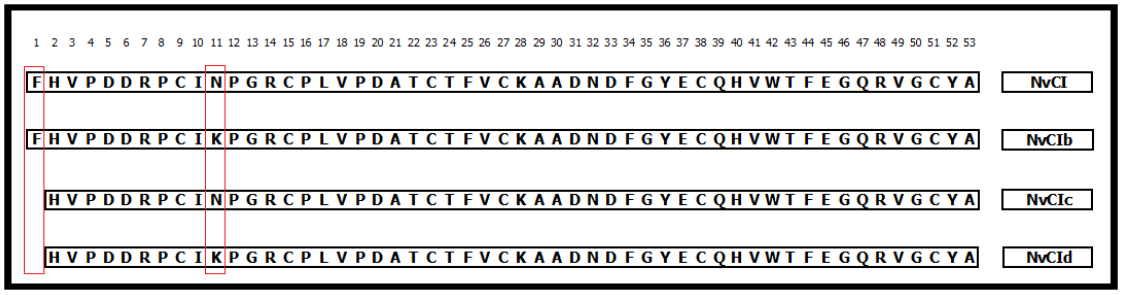


Figure 106. Comparison of amino acid sequence of NvCI isoforms.
The amino acid differences between NvCI isoforms are marked in red box.

IV.4.2.4. Computed parameters of NvCI

Known the amino acid sequence of NvCI, theoretical isoelectric point (pI), extinction coefficient, estimated half-life, instability index, aliphatic index and grand average of hydropathicity were estimated by using ExPaSy Proteomics Server (Swiss Institute of Bioinformatics):

Table 38. Computed parameters of natural NvCI

Theoretical pI: 4.91
Extinction coefficients:
Extinction coefficients expressed in units of M^{-1} cm^{-1}, at 280 nm measured in water.

Ext. coefficient 8855
Abs 0.1% (=1 g/l) 1.488, assuming all pairs of Cys residues forming disulfide bridges

Ext. coefficient 8480
Abs 0.1% (=1 g/l) 1.425, assuming all Cys residues are reduced

Estimated half-life:

The estimated half-life is: 1.1 hours (mammalian reticulocytes, *in vitro*)
3 min ((yeast, *in vivo*)
2 min (*Escherichia coli*, *in vivo*)

Instability index:

The instability index (II) is computed to be 29.77
This classifies the protein as stable.

Aliphatic index: 49.62

Grand average of hydropathicity (GRAVY): -0.279

IV.4.2.5. Comparison of the C-terminal amino acid sequence of NvCI with other proteinaceous carboxypeptidase inhibitors

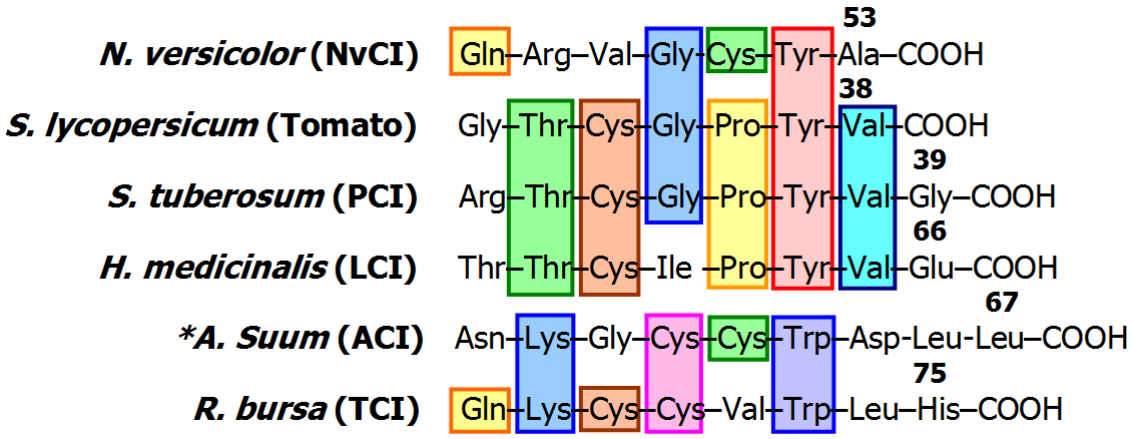


Figure 107. Comparison of the C-terminal amino acid sequence of NvCI with other proteinaceous carboxypeptidase inhibitors.
(Figure adapted from Arolas *et al.*, 2005) * Corrected amino acid sequence by Sanglas *et al.*, 2009

Comparison of the C-terminal of NvCI with other carboxypeptidase inhibitors is shown in figure 107. Some similarities were found regarding to other described inhibitors, as is the case of Tyr 52, Cys51, Gly50 and Gln47. However, all these carboxypeptidase inhibitors in comparison with NvCI (with the exception of tomato carboxypeptidase inhibitor) show an additional residue at the C-terminal tail. It has been described that this C-terminal residue in TCI, PCI and LCI, representing the P1' position in these inhibitors but absent in NvCI, is removed after binding to carboxypeptidases (Arolas *et al.*, 2005). However, this is not the case of NvCI, since this molecule was identified by IF MALDI-TOF MS from *N. versicolor* crude extract against CPA and CPB, with a molecular mass corresponding to the NvCI sequence previously determined. In addition, results obtained in the synthesis and cloning of cDNA encoding NvCI corroborated the alanine 53 as the C-terminal residue of this molecule. On the other hand, in spite of the absence of an extra C-terminal residue, the purified NvCI displayed a stronger inhibitory activity against CPs, in comparison to those described for TCI, PCI and LCI (see paragraph IV.8.1.4).

IV.5. Heterologous expression of NvCI in *P. pastoris*

Taking into account the exceptional kinetic characteristics of NvCI (K_i values in the picomolar range, see paragrahp IV.8.1.7.), which make it the most potent proteinaceous inhibitor against carboxypeptidases so far described and the importance of preserving natural sources such as marine invertebrates, heterologous expression of this inhibitor in *P. pastoris* system was carried out.

It is important to mention that in this work the kinetic characterization of natural NvCI was prior developed to its recombinant production. In order to compare the results obtained for these two molecules, the kinetic characterization of natural and recombinant NvCI is presented after the heterologous expression and purification procedure of recombinant NvCI.

Given the amino acid sequence of NvCI previously obtained by Edman degradation and MALDI-TOF MS, the design of a synthetic gene for its expression in *P. pastoris* system was performed. DNA corresponding to NvCI was fused in frame to the *S. cerevisiae* prepro-α-factor signal under the control of the *AOX1* gene promoter in *XhoI* site of pPICZαA expression vector for secretion into the culture media (figure 108).

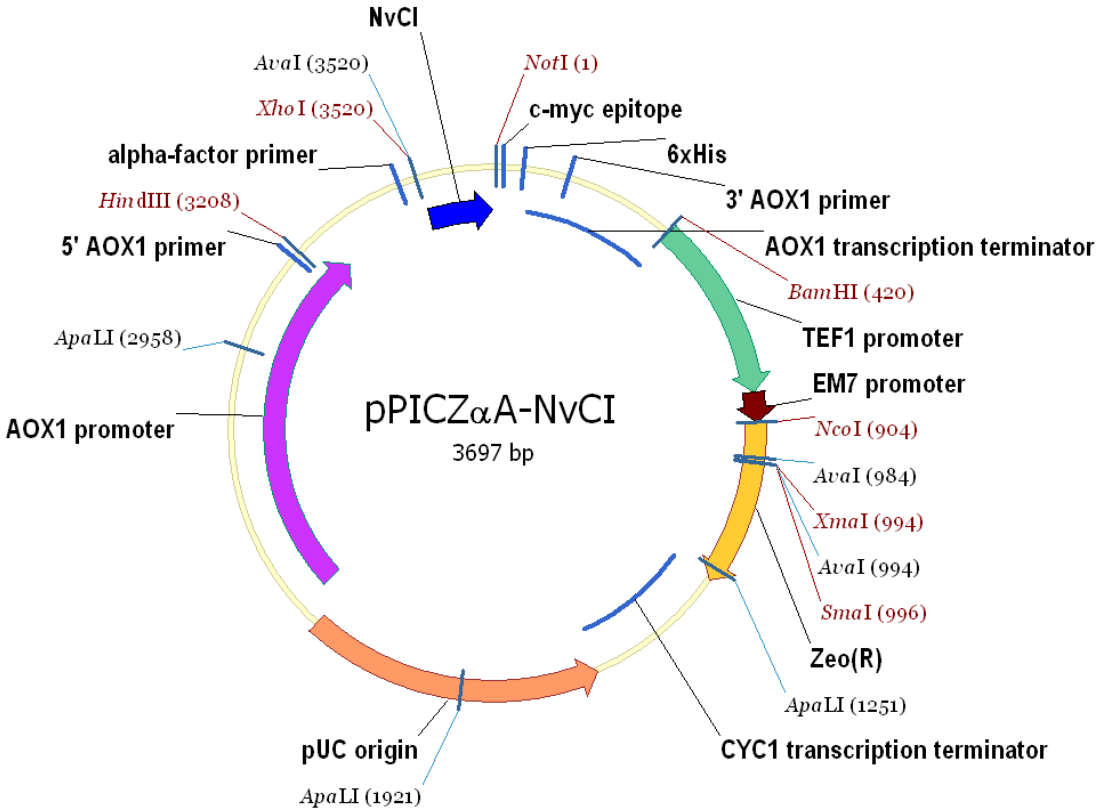


Figure 108. pPICZαA-NvCI construct for expression in *Pichia pastoris* system.
The synthetic DNA sequence was fused in frame to the *S. cerevisiae* prepro-α-actor signal under the *AOX1* gene promoter in *XhoI* site of pPICZαA vector for secretion into the culture medium.

Subsequently, the linearized plasmid of pPICZαA-NvCI was transformed into *P. pastoris* X-33 strain (Mut$^+$). Furthermore, selection of Zeocin-resistant transformants at high concentrations of Zeocin (i.e., Zeocin hyper-resistant transformants) generated enrichment in recombinant strains with multiple copies of the integrated vector. Direct selection of Zeocin hyper-resistant transformants was used to generate a population of multicopy clones that may ultimately result in an increase in the level of heterologous protein expression (Higgings *et al.*, 1998). The generation of recombinant strains with multiple copies of the expression plasmid integrated into the genome has shown to increase the heterologous protein production via a gene dosage effect for a number of different heterologous genes (Romanos, M. 1995).

Recombinant production of NvCI was carried out by using a Zeocin hyper-resistant strain in an autoclavable bioreactor (3 l) containing the ez-control module for measurements and system control (Applikon® Biotechnology, Netherlands). The production was monitored according to parameters such as wet cell weight (WCW), as well as by MALDI-TOF MS, determination of protein concentration by BCA method and bCPA inhibitory activity. Finally, process performance through the purification and quantification of recombinant NvCI was evaluated.

IV.5.1. Design of pPICZαA – NvCI construct

Plasmid DNA stocks of pPICZαA and pMA-T (which contains the DNA encoding for NvCI) were transformed in competent *E.coli* XL1-Blue cells in order to increase the amount of each plasmid. Plasmid DNA of each transformation was purified according to the method described in section III.10.1. Then, plasmid DNA of pPICZαA and pMA-T were cut by double digestion with *Not* I and *Xho* I restriction enzymes. Digestion products of pPICZαA were analyzed by 1% w/v agarose gel electrophoresis, where a band at approximately 3530 bp DNA corresponding to the cut vector was visualized (figure 109A). Moreover, analysis by 2% w/v agarose gel electrophoresis of the pMA-T digestion showed a band around 200 bp, which represents the DNA fragment encoding for NvCI (figure 109B).

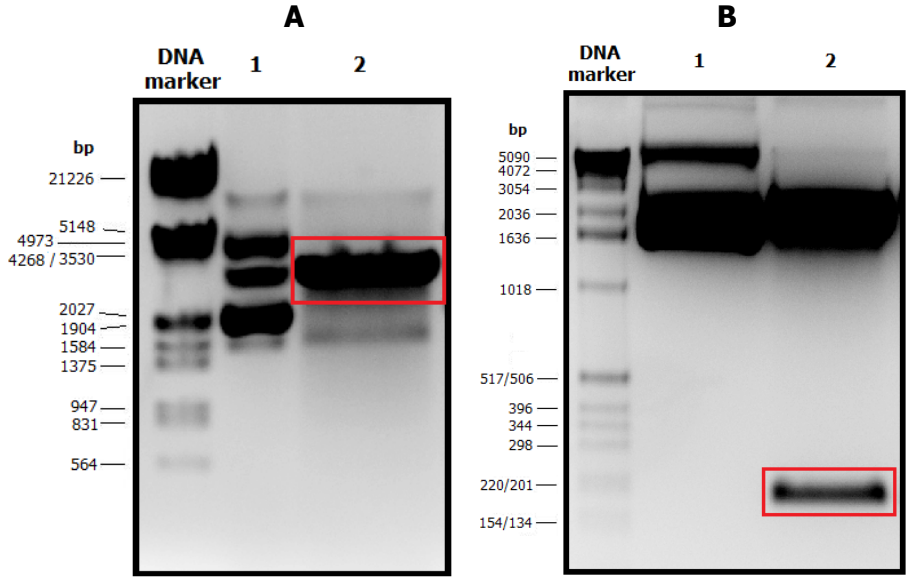


Figure 109. Double digestion with *Not* I and *Xho* I restriction enzymes of pPICZαA and pMA-T vectors. (A) pPICZαA digestion. 1% w/v agarose gel. Lane 1: mixture of plasmid and digestion buffer, lane 2: mixture of plasmid, digestion buffer and restriction enzymes. The band marked with red box corresponds to linearized pPICZαA. **(B) pMA-T digestion.** 2% w/v agarose gel. Lane 1: mixture of plasmid and digestion buffer, lane 2: mixture of plasmid, digestion buffer and restriction enzymes. The band marked with red box represents the linearized fragment corresponding to NvCI.

The bands corresponding to pPICZαA and NvCI were purified and subsequently subjected to overnight ligation. Ligation products were transformed into competent *E. coli* XL1-Blue cells and 10 colonies were selected. Plasmid DNA of pPICZαA- NvCI was purified and sequenced. Clone sequences were translated in three reading frames using a translate tool (ExPASy server, Swiss Institute of Bioinformatics). Sequencing results showed that 8 out of 10 colonies contained the insert NvCI (data not shown).

One selected clone of pPICZαA- NvCI purified above was linearized with *Sac* I restriction enzyme and digestion products were analyzed by 1% w/v agarose gel 1% electrophoresis, where a single band at approximately 3530 Da was clearly visualized. This DNA band corresponds to the linearized pPICZαA- NvCI construct (figure 110). The purified DNA from agarose gel was then transformed into *P. pastoris* X-33 strain (Mut^+) according to the protocol described by Cereghino *et al.*, 2005. After 3 days of growth on plates, 3 colonies were selected for growth in YPD medium containing 100 μg/ml Zeocin.

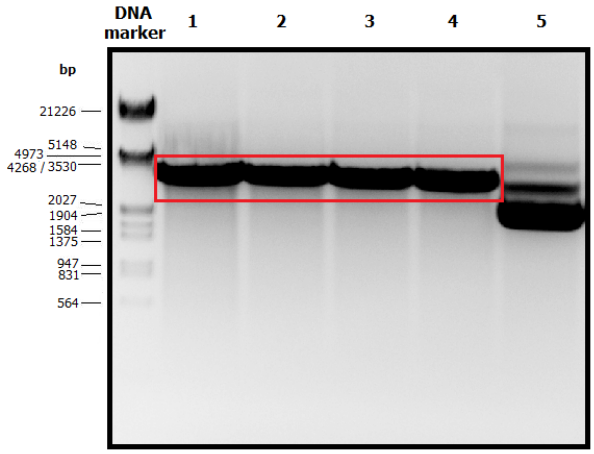


Figure 110. Digestion of pPICZαA-NvCI plasmid with *Sac* I restriction enzyme.
1% w/v agarose gel. Lanes 1 to 4: mixture of plasmid, digestion buffer and restriction enzyme. The band marked with red box corresponds to linearized pPICZαA-NvCI. Lane 5: mixture of plasmid and digestion buffer.

In the generation of Zeocin hyper-resistant strains, only one of the initially selected colonies could grow at Zeocin concentration of 5000 μg/ml. At this time, the hyper-resistant strain was stored in 15% v/v glycerol and it was also grown in Petri dish with YPD medium containing 500 μg/ml Zeocin. Finally, one isolated colony was selected for production in bioreactor of NvCI.

IV.5.2. Production monitoring and yield of NvCI in *P. pastoris*

Recombinant production of NvCI at pilot-plant-scale was performed in a 3 L bioreactor. During the growth phase with glycerol as carbon source, bioreactor temperature was maintained at 28ºC. After approximately 16 hours of growth with glycerol, an increase of the dissolved oxygen (DO) concentration to 100% was observed, indicative of total consumption of this carbon source. At this point, the addition of methanol was initiated under the conditions described in section III.10.4.3 at 25ºC.

At methanol induction time = 0 hours, a growth of cell mass on wet basis of 128.6 g/l was measured. Further analysis by RP-HPLC (figure 111) and MALDI-TOF MS did not reveal the presence of NvCI, nor CPA inhibitory activity was detected in the fermentation supernatant within the range of protein concentrations tested in the assay (0 -90 ng/ml).

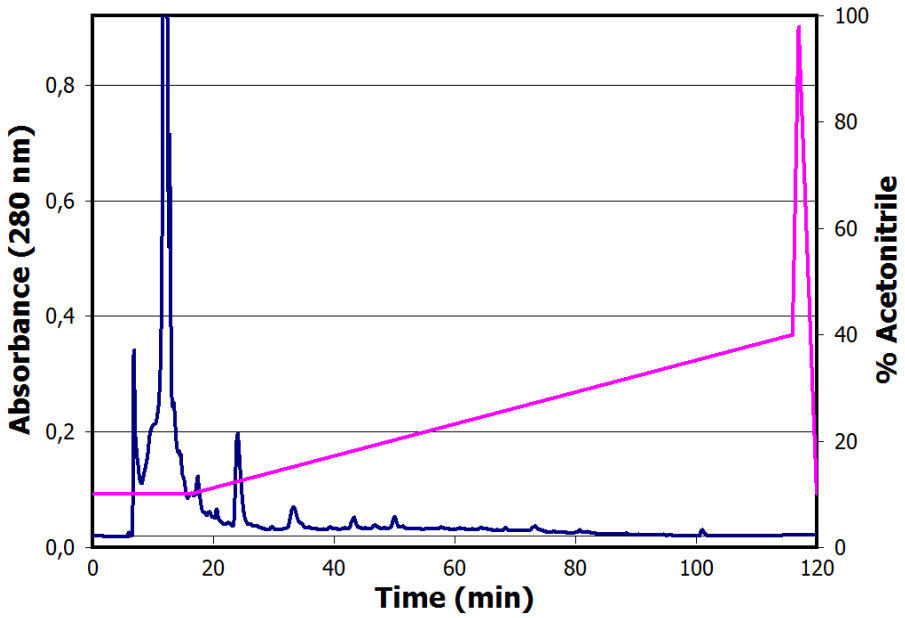


Figure Nº 111. RP-HPLC of the NvCI fermentation in *P. pastoris* at 0 h of induction time.
C4 (3.9 mm x 150 mm) column. Solution A was 0.1% v/v TFA in water, while solution B was 0.1% v/v TFA in acetonitrile. Elution was performed under the following experimental conditions: At the start, 10% of solution B during 15 min was used, followed by a linear gradient from 10 to 40% B over 100 min, a linear gradient from 40 to 98% B over 1 min and a linear gradient from 98 to 10% B over 3 min. The flow rate was 0.5 ml/min and room temperature. Absorbance at 280 nm, acetonitrile gradient (% volume).

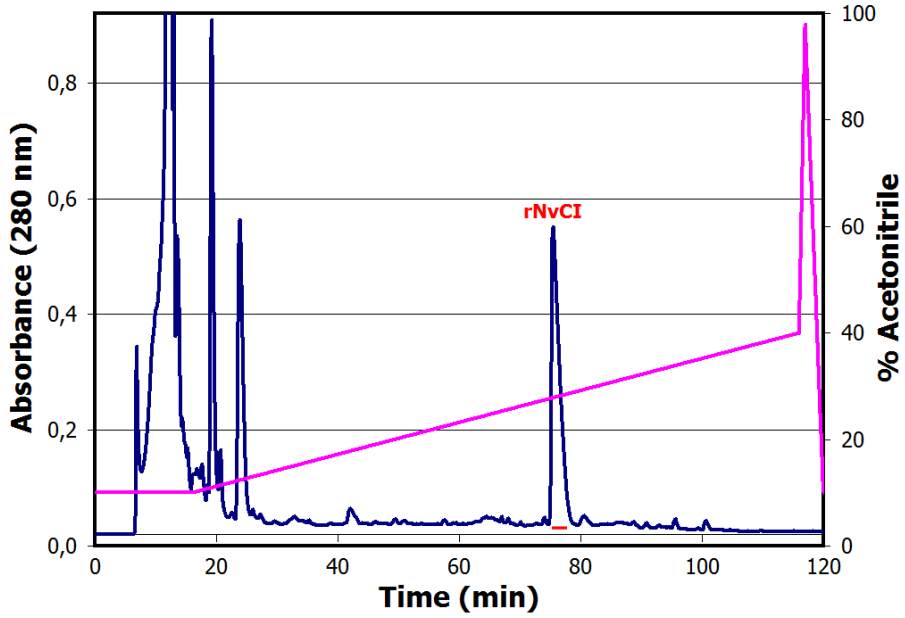


Figure Nº 112. RP-HPLC of the NvCI fermentation in *P. pastoris* at 24 h of induction time.
C4 (3.9 mm x 150 mm) column. Solution A was 0.1% v/v TFA in water, while solution B was 0.1% v/v TFA in acetonitrile. Elution was performed under the following experimental conditions: At the start, 10% of solution B during 15 min was used, followed by a linear gradient from 10 to 40% B over 100 min, a linear gradient from 40 to 98% B over 1 min and a linear gradient from 98 to 10% B over 3 min. The flow rate was 0.5 ml/min and room temperature. Absorbance at 280 nm, rNvCI peak, acetonitrile gradient (% volume).

After 24 hours of methanol induction, an increase in protein concentration in the fermentation supernatant was observed, as a result of extracellular expression of proteins from *P. pastoris*. The presence of NvCI in this supernatant was confirmed by purifying the inhibitor by RP-HPLC and determining its molecular mass in MALDI-TOF MS (figures 112 and 113). Furthermore, CPA inhibitory activity assays were carried out, yielding positive results.

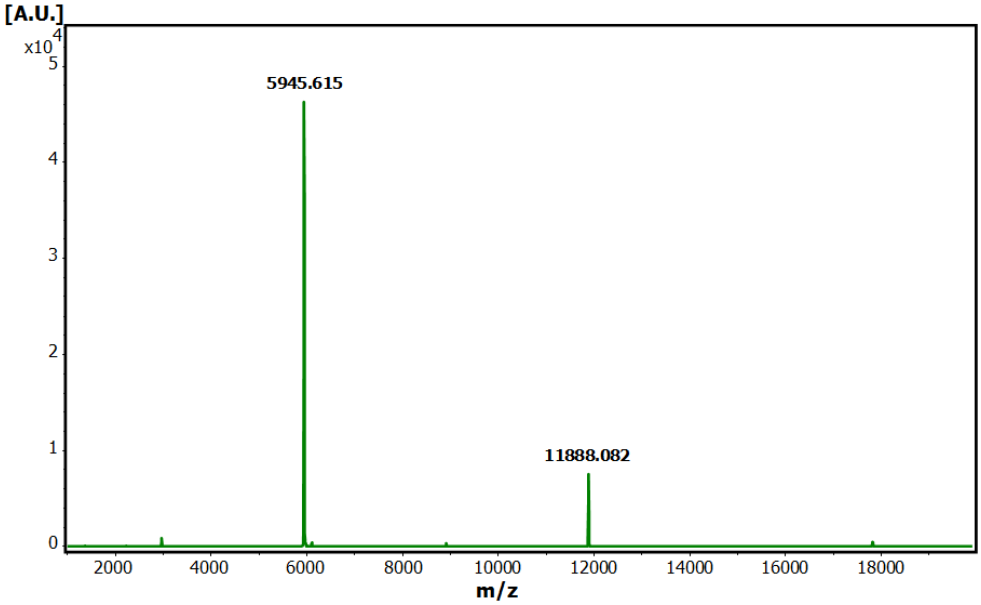


Figure Nº 113. MALDI-TOF spectrum of rNvCI obtained at 24 h of induction time and purified by RP-HPLC. 1 µl of sample was deposited on a MTP 384 target plate polished steel T F (Bruker Daltonics), followed by deposition of 1 µl of DHAP as a matrix. The mixture was allowed to dry at room temperature.

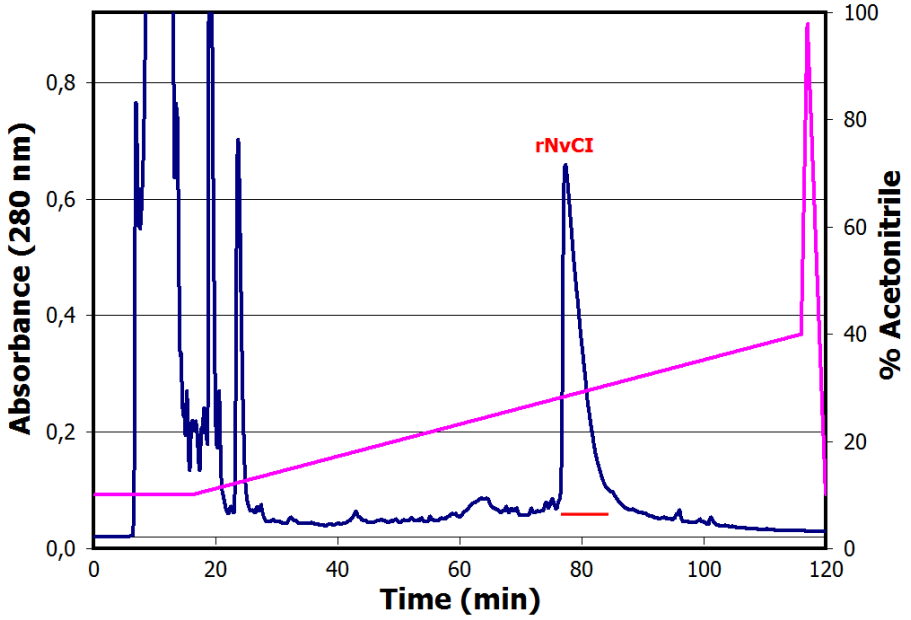


Figure Nº 114. RP-HPLC of the NvCI fermentation in *P. pastoris* at 48 h of induction time. C4 (3.9 mm x 150 mm) column. Solution A was 0.1% v/v TFA in water, while solution B was 0.1% v/v TFA in acetonitrile. Elution was performed under the following experimental conditions: At the start, 10% of solution B during 15 min was used, followed by a linear gradient from 10 to 40% B over 100 min, a linear gradient from 40 to 98% B over 1 min and a linear gradient from 98 to 10% B over 3 min. The flow rate was 0.5 ml/min and room temperature. Absorbance at 280 nm, rNvCI peak, acetonitrile gradient (% volume).

Finally after 48 hours of methanol induction, the protein concentration in the supernatant continued to enhance, resulting in an increase of 5.7 times compared to the initial time of methanol induction. The chromatographic profile of RP-HPLC showed the presence of an intense peak at a retention time of 80 min (figure 114). MALDI-TOF spectrum (figure 115) of the peak obtained from RP-HPLC confirmed the molecular mass of NvCI. Specific inhibitory activity of 75.0 U/mg in this supernatant was calculated. rNvCI concentration (mg inhibitor/ml culture medium) was determined by measuring the absorbance at 280 nm of the peak obtained from RP-HPLC using the computerized molar extinction coefficient of 8855 $M^{-1}cm^{-1}$. Results of monitoring and evaluation of fermentation performance are presented in table 39.

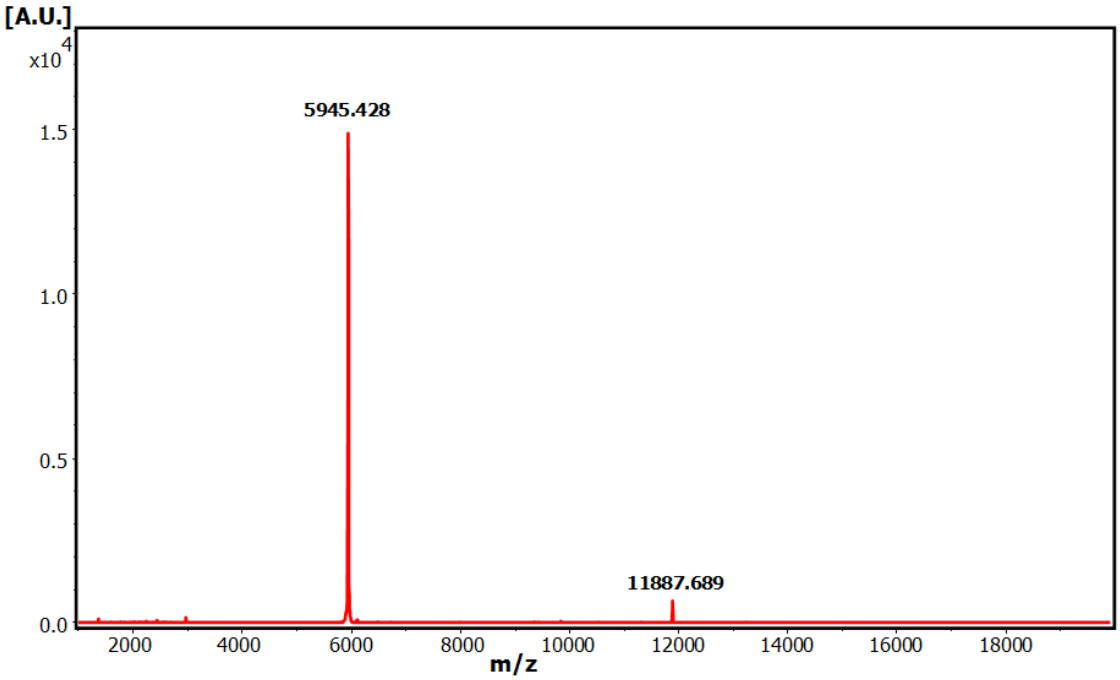


Figure Nº 115. MALDI-TOF spectrum of rNvCI obtained at 48 h of induction time and purified by RP-HPLC. 1 µl of sample was deposited on a MTP 384 target plate polished steel T F (Bruker Daltonics), followed by deposition of 1 µl of DHAP as a matrix. The mixture was allowed to dry at room temperature.

Table Nº 39. Fermentation performance of NvCI in *P. pastoris*

Induction time (h)	WCW* (g/l)	Total protein (mg/ml)	[rNvCI] (mg/l)	Specific inhibitory activity (U/mg)	IC_{50} (µg/ml)
0	128.6 ± 0.2	1.42 ± 0.04	0	0	----
24	222.9 ± 0.4	3.69 ± 0.07	190.0 ± 3.6	46.7 ± 5.0	0.037 ± 0.002
48	380.8 ± 1.2	8.15 ± 0.31	556.7 ± 20.9	75.0 ± 9.3	0.01935 ± 0.00003

*WCW: wet cell weight. Data are means (n=3) ± S.D. Inhibitory activity was measured against bovine CPA.

IV.6. Purification and characterization of recombinant NvCI

Purification of recombinant NvCI (rNvCI) was performed using a combination of two ion exchange chromatographic methods: an initial weak cation exchange (Accell™ Plus CM) and a second step of anion exchange (TSK® DEAE). Purified inhibitor was initially characterized by determining its molecular mass by MALDI-TOF MS and its functional behavior in IF MALDI-TOF MS (molecular interaction of immobilized CPA) along with inhibitory activity. After that, the N-terminal sequence of rNvCI was determined by automatic Edman degradation, as well as the number of cysteine residues and peptide mass fingerprinting. Strategy and results of rNvCI purification and characterization are described below:

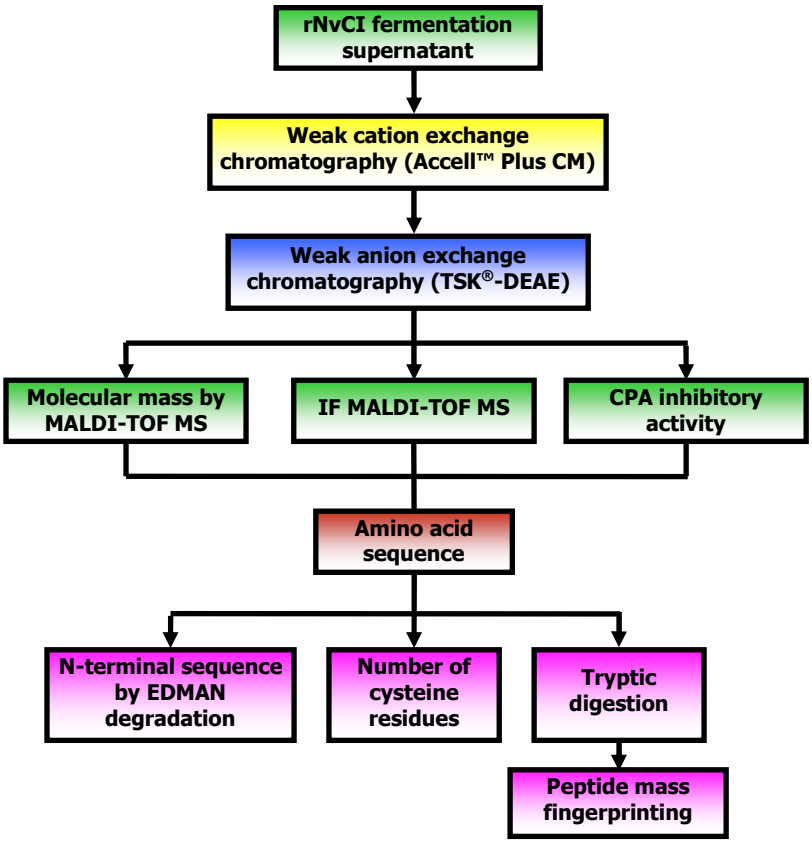


Figure 116. Flowchart process for purification and preliminary characterization of recombinant NvCI

IV.6.1. Purification of rNvCI

Fermentation supernatant was applied to a weak cation exchange chromatography column Accell™ Plus CM, equilibrated at pH 3. After removal of non-retained components and use of ionic strength gradient (up to 1 M NaCl in a buffer at pH 7.0), a protein peak containing CPA inhibitory activity was obtained (figure 117). Total inhibitory activity of the elution peak corresponded to 3778.6 Units, representing a yield of 20.6% regarding the initial stage and a specific inhibitory activity of 274.8 U/mg protein, equivalent to a 3.7-fold purification also related to the initial extract (table 40).

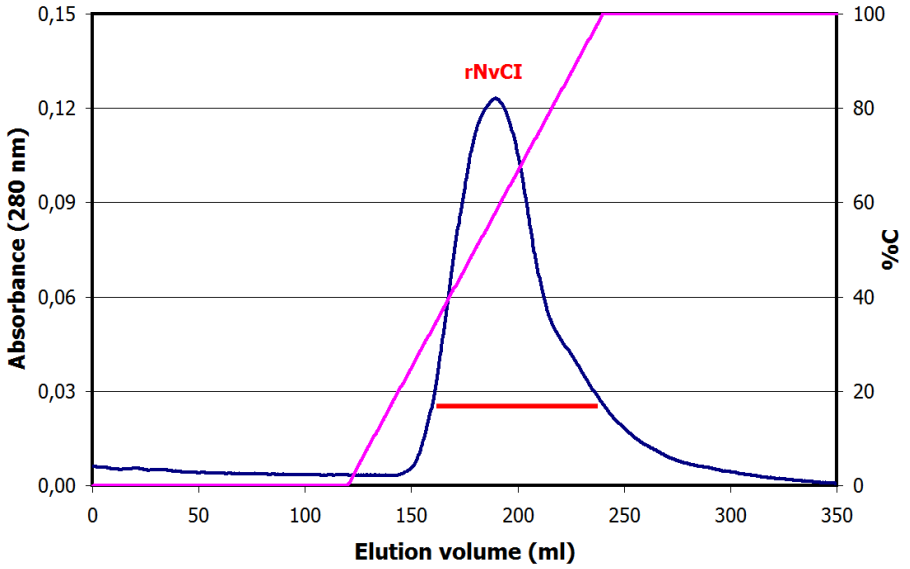


Figure 117. Weak cation exchange chromatography of rNvCI fermentation supernatant on a Waters Accell™ Plus CM column.
Column: 1.6 cm x 20.0 cm. Buffer A: 20 mM sodium citrate, pH 3.0; buffer B: 20 mM Tris-HCl pH 7.0, and buffer C: 20 mM Tris-HCl pH 7.0, containing 1 M NaCl. Column was equilibrated at 1 ml/min with buffer A over 3 column volumes (CV). Sample loading was performed at 1 ml/min in buffer A. Non-retained molecules were removed by washing the column at 1 ml/min with buffer A and 3 CV. Elution was performed under the following experimental conditions: At the start, 100% buffer B, 2 ml/min and 4 CV was used, followed by a linear gradient from 0 to 100% C at 2 ml/min with 4 CV, and 100% C at 2 ml/min and 1 CV. The whole process was carried out at room temperature.
Absorbance at 280 nm, rNvCI peak of CPA inhibitory activity, % buffer C (% volume).

The elution peak of cation exchange chromatography was analyzed by Tris-tricine SDS PAGE (figure 118), where a clear band around 6 kDa corresponding to the molecular mass of NvCI was visualized. This band was more intense at the center of the peak, although other protein contaminants were also displayed in this region. MALDI spectrum of the elution peak confirmed the presence of a major peak at 5945 Da (figure 119). The absence of other peaks in the spectrum may be due to signal suppression effects in MALDI.

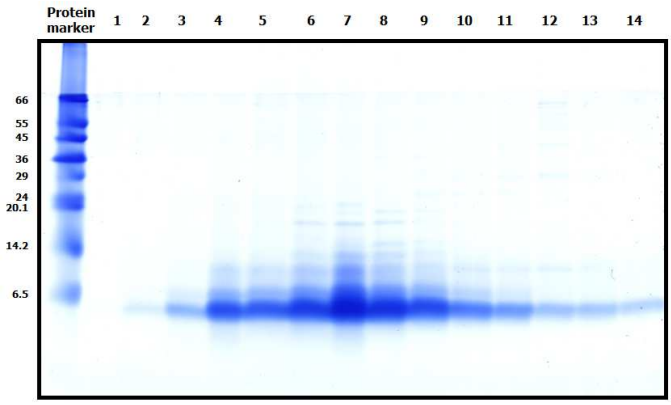


Figure 118. Tris-tricine SDS-PAGE of rNvCI elution peak obtained from CM weak cation exchange chromatography. Lane 1: elution fraction Nº 1 obtained at elution volume (V_e)=160 ml. Subsequent fractions were collected each 5 ml V_e. Lanes 2 to 14: elution fractions Nº 2 to Nº 14, respectively

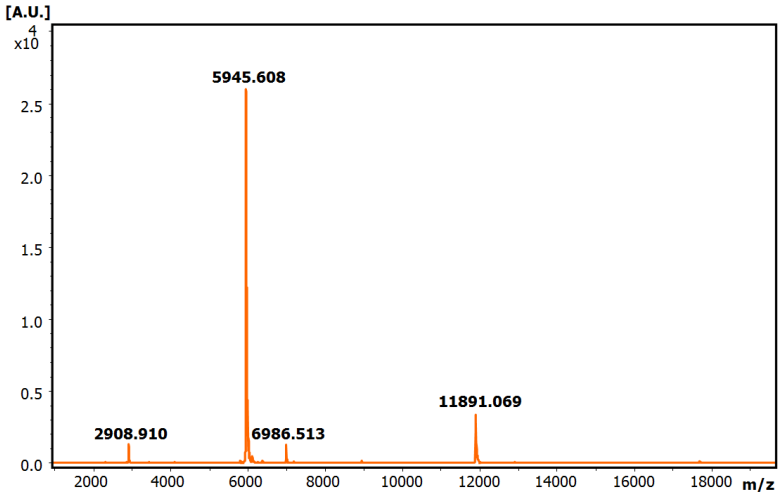


Figure Nº 119. MALDI-TOF spectrum of rNvCI elution peak obtained from CM weak cation exchange chromatography. 1 µl of sample was deposited on a MTP 384 target plate polished steel T F (Bruker Daltonics), followed by deposition of 1 µl of DHAP as a matrix. The mixture was allowed to dry at room temperature.

Active fractions representing the elution peak from weak cation exchange chromatography were dialyzed using a CelluSep H1 tubular dialysis membrane (molecular weight cut-Off of 1 kDa; Membrane filtration products Inc., USA) in order to remove salts from this separation step. Sample was equilibrated in 20 mM Tris-HCl pH 8.5 and applied to a weak anion exchange column TSK® DEAE equilibrated at pH 8.5, as a second purification step.

After removing non-retained contaminants, an elution step by a salt gradient was carried out according to the experimental conditions described in section III.11.2. A single and well-resolved peak containing CPA inhibitory activity was displayed in the chromatographic profile (figure 120). Collected fractions with CPA inhibitory activity (2905.0 U) contained a specific inhibitory activity of 293.7 U/mg, representing a 3.9-fold purification compared to the initial extract and a yield of 15.8% (table 40). A yield of 76.9% was achieved when compared to the previous purification step.

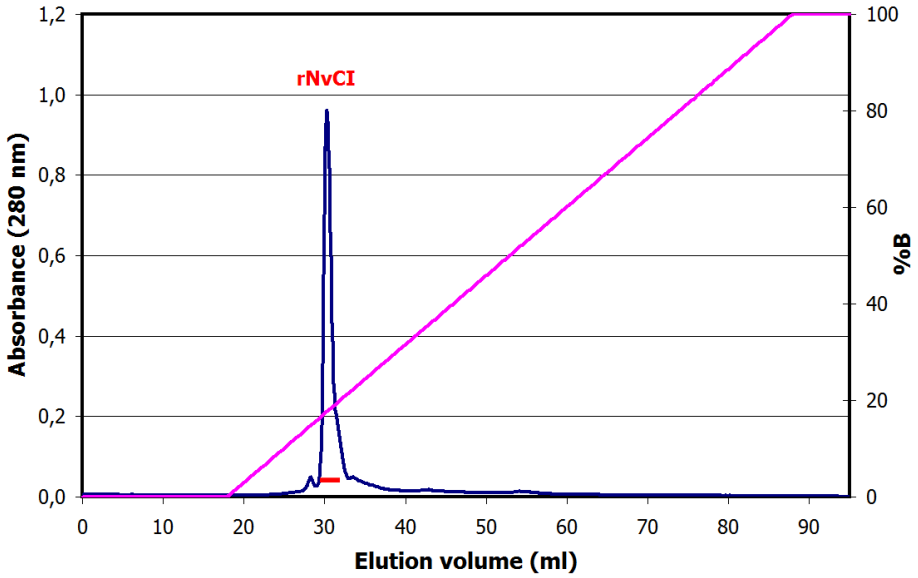


Figure 120. Weak anion exchange chromatography of CM CEC* elution peak on a TSK-GEL® DEAE-5PW column.

Column: 7.5 cm x 7.5 mm. Buffer A: 20 mM Tris-HCl pH 8.5; buffer B: 20 mM Tris-HCl pH 8.5, containing 1 M NaCl. Column was equilibrated with buffer A over 5 column volumes (CV). Sample loading was performed in buffer A. Non-retained molecules were removed by washing the column with buffer A over 5 CV. Elution was performed using a linear gradient from 0% to 100% B over 20 CV followed by 100% B over 10 CV. The flow rate was 1 ml/min and room temperature. ___ Absorbance at 280 nm, ___ rNvCI peak of CPA inhibitory activity, ___ % buffer B
*CEC: Cation exchange chromatography

Tris-tricine SDS-PAGE of the elution peak from TSK-GEL® DEAE-5PW column displayed a very intense band in the central region of the peak at approximately 6 kDa (figure 121). However, another intense band with double molecular weight was also visualized. It is likely that this second band represents a dimer of NvCI. This aggregation phenomenon has been described for other protease inhibitors at high concentrations, such as ShPI (García-Fernández *et al.*, personal communication). Purity of recombinant NvCI was also analyzed by MALDI-TOF MS, whose spectrum showed a homogeneous peak at m/z 5945.6 (figure 122).

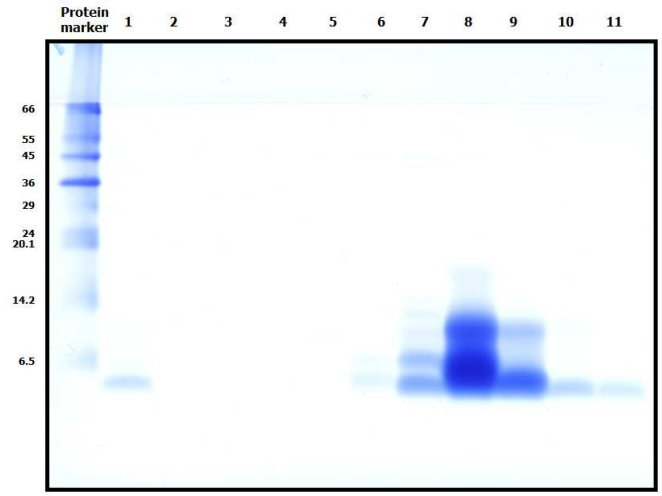


Figure 121. Tris-tricine SDS-PAGE of the elution peak obtained from TSK®-DEAE weak anion exchange chromatography. Lane 1: loading sample, lane 2: non-binding fraction, lane 3: washing fraction, lane 4: elution fraction Nº 1 obtained at elution volume (V_e)=24 ml. Subsequent fractions were collected each 2 ml V_e. Lanes 5 to 11: elution fractions Nº 2 to Nº 8, respectively

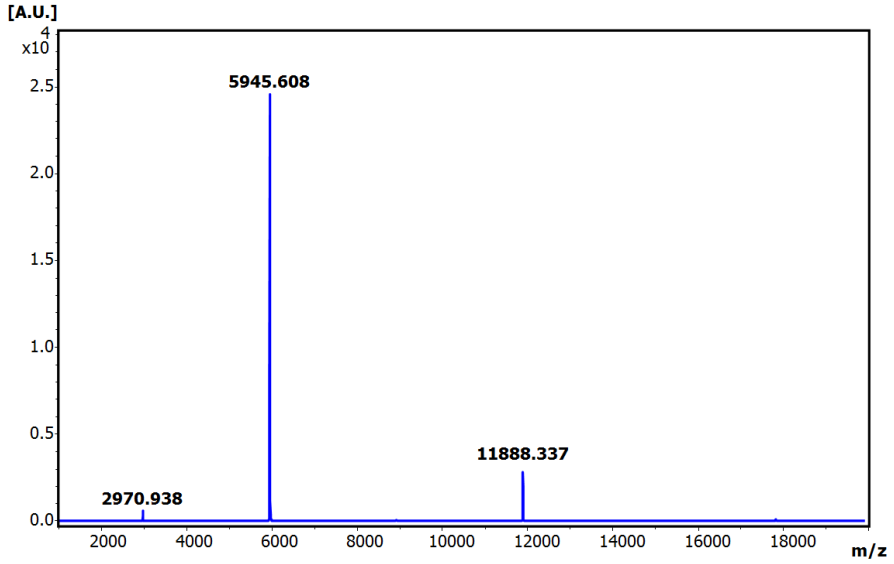


Figure Nº 122. MALDI-TOF spectrum of rNvCI elution peak obtained from TSK®-DEAE weak anion exchange chromatography. 1 µl of sample was deposited on a MTP 384 target plate polished steel T F (Bruker Daltonics), followed by deposition of 1 µl of DHAP as a matrix. The mixture was allowed to dry at room temperature.

Quantitative analysis of the purification process showed an apparently very low yield, due to the first chromatographic step (table 40). Taking into account that performance is expressed in terms of inhibitory activity, it is possible that this value in the fermentation supernatant could be affected by enzyme degradation/inactivation of bovine CPA used as a control in the activity assay. This is supported by the presence in the supernatant of proteases and other compounds that could inactivate the control enzyme. However, it should be noted that approximately 10 mg of active inhibitor were obtained, representing 329.7 mg rNvCI/l culture medium.

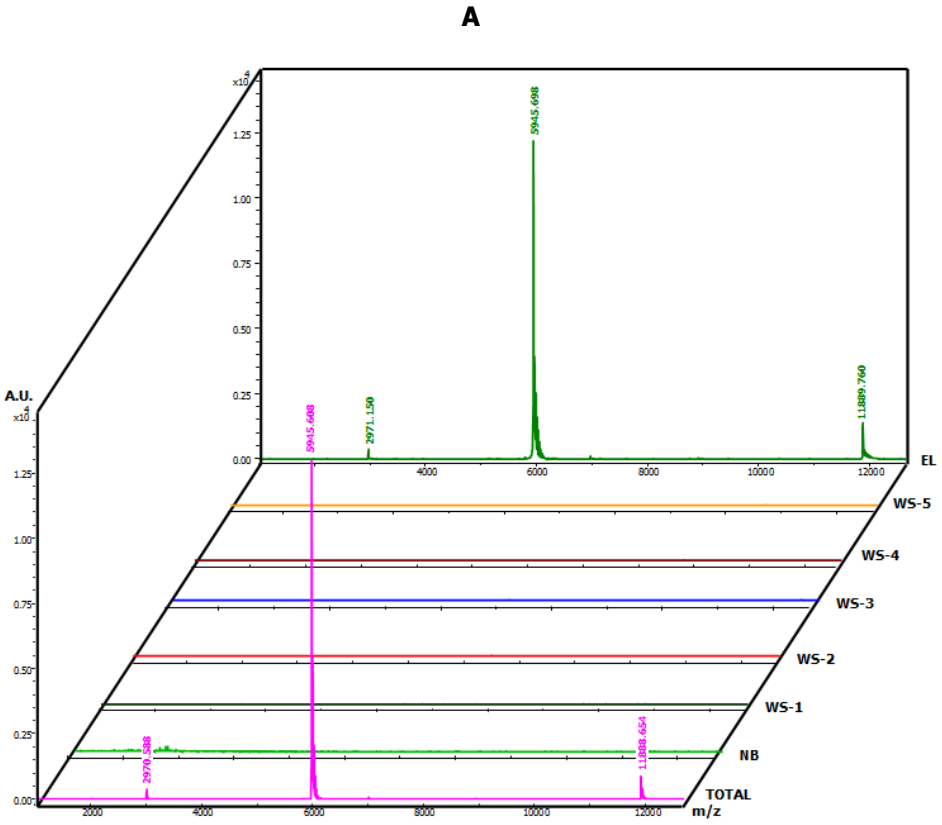
Table 40. Summary of the purification procedure of recombinant NvCI

Step	Total protein (mg)	Inhibitory activity (U)	Specific inhibitory activity (U/mg)	Yield (%)	Purification (-fold)
Fermentation supernatant	244.5 ± 9.2	18340 ± 2262	75.0 ± 9.3	100.0 ± 0.0	1.0 ± 0.0
CM weak cation exchange	13.7 ± 0.6	3779 ± 464	274.8 ± 33.7	20.6 ± 3.6	3.7 ± 0.6
TSK weak anion exchange	9.89 ± 0.02	2905 ± 383	293.7 ± 38.8	15.8 ± 2.9	3.9 ± 0.7

Data are means (n=3) ± S.D.

IV.6.2. Molecular characterization of rNvCI

The purity of recombinant NvCI and its ability to interact selectively with bovine CPA was demonstrated by IF MALDI-TOF MS. MALDI spectrum of the inhibitor obtained from the anion-exchange chromatography showing a single peak with the same molecular mass of natural NvCI (figure 122). This peak was completely captured by the immobilized CPA since any MALDI signal in the non-retained fraction and washings was detected. Subsequently, this initial peak was recovered after elution with 0.5% v/v TFA. This result demonstrates that the obtained recombinant inhibitor is pure and able to selectively interact with bovine CPA (figure 123).



B

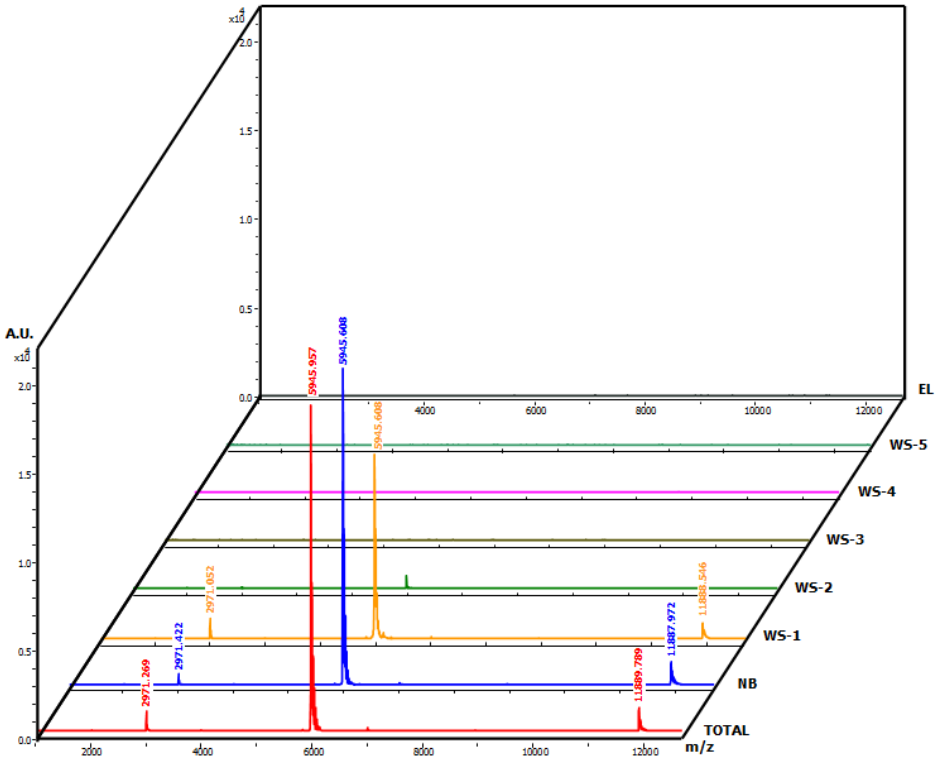


Figure 123. Intensity Fading MALDI-TOF MS of purified rNvCI on CPA-glyoxal Sepharose® CL-4B (A) and Sepharose® CL-4B (B). TOTAL: rNvCI in 20 mM Tris-HCl buffer pH 7.5, 500 mM NaCl. **NB:** non-binding molecules, interaction at pH 7.5, 10 min. **WS-1 to WS-5:** washing at pH 7.5. **EL:** elution with 0.5% v/v TFA, 10 min. Interaction assays were carried out in duplicate at room temperature. The experimental conditions of MALDI-TOF MS are described in section III.7.2**.**

On the other hand, the N-terminal sequence of rNvCI and the number of cysteine residues coincided with those previously established for the natural inhibitor (table 41 and figure 124). In addition, tryptic peptide map of the recombinant molecule was identical to that obtained for the natural inhibitor, where 100% NvCI sequence was covered (figure 125).

Table 41. N-terminus sequence by EDMAN degradation of rNvCI

F	H	V	P	D	D	R	P	C	I	N	P	G	R	C	P	L	V	P	D
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20

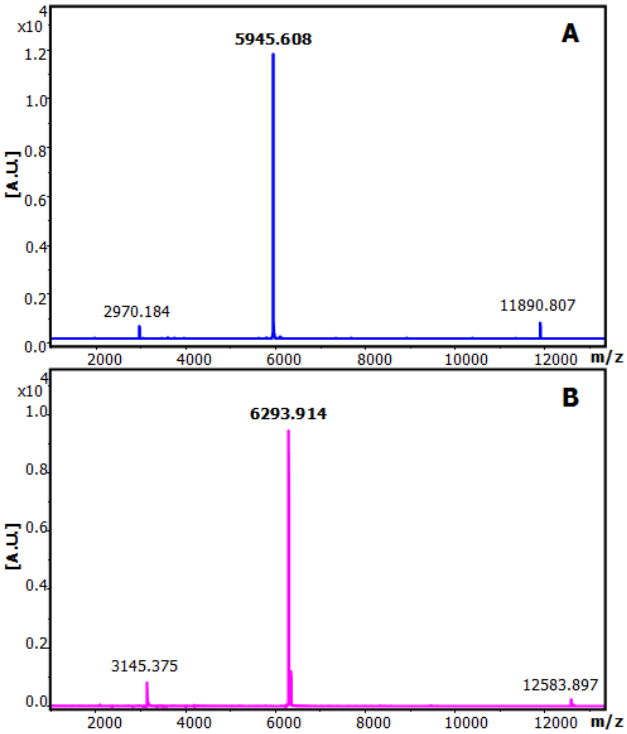
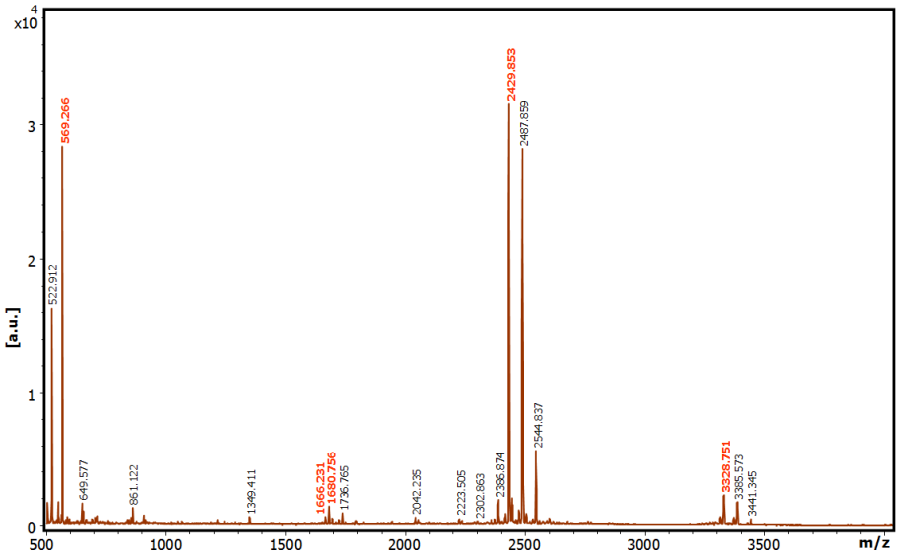


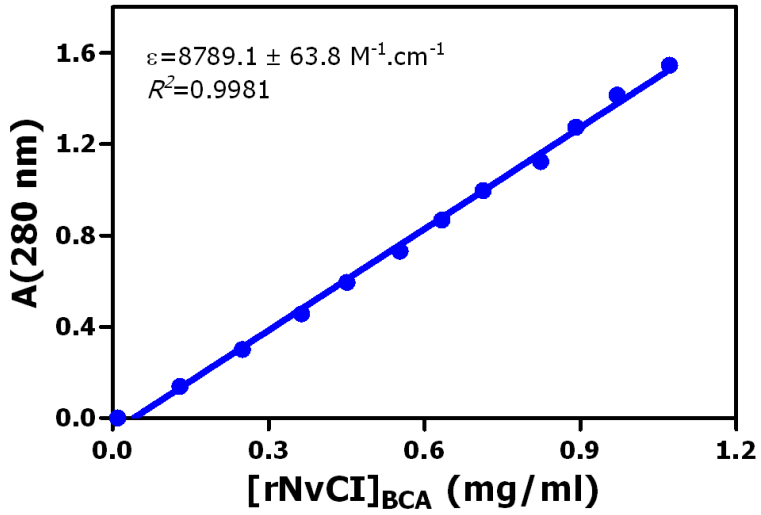
Figure 124. Reduction and S-carbamidomethylation of recombinant NvCI.
A. MALDI-TOF MS spectrum of rNvCI, using DHAP as a matrix. **B.** MALDI-TOF MS spectrum after reduction and S-carbamidomethylation of rNvCI, using DHAP as a matrix



Peptide (m/z)	Aminoacid range	Sequence
569.266	49-53	FHVPDDRPCI NPGRCPLVPD ATCTFVCKAA DNDFGYECQH VWTFEGQRVG CYA
1666.231	15-28	FHVPDDRPCI NPGRCPLVPD ATCTFVCKAA DNDFGYECQH VWTFEGQRVG CYA
1680.756	1-14	FHVPDDRPCI NPGRCPLVPD ATCTFVCKAA DNDFGYECQH VWTFEGQRVG CYA
2429.853	29-48	FHVPDDRPCI NPGRCPLVPD ATCTFVCKAA DNDFGYECQH VWTFEGQRVG CYA
3328.751	1-28	FHVPDDRPCI NPGRCPLVPD ATCTFVCKAA DNDFGYECQH VWTFEGQRVG CYA

Figure 125. Peptide mass fingerprinting of the tryptic digestion of rNvCI.
Analysis by MALDI-TOF MS using Biotools 3.2. 1 µl of sample was deposited on a MTP 384 target plate polished steel T F (Bruker Daltonics), followed by deposition of 1 µl of α-CHCA as a matrix. The mixture was allowed to dry at room temperature. The experimental conditions of tryptic digestion are described in section III.9.2.3.

IV.6.3. Determination of the molar extinction coefficient of rNvCI



Figue Nº 126. Determination of the molar extinction coefficient of rNvCI.
Date are means (n=3) ± S.D.

Molar extinction coefficient is a parameter of practical importance for determining the concentration of a purified protein by an accurate and faster method. The molar extinction coefficient measured as the slope of the straight line (absorbance vs. protein concentration determined by BCA method) was 8789 $M^{-1}.cm^{-1}$ (equivalent to 1.48 $ml.mg^{-1}.cm^{-1}$).

This value was very similar to that determined by computerized methods from the amino acid sequence (8855 $M^{-1}.cm^{-1}$). It is important to note that the molar extinction coefficient of NvCI is relatively high, due to the presence in this small molecule (53 amino acids residues) of 7 aromatic residues (1 Tryp, 2 Tyr and 4 Phe residues).

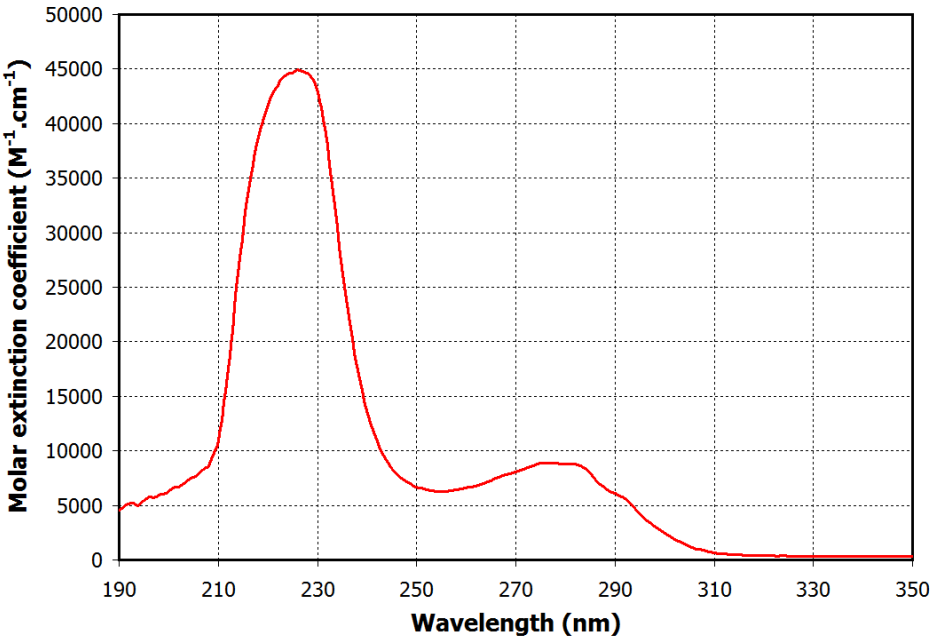


Figure Nº 127. rNvCI extinction profile. Spectrum was measured in 20 mM Tris-HCl pH 8.5, 150 mM NaCl at 37ºC, protein concentration of 0.43 mg/ml, with a Cary 400 Bio spectrophotometer in a final volume of 1 mL and 1 cm path length.

The molar extinction coefficient profile of rNvCI at different wavelengths is shown in figure 127. Table 42 summarizes the values of molar extinction coefficient at the wavelengths most commonly used.

Table 42. Molar extinction coefficient of rNvCI at different wavelengths

λ (nm)	ε ($M^{-1}.cm^{-1}$)	κ ($ml.mg^{-1}.cm^{-1}$)
320	393.9 ± 28.9	0.066 ± 0.005
310	657.9 ± 29.1	0.111 ± 0.005
300	2435.7 ± 14.0	0.410 ± 0.002
295	4267.6 ± 11.9	0.718 ± 0.002
290	6085.9 ± 17.2	1.024 ± 0.003
285	7957.1 ± 13.4	1.339 ± 0.002
280	8783.1 ± 16.9	1.477 ± 0.003
278	8844.8 ± 11.5	1.488 ± 0.002
275	8846.2 ± 18.2	1.488 ± 0.003
254	6291.2 ± 24.1	1.058 ± 0.004
225	44642.9 ± 53.1	7.510 ± 0.009
220	41692.3 ± 61.6	7.013 ± 0.010
215	30213.3 ± 107.4	5.082 ± 0.018
214	26328.1 ± 55.7	4.429 ± 0.009

Date are means (n=3) ± S.D.

IV.7. Analysis of natural and recombinant NvCI by circular dichroism spectroscopy

In order to analyze some conformational characteristics of NvCI, Circular Dichroism (CD) spectroscopy was employed. Figure 128A displays far-ultra violet (far-UV) CD spectra obtained with NvCI at different pH values. At pH 8.0, a strong minimum of ellipticity at 205 nm was observed, which could be related to the presence of a high percentage of residues in β-structure (Chen *at al.*, 1972). However, NvCI did not show a maximum at 228 nm, as in the case of PCI and LCI, where this band was attributed to an asymmetric environment of the tyrosine residue located at the C-terminal of these molecules (Venhudova *et al.*, 2001; Reverter *et al.*, 1998). It is important to mention that CD spectra of disulfide rich molecules like NvCI, LCI, among others are difficult to interpret due to the occurrence of anomalous dichroic bands such as that at 228 nm (Chen *et al.*, 1972; Reverter *et al.*, 1998). At lower pH values, a decrease in the intensity of the minimum of ellipticity at 205 nm was observed, as well as a slight rightward shift of this minimum, which may indicate conformational changes in the protein.

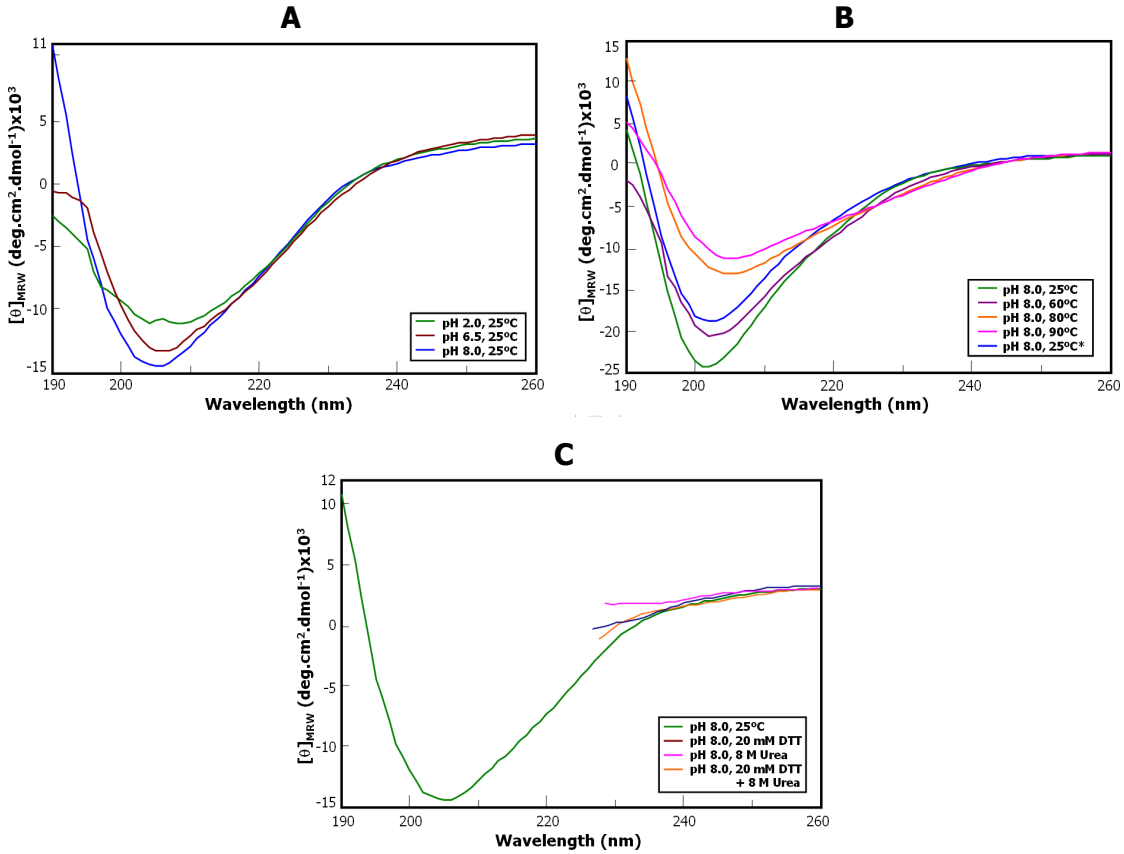


Figure 128. Circular dichroism analyses of natural NvCI.
A.CD analysis of NvCI at different pH values: spectra were recorded at 25ºC in 0.1% v/v aqueous TFA (pH 2.0), 20 mM sodium phosphate buffer pH 6.5 and 20 mM sodium phosphate buffer pH 8.0 at a protein concentration of 0.05 mg/ml. **B.** CD analysis of thermal denaturation of NvCI: assays were performed in 20 mM sodium phosphate buffer pH 8.0 at 25, 60, 80, 90ºC and protein concentration of 0.05 mg/ml. (*) indicates the initial temperature reached again after the heat treatment. **C.** CD analysis of NvCI denaturation by chemical denaturants: assays were performed in 20 mM sodium phosphate buffer pH 8.0 at 25ºC and protein concentration of 0.05 mg/ml, after incubation at 25ºC for 24 h in buffer containing 20 mM DTT, 8 M Urea or both reagents.

CD spectra of thermal denaturation of NvCI are shown in figure 128B. As the temperature was increased from 25 to 90ºC, a decrease in intensity and a progressive rightward shift of the minimum at 205 nm were observed, which could be related to the partial thermal denaturation of the protein. Once the temperature was brought back to 25°C, both the minimum and the position of minimum ellipticity band were partially recovered. However, compared to the initial spectrum of NvCI at 25°C, the spectrum was not completely identical, indicating the occurrence of irreversible conformational changes in NvCI and/or a different reshuffling stage caused during the heat treatment.

It is noteworthy that the pH value at which thermal denaturation of NvCI was performed (pH=8.0) is close to the pKa value of cysteine thiol group, where its reactivity is enhanced and thiol-disulfide exchange takes place. A different oxidation pattern of cysteine residues in the molecule could occur during this thermal denaturation, resulting in a different refolding stage respect to the native protein. In all the studied cases, an almost complete disappearance of CD ellipticity bands were observed, as a result of denaturation with urea, disulfide bonds reduction by DTT, and denaturation-reduction of NvCI (figure 128C).

Circular dichroism spectroscopy of recombinant NvCI showed the same characteristics and behaviours of the natural inhibitor, confirming the correctly folding of recombinant NvCI.

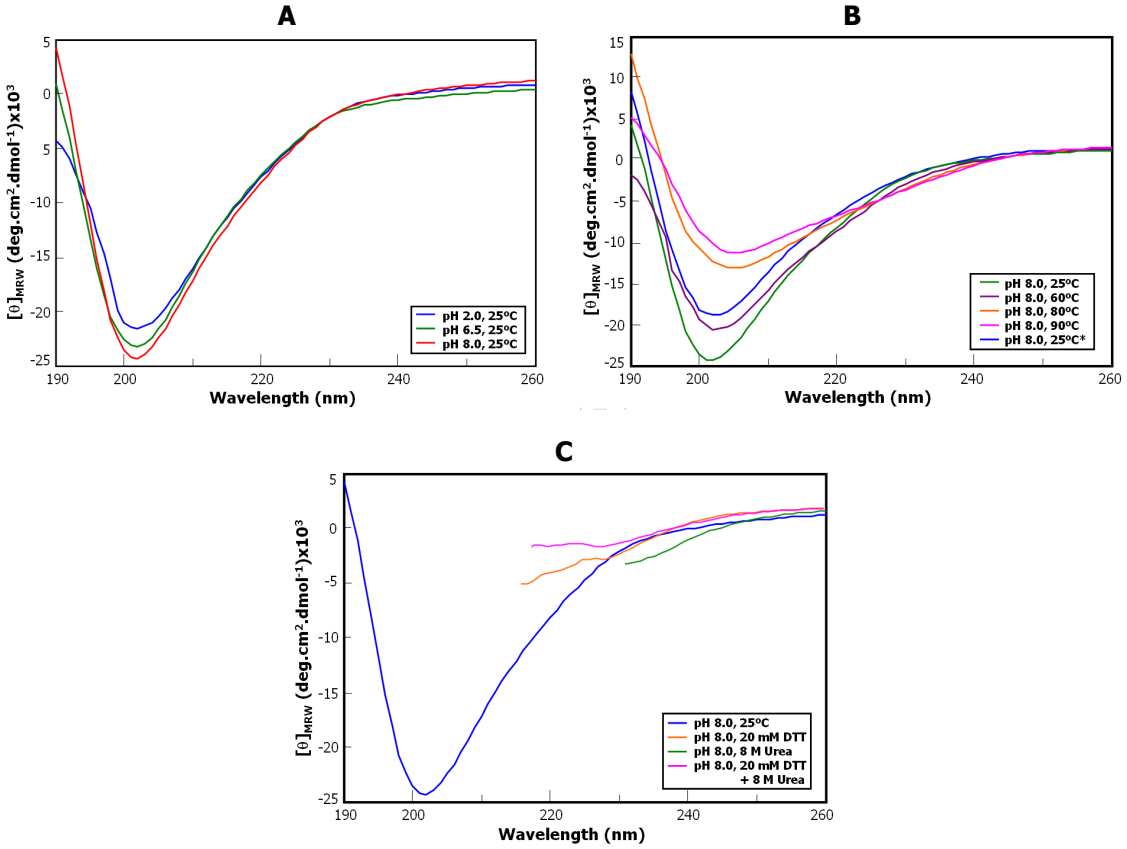


Figure 129. Circular dichroism analyses of recombinant NvCI.

A.CD analysis of rNvCI at different pH values: spectra were recorded at 25ºC in 0.1% v/v aqueous TFA (pH 2.0), 20 mM sodium phosphate buffer pH 6.5 and 20 mM sodium phosphate buffer pH 8.0 at a protein concentration of 0.06 mg/ml. **B.** CD analysis of thermal denaturation of rNvCI: assays were performed in 20 mM sodium phosphate buffer pH 8.0 at 25, 60, 80, 90ºC and protein concentration of 0.06 mg/ml. (*) indicates the initial temperature reached again after the heat treatment. **C.** CD analysis of rNvCI denaturation by chemical denaturants: assays were performed in 20 mM sodium phosphate buffer pH 8.0 at 25ºC and protein concentration of 0.06 mg/ml, after incubation at 25ºC for 24 h in buffer containing 20 mM DTT, 8 M Urea or both reagents.

Figure 129A displays the CD spectra of rNvCI at different pH values, where unlike natural NvCI, a minimum of ellipticity at 202 nm was visualized. Thermal denaturation of rNvCI presented a similar pattern to natural NvCI (figure 129B) with a decrease and rightward shift of the minimum at 202 nm as the temperature was increased. Analysis of chemical denaturation of rNvCI by CD spectroscopy, as the natural inhibitor, displayed a complete destruction of CD ellipticity profile of the native protein after treatment with DTT and Urea (figure 129C).

It is important to be mentioned that, as shown in figures 128B and 129B, representing the thermal denaturation of natural and recombinant NvCI, respectively, a rightward shift in the minimum of ellipticity was observed as temperature increased. Taking into account that a heat treatment (60ºC, 30 min) was carried out during the purification process of natural NvCI, the shift in the minimum ellipticity from 202 to 205 nm compared to the recombinant molecule could be related to partial protein denaturation occurred in that purification step.

IV.8. Kinetic characterization of natural and recombinant NvCI

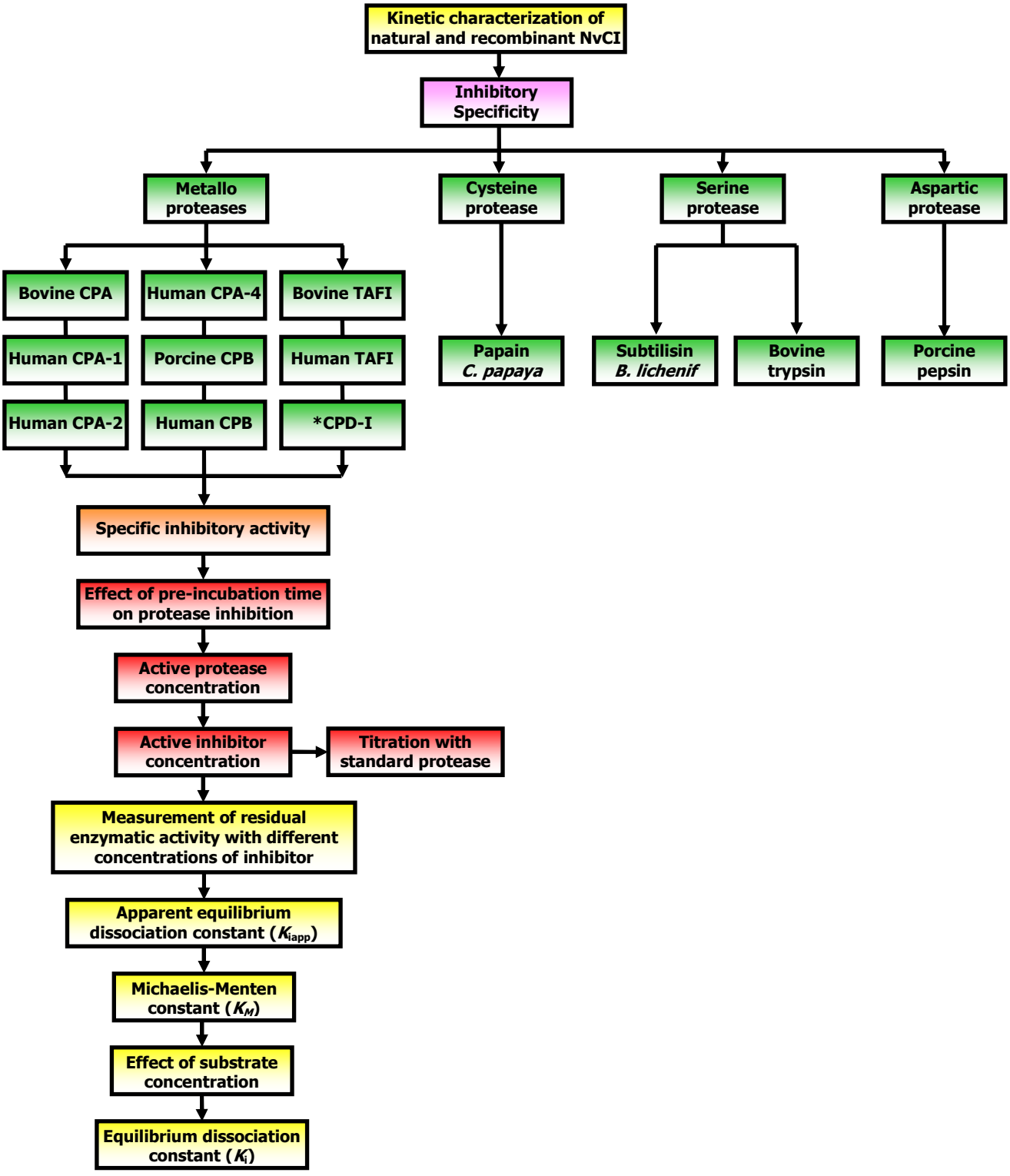


Figure 130. Flowchart process for the kinetic characterization of natural and recombinant NvCI. CPD-I: Carboxypeptidase D domain I from *Drosophila melanogaster*. Subtilisin from *B. licheniformis.* TAFI: thrombin-activable fibrinolysis inhibitor. *NvCI and rNvCI were not active against CPD-I, therefore, the strategy followed by K_i determination was not performed with this enzyme.

Kinetic characterization of native and recombinant NvCI was initiated by studying the specificity of the inhibitor against proteases from different mechanistic classes: trypsin and subtilisin, representatives of two families of serine proteases, papain and pepsin as cysteine and aspartic proteases, respectively and finally, different A- and B-type carboxypeptidases. Subsequently, kinetic characterization against the above established positive to inhibition proteases was performed in terms of specific inhibitory activity and dose-response relationships.

Subsequently, kinetic characterization including the determination of NvCI active concentration, time required to establish reaction equilibrium, analysis of complex dissociation induced by the substrate and the determination of the equilibrium dissociation constant (K_i) was performed against the proteases positive to inhibition.

The determination of K_i values was performed according to the strategy described for tight-binding inhibitors (Bieth, J.G. 1995; Morrison, J.F., 1982) that includes as requirement, prior knowledge of the active inhibitor concentration. For this purpose, selected proteases were initially titrated with a well-known inhibitor, which allowed calculating their active concentrations. Then, active NvCI concentration was determined by titration with several proteases. Taking into account that most of tight-biding inhibitors are slow, time required to reach equilibrium for titration, which is a previous and essential prerequisite, as well as other experimental conditions such as high enzyme concentrations and low substrate concentrations were established.

In a second stage, the determination of apparent K_i values was performed by measuring dose-response relationships ($a=v_i/v_0$ *vs* $[I_0]$) using new experimental conditions according to the requirements established by Morrison equation, such as equilibrium time (generally longer) and low enzyme and substrate concentrations. Finally, the study of the effect of substrate concentrations on the inhibitory activities, in order to demonstrate dissociation of enzyme-inhibitor complexes induced by the substrates, was performed allowing the determination of the real K_i values. The entire process of kinetic characterization of natural and recombinant NvCI is summarized in figure 130.

IV.8.1. Kinetic characterization of natural NvCI

IV.8.1.1. Inhibitory specificity study

Results of the specificity study are shown in table 43. Natural NvCI in the range of 1-500 nM did not show inhibitory activity against serine, cysteine and aspartic proteases used. However, this molecule was active against different metalloproteases: bovine, porcine and human carboxypeptidases A and B. NvCI capacity of inhibiting only MCPs is characteristic of the MCPs inhibitors described with the only exception of SmCI, a BPTI Kunitz inhibitor isolated from the marine invertebrate *S. magnifica* (Alonso del Rivero *et al.*, 2012), which is capable of inhibiting some serine proteases. Therefore, this is the first typical CPA and CPB-like CPs inhibitor described from marine invertebrates.

Table 43. Inhibitory activity of NvCI against various proteases

Protease	$[E_0]$ (nM)	$[I_0]$ (nM)	Res.E.A.(v_i/v_0)
Bovine CPA	3.5	3.0	0.52 ± 0.04
Human CPA-1	2.6	2.5	0.42 ± 0.04
Human CPA-2	11,0	16.5	0.65 ± 0.01
Human CPA-4	26.2	8.5	0.67 ± 0.04
Porcine CPB	1.7	10.5	0.38 ± 0.03
Human CPB	8.4	4.5	0.41 ± 0.01
Bovine TAFI	5.9	3.5	0.34 ± 0.02
Human TAFI	54.2	5.5	0.53 ± 0.02
CPD domain I	16.0	1-160	1.00 ±0.00
Porcine pepsin	10.0	1-500	1.00 ±0.00
Papain	45.0	1-500	1.00 ±0.00
Bovine trypsin	280.0	1-500	1.00 ±0.00
Subtilisin	200.0	1-500	1.00 ±0.00

$[E_0]$ and $[I_0]$: total enzyme and inhibitor concentration. Res.E.A.: Residual Enzymatic Activity. v_i/v_0: fraction of enzymatic activity in the presence (v_i) and absence of inhibitor (v_0) in terms of initial velocities. Date are means (n=3) ± S.D. Other specifications are described in Materials and Methods section

IV.8.1.2. Determination of the active enzyme concentrations

Prior to determining the active NvCI concentration, different CPAs and CPBs were titrated with TCI (tick carboxypeptidase inhibitor, Arolas *et al.*, 2005a). Therefore, a previous study of the effect of preincubation time on residual enzymatic activity (v_i/v_0=a) was carried out. These results are shown in table 44, where in none of the studied cases, an increase in the inhibitory activity (decrease in residual enzymatic activity) until 25 minutes of preincubation time was found, indicating that TCI is not a slow inhibitor in the established experimental conditions for CPs titration.

Table 44. Effect of preincubation time on the inhibitory activity of TCI against various carboxypeptidases

Preincubation time (min)	v_i/v_0 against bCPA	v_i/v_0 against hCPA-1	v_i/v_0 against hCPA-2	v_i/v_0 against hCPA-4	v_i/v_0 against pCPB
1	0.49 ± 0.01	0.34 ± 0.05	0.28 ± 0.04	0.38 ± 0.02	0.72 ± 0.02
5	0.50 ± 0.01	0.37 ± 0.02	0.30 ± 0.01	0.39 ± 0.01	0.72 ± 0.02
10	0.54 ± 0.04	0.35 ± 0.03	0.29 ± 0.02	0.38 ± 0.01	0.73 ± 0.01
15	0.52 ± 0.04	0.39 ± 0.02	0.31 ± 0.03	0.38 ± 0.03	0.75 ± 0.02
20	0.51 ± 0.05	0.39 ± 0.02	0.31 ± 0.03	0.42 ± 0.03	0.73 ± 0.01
25	0.53 ± 0.02	0.40 ± 0.02	0.32 ± 0.02	0.40 ± 0.01	0.74 ± 0.01

Experimental conditions: [bCPA]=25.0 nM / [TCI]=11.0 nM, [hCPA-1]=22.0 nM / [TCI]=10.0 nM, [hCPA-2]=48.0 nM / [TCI]=28.0 nM, [hCPA-4]=45.0 nM / [TCI]=12.0 nM, [pCPB]=17.5 nM / [TCI]=4.0 nM. Other specifications are described in Materials and Methods. v_i/v_0: fraction of enzymatic activity in the presence (v_i) and absence of inhibitor (v_0) in terms of initial velocities. Data are means (n=3) ± S.D. One-way ANOVA did not find significant differences between groups at $p<0.05$.

Under selected experimental conditions of preincubation time at 15 min, enzyme and substrate concentrations, CPs titrations by TCI were performed and percentages of active concentrations determined. Figure 131 shows the results obtained with hCPA1, where the equivalence point is reached at 18.9 nM (enzyme concentration is equal to inhibitor concentration), allowing to calculate its active concentration. Table 45 summarizes concentration percentages of active enzymes. Titration curves of the other CPs are shown in annex IV. It is important to point out that the percentage of active concentrations for the majority of CPs were up to 85% with the exception of human CPA-4, human CPB and bovine TAFI for which values less than 50% were obtained.

These results could be explained in the case of TAFI for its low half-life (Sanglas *et al.*, 2010). On the other hand, in the case of human CPA-4 and human CPB, these results could be related to the purification procedures of these enzymes, leading to partial denaturation.

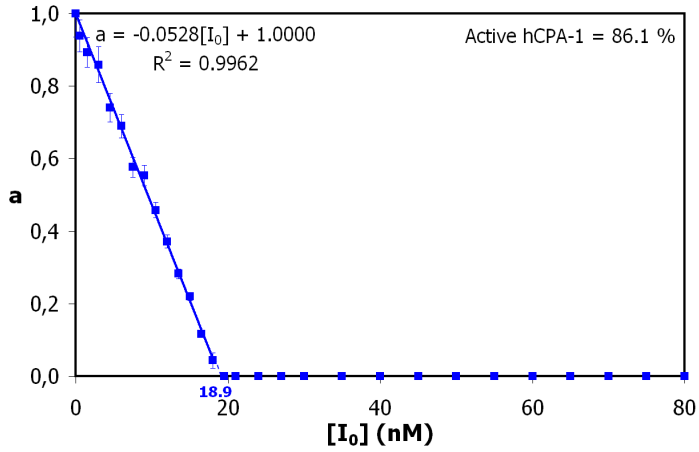


Figure 131. hCPA1 titration curve with tick carboxypeptidase inhibitor (TCI).
[hCPA-1]=22.0 nM, $[E_0]/K_i$=18.3. Substrate: AAFP. Preincubation time: 15 min, T=37ºC. Other specifications are described in Materials and Methods. $a=v_i/v_0$: fraction of enzymatic activity in the presence (v_i) and absence of inhibitor (v_0) in terms of initial velocities. Data are means (n=3) ± S.D.

Figure 45. Enzyme titrations with tick carboxypeptidase inhibitor (TCI)

Enzyme	$[E_0]_{total}$ (nM)	$[E_0]_{active}$ (nM)	Percentage of active enzyme (%)
bCPA	25.0	22.2	88.9
hCPA-1	22.0	18.9	86.1
hCPA-2	48.0	44.4	92.6
hCPA-4	45.0	21.5	47.8
pCPB	17.5	16.2	92.5
hCPB	29.0	14.1	48.8
bTAFI	95.0	28.9	30.4

[bCPA]=25.0 nM, $[E_0]/K_i$=22.7. [hCPA-1]=22.0 nM, $[E_0]/K_i$=18.3. [hCPA-2]=48.0 nM, $[E_0]/K_i$=13.3. [hCPA-4]=45.0 nM, $[E_0]/K_i$=56.2. [pCPB]=17.5 nM, $[E_0]/K_i$=10.7. [hCPB]=29.0 nM, $[E_0]/K_i$=22.2. [bTAFI]=95.0 nM, $[E_0]/K_i$=73.1.Substrate: AAFP. Preincubation time: 15 min, T=37ºC. Other conditions are described in Materials and Methods. $a=v_i/v_0$: fraction of enzymatic activity in the presence (v_i) and absence of inhibitor (v_0) in terms of initial velocities. Data are means (n=3) ± S.D.

IV.8.1.3. Determination of active NvCI concentration

Active NvCI concentration was determined by titration of the inhibitor with the enzymes previously titrated with TCI. For this purpose, a similar procedure to the enzyme titration was used and experimental conditions such as time to reach equilibrium for each CP and NvCI were established.

Table 46 shows the results of the effect of preincubation time of NvCI against hCPA1, where significant differences in vi/vo values between 5 and 10 min were obtained, indicating a small decrease in residual activity (increase of inhibitory activity) at 10 min of incubation. Therefore, NvCI titration with this enzyme was performed using a preincubation time of 15 min. Results obtained with other enzymes (except for human TAFI) were similar in terms of percentage of active inhibitor. Figure 132 shows the titration curve of NvCI with hCPA1, where the percentage of active inhibitor found to be 50.4%. Results obtained with other enzymes (except for human TAFI) were similar in terms of percentage of active inhibitor.

Table 46. Effect of preincubation time on the inhibitory activity of NvCI against hCPA1

Preincubation time (min)	v_i/v_0 against hCPA-1
1	0.47 ± 0.02 b
5	0.46 ± 0.01 b
10	0.43 ± 0.01 c
15	0.44 ± 0.01 c
20	0.42 ± 0.02 c
25	0.431 ± 0.004 c

Experimental conditions: [hCPA-1]=19.5 nM / [NvCI]=19.0 nM. Other specifications are described in Materials and Methods. v_i/v_0: fraction of enzymatic activity in the presence (v_i) and absence of inhibitor (v_0) in terms of initial velocities. Data are means (n=3) ± S.D. One-way ANOVA found significant differences between groups at $p<0.05$. Tukey HSD test found significant differences between sample means at $p<0.05$. Different letters indicate significant differences at $p<0.05$.

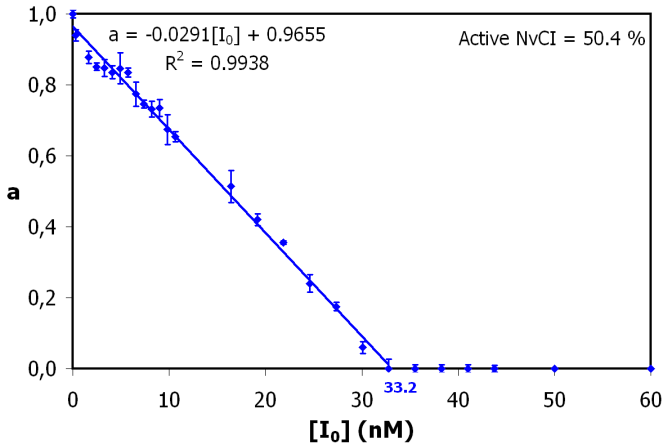


Figure 132. NvCI titration curve with hCPA1.
[hCPA-1]=19.5 nM. Substrate: AAFA. Preincubation time: 15 min, T=37ºC. Other specifications are described in Materials and Methods. $a=v_i/v_0$: fraction of enzymatic activity in the presence (v_i) and absence of inhibitor (v_0) in terms of initial velocities. Data are means (n=3) ± S.D.

IV.8.1.4. Determination of K_i values of NvCI against various carboxypeptidases

Once the active inhibitor concentration was known, determination of K_i values for selected carboxypeptidases was performed. Prior to this procedure, time required to reach equilibrium was evaluated under these new conditions: low enzyme and substrate concentrations. Results obtained with hCPA-1 confirmed that NvCI behaves as a slow inhibitor against this enzyme (equilibrium time of 10 min). For bCPA, hCPA2 and hCPA-4 no significant differences were found between the studied incubation times, while in the case of pCPB, significant differences were only determined for 5 minutes of preincubation time (table 47). However, the value of pCPB inhibitory activity found at preincubation time of 1 min seems to be an outlier, given the considerable difference only found between this value when compared to the rest of the data.

Selected preincubation time was 15 minutes for all the analyzed enzymes, except for human CPB and bovine and human TAFI. In the case of human CPB the presence of trypsin used for its activation could affect the inhibitory activity results due to degradation of the enzyme and the inhibitor during the preincubation time. On the other hand, a half-life of few minutes for bovine and human TAFI has been described (Sanglas *et al.*, 2010). For this reason, these two latter enzymes were not incubated with the inhibitor.

Table 47. Effect of preincubation time on the inhibitory activity of NvCI against various carboxypeptidases

Preincubation time (min)	v_i/v_0 against bCPA	v_i/v_0 against hCPA-1	v_i/v_0 against hCPA-2	v_i/v_0 against hCPA-4	v_i/v_0 against pCPB
1	0.59 ± 0.03	0.45 ± 0.02 b	0.63 ± 0.02	0.53 ± 0.01	0.36 ± 0.03 b
5	0.53 ± 0.01	0.44 ± 0.02 b	0.66 ± 0.03	0.52 ± 0.02	0.46 ± 0.01 c
10	0.53 ± 0.03	0.391 ± 0.004 c	0.70 ± 0.02	0.55 ± 0.01	0.47 ± 0.02 c
15	0.57 ± 0.04	0.39 ± 0.01 c	0.69 ± 0.06	0.53 ± 0.04	0.47 ± 0.02 c
20	0.56 ± 0.03	0.39 ± 0.02 c	0.70 ± 0.02	0.53 ± 0.02	0.46 ± 0.01 c
25	0.56 ± 0.05	0.39 ± 0.02 c	0.69 ± 0.01	0.53 ± 0.03	0.46 ± 0.01 c

Experimental conditions: [bCPA]=3.5 nM / [NvCI]=4.0 nM, [hCPA-1]=2.6 nM / [NvCI]=2.7 nM, [hCPA-2]=11.0 nM / [NvCI]=11.0 nM, [hCPA-4]=26.2 nM / [NvCI]=12.0 nM, [pCPB]=1.7 nM / [NvCI]=6.5 nM. Other specifications are described in Materials and Methods. v_i/v_0: fraction of enzymatic activity in the presence (v_i) and absence of inhibitor (v_0) in terms of initial velocities. Data are means (n=3) ± S.D. One-way ANOVA found significant differences between groups at $p<0.05$ for hCPA-1 and pCPB. Tukey HSD test found significant differences between sample means at $p<0.05$ for hCPA-1 and pCPB. Different letters indicate significant differences at $p<0.05$.

Knowledge of Michaelis-Menten constant (K_M) is a requirement to calculate real K_i values. To this purpose, K_M values were calculated for those enzymes whose values are not described or determined in different experimental conditions used in this work (tables 48 and 49). The obtained $[S_0]$ *vs* v_0 curves are shown in annex V.

Table 48. Michaelis-Menten constant (K_M) for AAFP hydrolysis catalyzed by various A-type carboxypeptidases

Enzyme	$[E_0]$ (nM)	$[S_0]$ range (μM)	K_M (μM)
Bovine CPA	3.5	0 – 200	146.2 ± 10.4
Human CPA1	2.6	0 – 200	43.3 ± 1.8
Human CPA2	6.8	0 – 210	98.5 ± 14.6
Human CPA4	22.6	0 – 200	57.6 ± 1.8

Proteolytic activity was followed by its ability to hydrolyse AAFP synthetic substrate in a final volume of 250 μl, buffer: 20 mM Tris-HCl pH 7.5, 500 mM NaCl, 1% v/v DMSO, 0.05% w/v BRIJ-35 and T=37ºC. The best-fit value of K_M was performed by adjusting the experimental values to the Michaelis-Menten equation (Eq. 22) using the program GraphPad Prism 5 (GraphPad Software, Inc.) at $p<0.05$. Data are means (n=3) ± S.D.

Table 49. Michaelis-Menten constant (K_M) for AAFA hydrolysis catalyzed by various B-type carboxypeptidases

Enzyme	$[E_0]$ (M)	$[S_0]$ range (M)	K_M (M)
Porcine CPB	1.7	0 – 200	37.0 ± 3.3
Human CPB	8.4	0 – 200	33.2 ± 0.7
Bovine TAFI	14.0	0 – 200	79.7 ± 6.7

Proteolytic activity was followed by its ability to hydrolyse AAFA synthetic substrate in a final volume of 250 μl, buffer: 20 mM Tris-HCl pH 7.5, 100 mM NaCl, 1% v/v DMSO, 0.05% w/v BRIJ-35 and T=37ºC. The best-fit value of K_M was performed by adjusting the experimental values to the Michaelis-Menten equation (Eq. 22) using the program GraphPad Prism 5 (GraphPad Software, Inc.) at $p<0.05$. Data are means (n=3) ± S.D.

Prior to determine Ki values, it is interesting to experimentally study the effect of substrate concentrations over the dissociation of the enzyme-inhibitor complexes. This effect is generally produced by competitive inhibitors, such as NvCI, demonstrated by the elucidation of its X-ray structure in complex with hCPA4 (see later).

Analysis of the effect of substrate concentrations on the residual enzymatic activity (inverse of inhibitory activity) showed a similar behavior for almost all proteases, where at least for $[S_0]/K_M$=4.5 in some cases there was not evidenced any effect. For hCPA2, an increase of residual enzymatic activity (indicative of dissociation) was determined. This behaviour could be associated with its lower inhibitory capacity (10-1000 times lower in terms of K_i, compared with the rest of the MCPs studied), being more susceptible to substrate dissociation in these conditions. In the case of hCPA4, a slight decrease in residual activity was observed, attributable to other factors unrelated to dissociation phenomenon. Figures 133 and 134 summarize the results obtained.

It is important to note that the apparent absence of dissociation induced by the substrate in the experimental range studied for the MCPs used, with the exception of hCPA4, is not a demonstration that substrate is not able to induce it in other conditions. This is supported by the competitive character of the inhibitor. In addition, dissociation could be induced during the mix of the reagents in the assay, prior to the first measurement of residual activity (Bieth, J.G., 1995). Therefore, the apparent K_i values in all cases must be rearranged for this effect, according to the equation described by Morrison, J.F., 1982 (Eq. Nº 24).

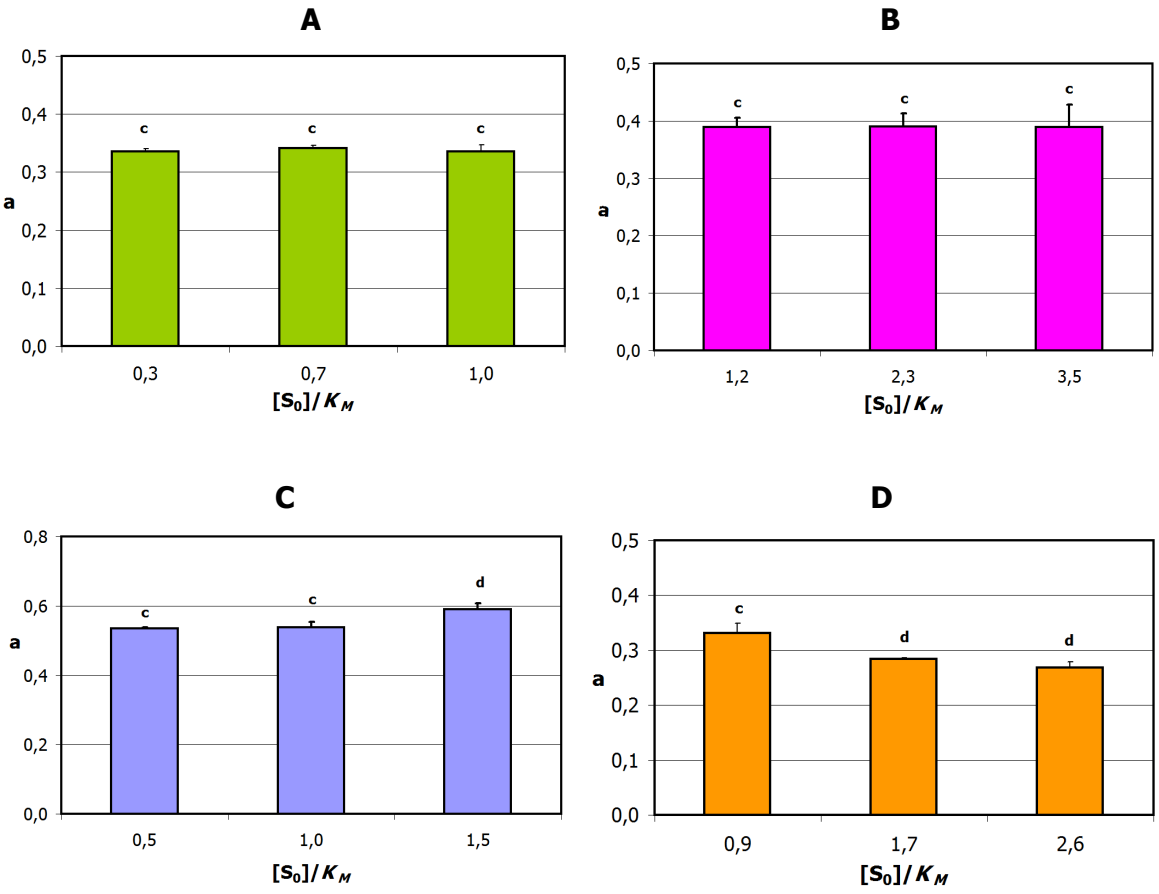


Figure 133. Effect of the substrate concentration on the inhibitory activity of NvCI against various A-type carboxypeptidases.
A. [bCPA]=3.5 nM, [NvCI]=4.0 nM. **B.** [hCPA-1]=2.6 nM, [NvCI]=2.7 nM. **C.** [hCPA-2]=11.0 nM, [NvCI]=25.0 nM. **D.** [hCPA-4]=26.2 nM, [NvCI]=8.5 nM. Substrate: AAFP. Other specifications are described in Materials and Methods. $a=v_i/v_0$: fraction of enzymatic activity in the presence (v_i) and absence of inhibitor (v_0) in terms of initial velocities. Data are means (n=3) ± S.D. One-way ANOVA found significant differences between groups at $p<0.05$ for hCPA-2. Tukey HSD test found significant differences between sample means at $p<0.05$ for hCPA-2. Different letters indicate significant differences at $p<0.05$.

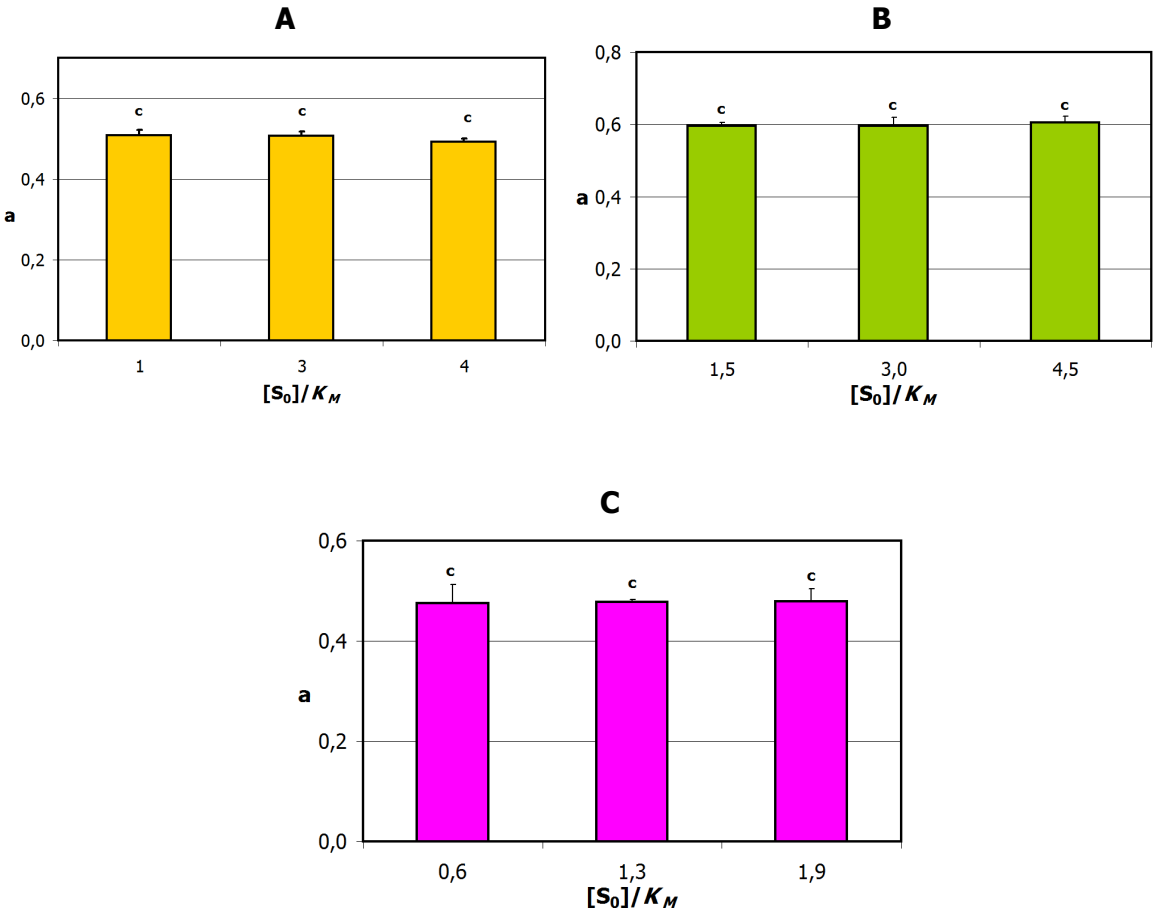


Figure 134. Effect of the substrate concentration on the inhibitory activity of NvCI against various B-type carboxypeptidases.

A. [pCPB]=1.7 nM, [NvCI]=5.5 nM. **B.** [hCPB]=8.4 nM, [NvCI]=1.5 nM. **C.** [bTAFI]=5.9 nM, [NvCI]=1.5 nM. Substrate: AAFA. Other specifications are described in Materials and Methods. $a=v_i/v_0$: fraction of enzymatic activity in the presence (v_i) and absence of inhibitor (v_0) in terms of initial velocities. Data are means (n=3) ± S.D. One-way ANOVA found significant differences between groups at $p<0.05$. Tukey HSD test found significant differences between sample means at $p<0.05$. Different letters indicate significant differences at $p<0.05$.

The determination of K_i according Morrison, J.F., 1982 was performed by measuring the fractional velocities ($a=v_i/v_0$) vs [NvCI], using equilibrium experimental conditions, such as preincubation time, which is generally longer, due to a more diluted environment with lower enzyme concentrations and maintaining, in most cases, the same low substrate concentrations used during titration. Figures 135 and 136 show dose-response relationships of NvCI against all the analyzed carboxypeptidases. Adjusting these values to the Morrison equation (Eq. Nº 23) allowed the determination of apparent K_i values. These values were subsequently rearranged assuming dissociation induced by substrate, according to the equation described by Morrison, J.F., 1982 (Eq. Nº 24), using the K_M values for each enzyme against its substrate and the substrate concentration in equilibrium conditions.

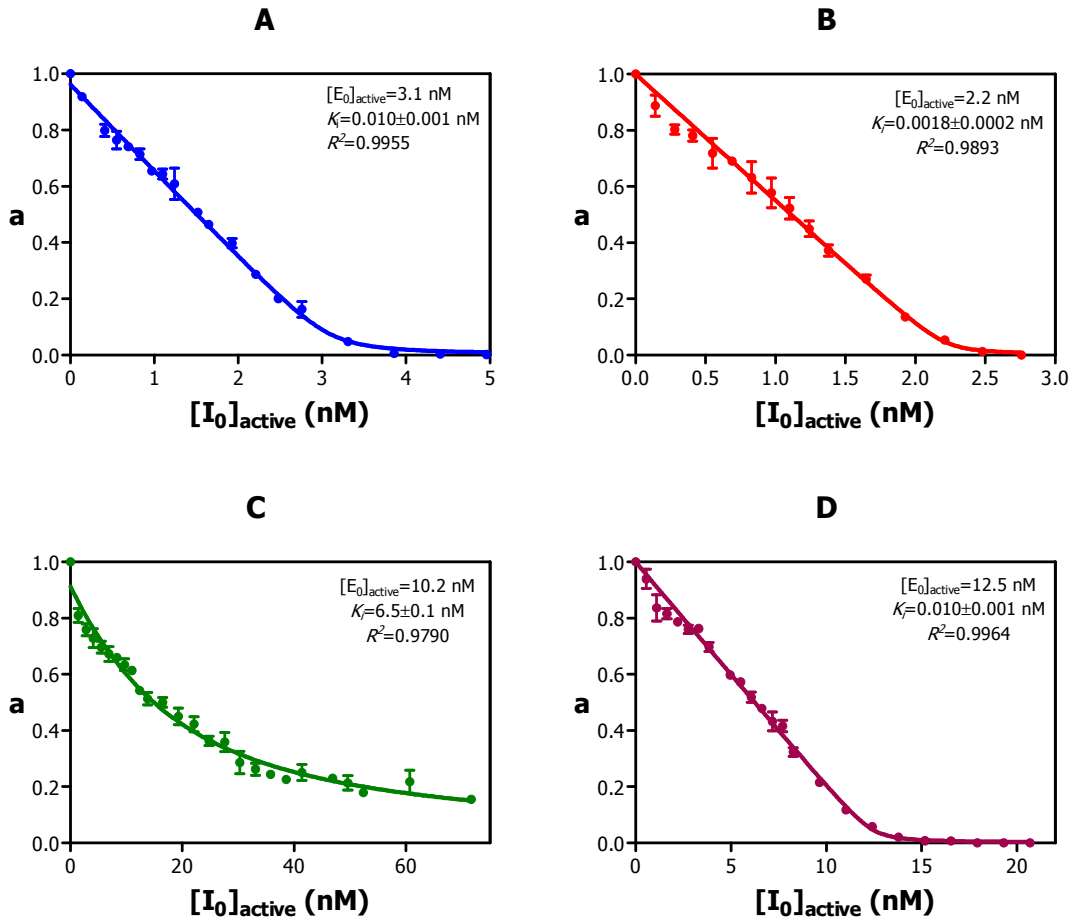
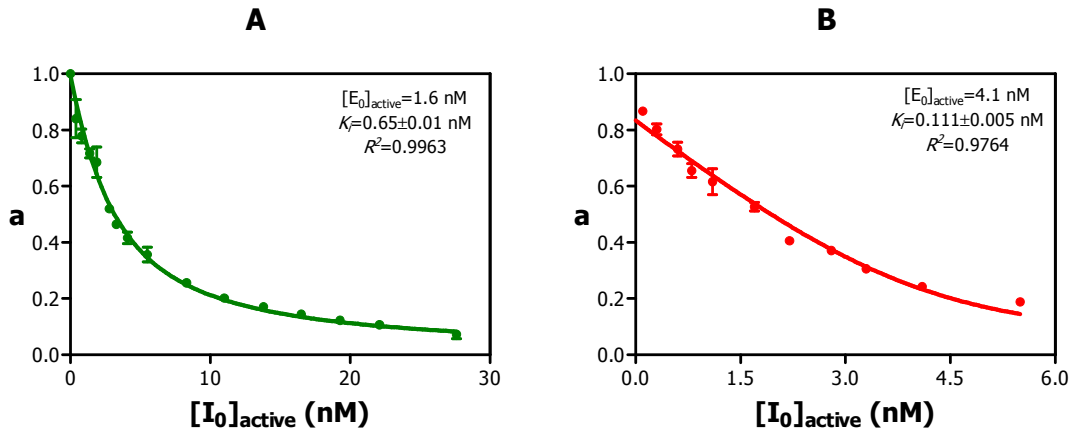


Figure 135. Equilibrium dissociation constant (K_i) of NvCI against various A-type carboxypeptidases. (A). [bCPA]=3.5 nM. **(B).** [hCPA-1]=2.6 nM. **(C).** [hCPA-2]=11.0 nM. **(D).** [hCPA-4]=26.2 nM. Substrate: AAFP. Preincubation time: 15 min, T=37ºC. Other specifications are described in Materials and Methods. $a=v_i/v_0$: fraction of enzymatic activity in the presence (v_i) and absence of inhibitor (v_0) in terms of initial velocities. The best-fit value of K_i was performed by adjusting the experimental values to the Morrison equation (Eq. 23) using the program GraphPad Prism 5 (GraphPad Software, Inc.) at $p<0.05$. Data are means (n=3) ± S.D.



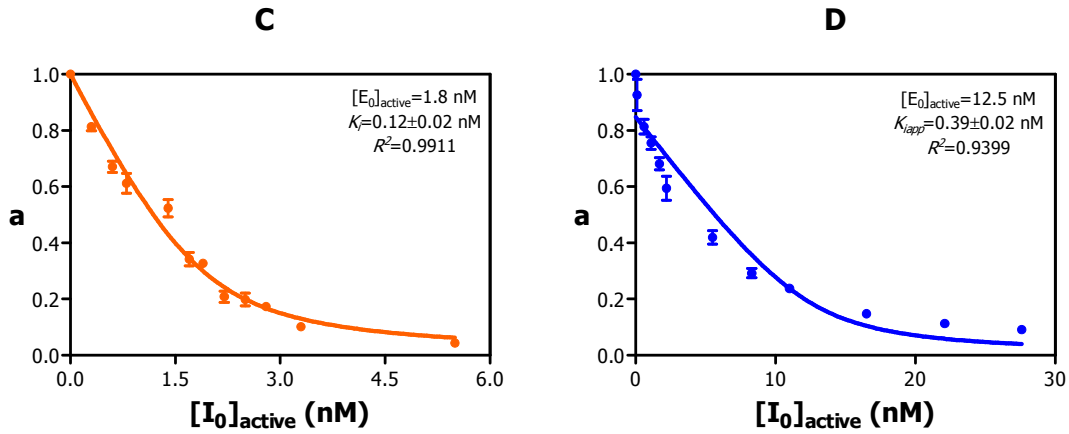


Figure 136. Equilibrium dissociation constant (K_i) of NvCI against various B-type carboxypeptidases. (A). [pCPB]=1.7 nM. **(B).** [hCPB]=8.4 nM. **(C).** [bTAFI]=5.9 nM. **(D).** [hTAFI]=54.2 nM. Substrate: AAFA. Preincubation time: 15 min, T=37ºC. Other specifications are described in Materials and Methods. $a=v_i/v_0$: fraction of enzymatic activity in the presence (v_i) and absence of inhibitor (v_0) in terms of initial velocities. The best-fit value of K_i was performed by adjusting the experimental values to the Morrison equation (Eq. 23) using the program GraphPad Prism 5 (GraphPad Software, Inc.) at $p<0.05$. Data are means (n=3) ± S.D.

IV.8.2. Kinetic characterization of recombinant NvCI

IV.8.2.1. Inhibitory specificity study

Recombinant NvCI was kinetically characterized according to the procedure used for the natural inhibitor. The specificity displayed by this molecule also proved to be similar to that obtained by the natural inhibitor, indicating that NvCI is a specific carboxypeptidase inhibitor which has no ability to inhibit proteases from other mechanistic classes. This greater specificity is an advantage for this molecule. Results of the specificity study are presented in table 50.

Table 50. Inhibitory activity of rNvCI against various proteases

Protease	$[E_0]$ (nM)	$[I_0]$ (nM)	Res.E.A.(v_i/v_0)
Bovine CPA	2.0	0.7	0.59 ±0.01
Human CPA-1	4.8	0.6	0.756 ±0.005
Human CPA-2	7.4	20.0	0.31 ±0.01
Human CPA-4	22.6	5.0	0.70 ±0.01
Porcine CPB	1.9	2.5	0.38 ±0.01
Human CPB	8.9	2.0	0.62 ±0.03
Bovine TAFI	23.0	4.0	0.49 ±0.02
Human TAFI	31.2	4.0	0.66 ±0.03
CPD domain I	16.0	1-160	1.00 ±0.00
Porcine pepsin	10.0	1-500	1.00 ±0.00
Papain	45.0	1-500	1.00 ±0.00
Bovine trypsin	280.0	1-500	1.00 ±0.00
Subtilisin	200.0	1-500	1.00 ±0.00

$[E_0]$ and $[I_0]$: total enzyme and inhibitor concentration. Res.E.A.: Residual Enzymatic Activity. v_i/v_0: fraction of enzymatic activity in the presence (v_i) and absence of inhibitor (v_0) in terms of initial velocities. Other specifications are described in Materials and Methods section

IV.8.2.2. Determination of active enzyme concentrations

Given that enzyme stocks were different, the study of preincubation time of MCPs with TCI on titration condition ($[E_0]/K_i \geq 100$) was performed again. The behavior was similar to that previously obtained, since there were no significant differences between the different preincubation times, except for hCPA1, where an outlier at 10 minutes of preincubation time was determined (table 51).

Table 51. Effect of preincubation time on the inhibitory activity of TCI against various carboxypeptidases

Preincubation time (min)	v_i/v_0 against bCPA	v_i/v_0 against hCPA-1	v_i/v_0 against hCPA-2	v_i/v_0 against hCPA-4	v_i/v_0 against pCPB
1	0.46 ± 0.01	0.50 ± 0.01 b	0.60 ± 0.01	0.33 ± 0.03	0.70 ± 0.01
5	0.47 ± 0.02	0.49 ± 0.01 b	0.57 ± 0.02	0.34 ± 0.02	0.70 ± 0.01
10	0.45 ± 0.02	0.47 ± 0.01 c	0.57 ± 0.01	0.32 ± 0.02	0.72 ± 0.01
15	0.472 ± 0.003	0.50 ± 0.01 b	0.555 ± 0.002	0.34 ± 0.02	0.73 ± 0.01
20	0.46 ± 0.01	0.50 ± 0.01 b	0.56 ± 0.02	0.33 ± 0.02	0.70 ± 0.02
25	0.46 ± 0.01	0.50 ± 0.01 b	0.57 ± 0.03	0.32 ± 0.01	0.70 ± 0.03

Experimental conditions: [bCPA]=20.0 nM / [TCI]=11.0 nM, [hCPA-1]=26.0 nM / [TCI]=7.0 nM, [hCPA-2]=27.0 nM / [TCI]=12.0 nM, [hCPA-4]=73.5 nM / [TCI]=40.0 nM, [pCPB]=17.0 nM / [TCI]=3 nM. Other specifications are described in Materials and Methods. v_i/v_0: fraction of enzymatic activity in the presence (v_i) and absence of inhibitor (v_0) in terms of initial velocities. Data are means (n=3) ± S.D. One-way ANOVA found significant differences between groups at $p<0.05$ for hCPA-1- Tukey HSD test found significant differences between sample means at $p<0.05$ for hCPA-1. Different letters indicate significant differences at $p<0.05$.

Figure 137 shows the titration curve of hCPA1 and TCI. The equivalence point was reached at 14.1 nM, representing 53.9% of active enzyme. This percentage showed a lower value than that obtained with the enzyme stock previously used, demonstrating that enzyme titration must be repeated when using different preparations. It is important to note that the percentages of enzyme active concentration depends on several factors including their production process, purification, conservation, purity, lability, among others. Table 52 summarizes the concentration and percentage of active concentration of all enzymes used except for hCPB, bTAFI and hTAFI because the same stocks previously titrated were used. Titration curves for other enzymes are shown in annex VI.

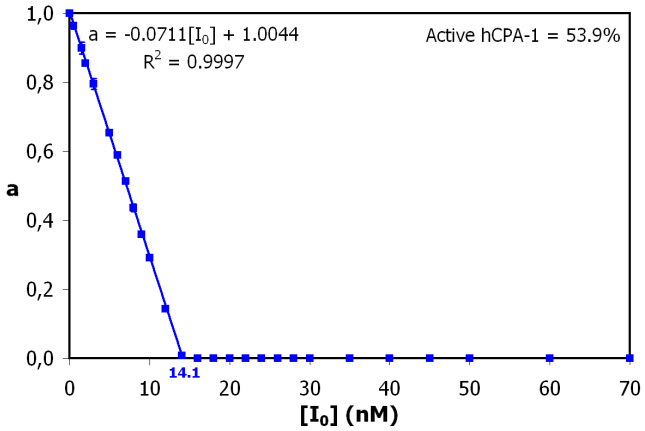


Figure 137. hCPA1 titration curve with tick carboxypeptidase inhibitor (TCI)
[hCPA-1]=26.0 nM, $[E_0]/K_i$=21.7. Substrate: AAFP. Preincubation time: 15 min, T=37ºC. Other specifications are described in Materials and Methods. $a=v_i/v_0$: fraction of enzymatic activity in the presence (v_i) and absence of inhibitor (v_0) in terms of initial velocities. Data are means (n=3) ± S.D.

Table 52. Enzyme titration with tick carboxypeptidase inhibitor (TCI)

Enzyme	$[E_0]_{total}$ (nM)	$[E_0]_{active}$ (nM)	Percentage of active enzyme (%)
bCPA	20.0	20.0	100.0
hCPA-1	26.0	14.1	53.9
hCPA-2	27.0	26.7	99.2
hCPA-4	73.5	57.5	78.0
pCPB	17.0	9.2	53.7

[bCPA]=20.0 nM, $[E_0]/K_i$=18.2. [hCPA-1]=26.0 nM, $[E_0]/K_i$=21.7. [hCPA-2]=27.0 nM, $[E_0]/K_i$=7.5. [hCPA-4]=73.5 nM, $[E_0]/K_i$=91.9 nM, pH=7.5. [pCPB]=17.0 nM, $[E_0]/K_i$=10.7. Substrate: AAFP. Preincubation time: 15 min, T=37ºC. Other specifications are described in Materials and Methods. $a=v_i/v_0$: fraction of enzymatic activity in the presence (v_i) and absence of inhibitor (v_0) in terms of initial velocities. Data are means (n=3) ± S.D.

IV.8.2.3. Determination of active rNvCI concentration

hCPA1 was selected for NvCI titration, based on previous experience with the natural inhibitor, where similar results were obtained with different proteases. The effect of preincubation time as shown in table 53 indicates no significant differences between the residual enzymatic activity values measured at different times.

However, mixtures of hCPA1–NvCI were preincubated 15 min for the titration procedure as in the case of natural NvCI. Active rNvCI concentration determined in the equivalence point showed a value of 14.3 nM, representing 100% of active recombinant inhibitor (figure 138), which is an excellent result compared to the natural inhibitor (50.4%). Natural inhibitor is a molecule present in a very complex natural medium and submitted to relatively harmful processes since its isolation from the animal cells to drastic conditions, particularly reversed phase HPLC could attempt to its structure integrity. In contrast, recombinant inhibitor was produced and purified in milder conditions. In addition, this value also evidences the quality of the production and purification processes of rNvCI, as well as its stability.

Table 53. Effect of preincubation time on the inhibitory activity of rNvCI against hCPA1

Time (min)	V_i/V_0 against hCPA1
1	0.709 ± 0.003
5	0.718 ± 0.003
10	0.71 ± 0.02
15	0.73 ± 0.01
20	0.72 ± 0.01
25	0.73 ± 0.01

Experimental conditions: [hCPA-1]=26.5 nM / [rNvCI]=4.0 nM. Other specifications are described in Materials and Methods. v_i/v_0: fraction of enzymatic activity in the presence (v_i) and absence of inhibitor (v_0) in terms of initial velocities. Data are means (n=3) ± S.D. One-way ANOVA did not find significant differences between groups at $p<0.05$.

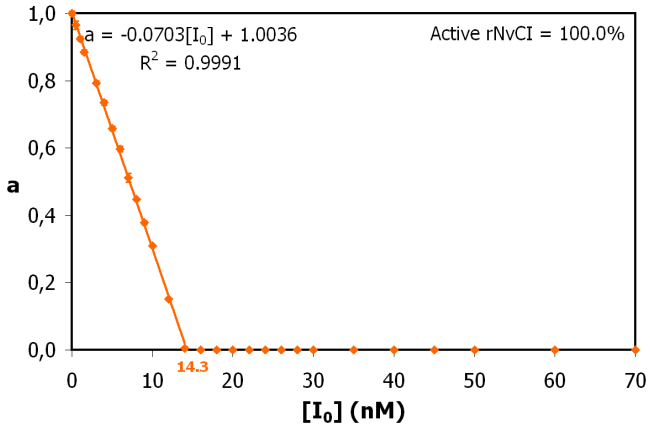


Figure 138. rNvCI titration curve with hCPA1

[hCPA-1]=26.5 nM. Substrate: AAFA. Preincubation time: 15 min, T=37ºC. Other specifications are described in Materials and Methods. $a=v_i/v_0$: fraction of enzymatic activity in the presence (v_i) and absence of inhibitor (v_0) in terms of initial velocities. Data are means (n=3) ± S.D.

IV.8.2.4. Determination of K_i values of rNvCI against various carboxypeptidases

At equilibrium conditions ($[E_0]/K_i \leq 10$), time to establish equilibrium towards different CPs were evaluated. Results are shown in table 54. No significant differences for bCPA, hCPA-1 and hCPA-2 were found, unlike hCPA-4 and pCPB, where equilibrium times of 5 and 15 min were respectively determined.

Table 54. Effect of preincubation time on the inhibitory activity of rNvCI against various carboxypeptidases

Preincubation time (min)	v_i/v_0 against bCPA	v_i/v_0 against hCPA1	v_i/v_0 against hCPA2	v_i/v_0 against hCPA4	v_i/v_0 against pCPB
1	0.35 ± 0.01	0.79 ± 0.02	0.60 ± 0.04	0.346 ± 0.002 b	0.27 ± 0.01 b
5	0.38 ± 0.01	0.79 ± 0.02	0.598 ± 0.002	0.33 ± 0.01 c	0.25 ± 0.01 b
10	0.38 ± 0.01	0.78 ± 0.02	0.61 ± 0.01	0.31 ± 0.02 c	0.267 ± 0.002 b
15	0.36 ± 0.01	0.79 ± 0.01	0.63 ± 0.02	0.32 ± 0.01 c	0.29 ± 0.01 c
20	0.37 ± 0.01	0.78 ± 0.02	0.62 ± 0.01	0.31 ± 0.01 c	0.282 ± 0.004 c
25	0.36 ± 0.01	0.794 ± 0.002	0.63 ± 0.02	0.31 ± 0.01 c	0.28 ± 0.01 c

Experimental conditions: [bCPA]=2.0 nM / [rNvCI]=1.3 nM, [hCPA1]=4.8 nM / [rNvCI]=0.5 nM, [hCPA2]=7.4 nM / [rNvCI]=10.0 nM, [hCPA4]=22.6 nM / [rNvCI]=12.0 nM, [pCPB]=1.9 nM / [rNvCI]=3.5 nM. Other specifications are described in Materials and Methods. v_i/v_0: fraction of enzymatic activity in the presence (v_i) and absence of inhibitor (v_0) in terms of initial velocities. Data are means (n=3) ± S.D. One-way ANOVA found significant differences between groups at p<0.05 for hCPA-4 and pCPB. Tukey HSD test found significant differences between sample means at p <0.05 for hCPA-4 and pCPB. Different letters indicate significant differences at p <0.05.

The study of the effect of substrate concentrations on the residual enzymatic activities is shown in figures 139 and 140, where in most cases no significant differences in v_i/v_0 values were observed, at least in the range of $[S_0]/K_M$ used. For hCPA-2 and pCPB, an increase in residual enzymatic activity was observed, as a result of dissociation induced by the substrate. In other cases a decrease of residual enzymatic activity was evaluated, although as mentioned above, it could be related to other causes. As it was mentioned before, independently of these results, rearrangement of the K_i values were performed, taking into account its reversible competitive character (see below).

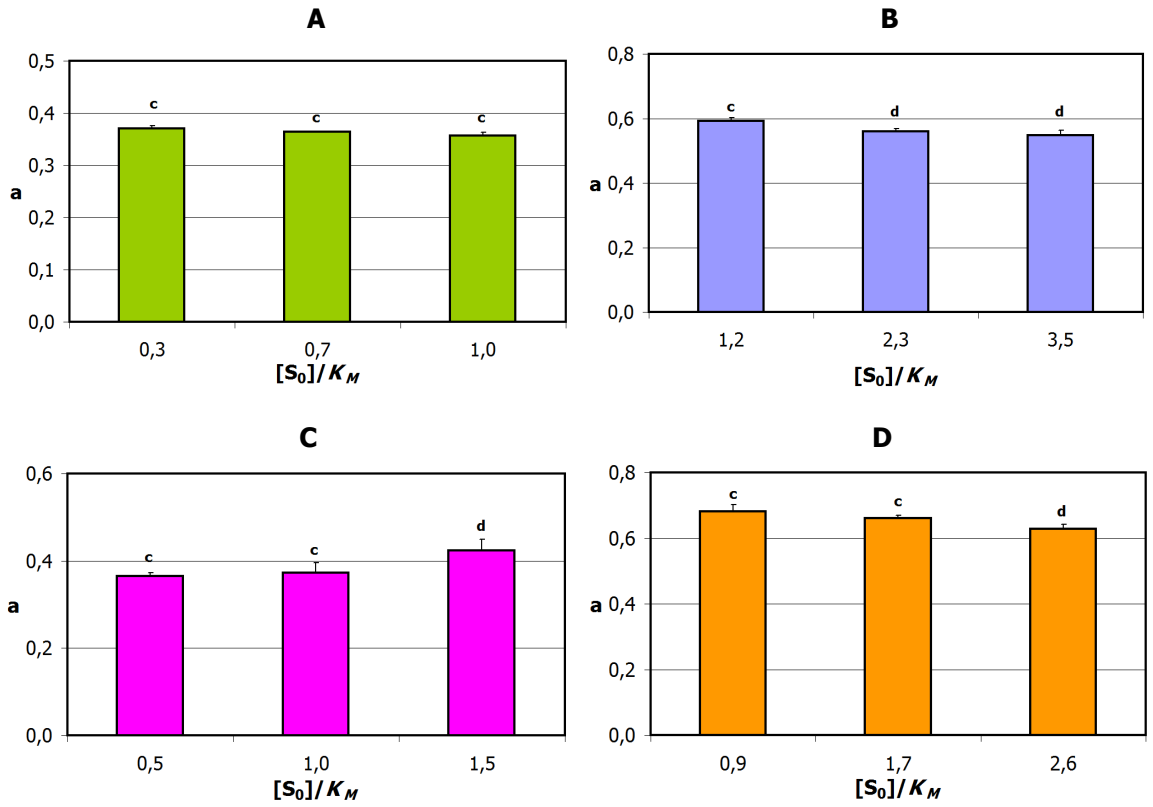


Figure 139. Effect of the substrate concentration on the inhibitory activity of rNvCI against various A-type carboxypeptidases.

A. [bCPA]=2.0 nM, [rNvCI]=1.2 nM. **B.** [hCPA-1]=4.8 nM, [rNvCI]=1.2 nM. **C.** [hCPA-2]=7.4 nM, [rNvCI]=15.0 nM. **D.** [hCPA-4]=22.6 nM, [rNvCI]=7.0 nM. Substrate: AAFP. Other specifications are described in Materials and Methods. a=v_i/v_0: fraction of enzymatic activity in the presence (v_i) and absence of inhibitor (v_0) in terms of initial velocities. Data are means (n=3) ± S.D. One-way ANOVA found significant differences between groups at p<0.05 for hCPA-2. Tukey HSD test found significant differences between sample means at p <0.05 for hCPA-2. Different letters indicate significant differences at p <0.05.

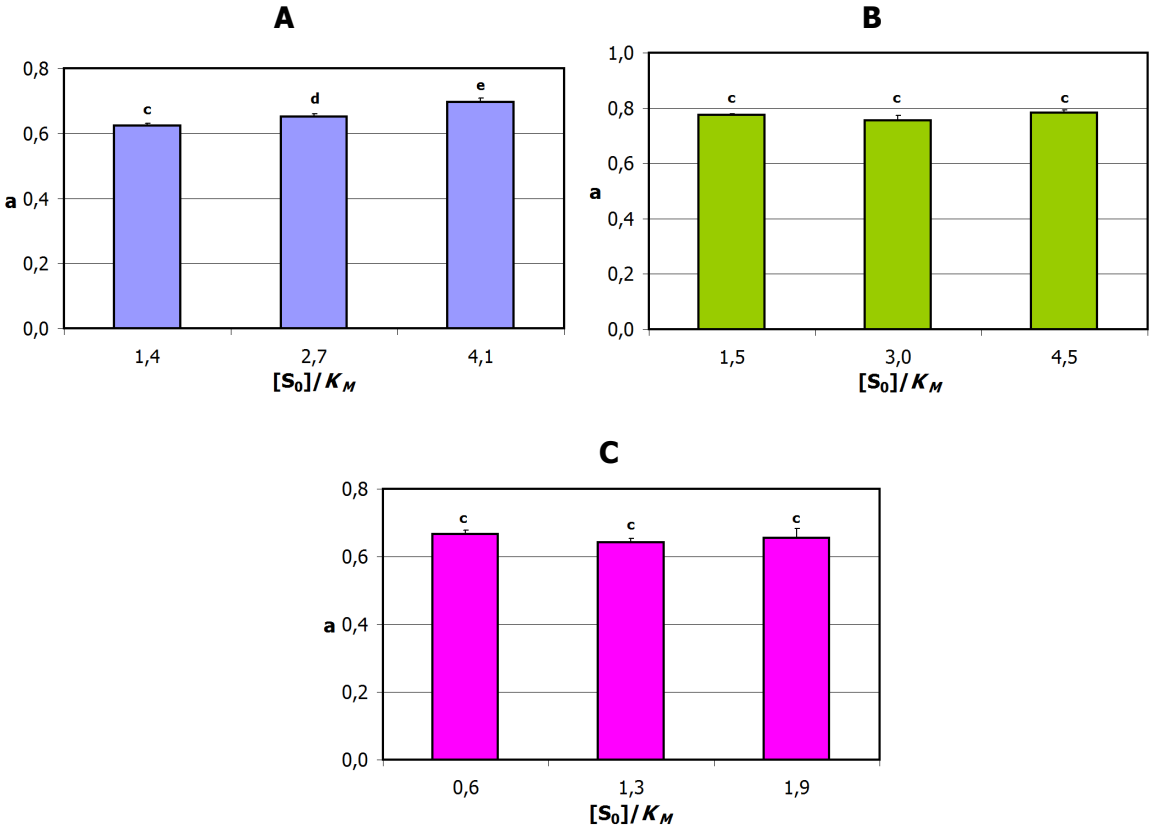
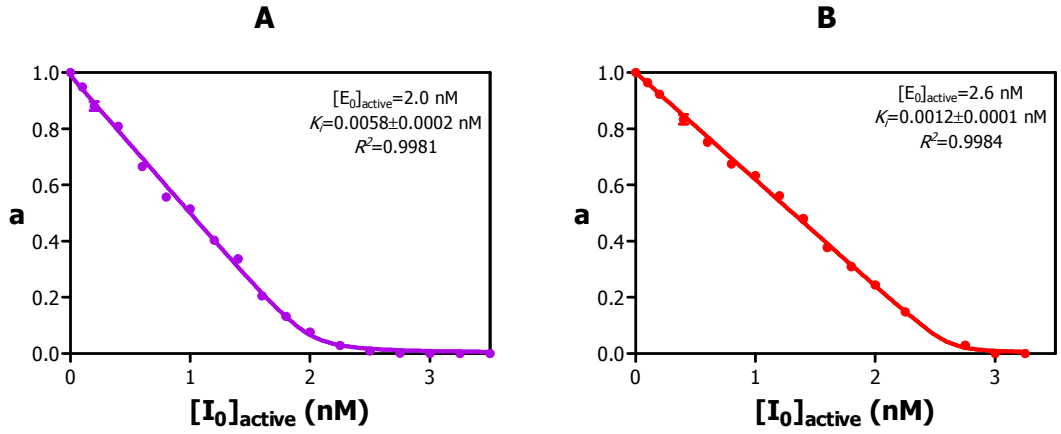


Figure 140. Effect of the substrate concentration on the inhibitory activity of NvCI against various B-type carboxypeptidases.

A. [pCPB]=1.9 nM, [NvCI]=1.0 nM. **B.** [hCPB]=8.9 nM, [NvCI]=1.0 nM. **C.** [bTAFI]=23.0 nM, [NvCI]=2.5 nM. Substrate: AAFA. Other specifications are described in Materials and Methods. a=v_i/v_0: fraction of enzymatic activity in the presence (v_i) and absence of inhibitor (v_0) in terms of initial velocities. Data are means (n=3) ± S.D. One-way ANOVA found significant differences between groups at $p<0.05$. Tukey HSD test found significant differences between sample means at $p<0.05$. Different letters indicate significant differences at $p<0.05$.

The determination of K_i values were performed at preincubation time of 15 min (except for human CPB, bovine and human TAFI), although in some cases NvCI proved to be a fast-tight-binding inhibitor. Lower enzyme concentrations were used in order to ensure equilibrium. Dose-response curves against the different carboxypeptidases as well as real K_i values are shown in figures 141 and 142.



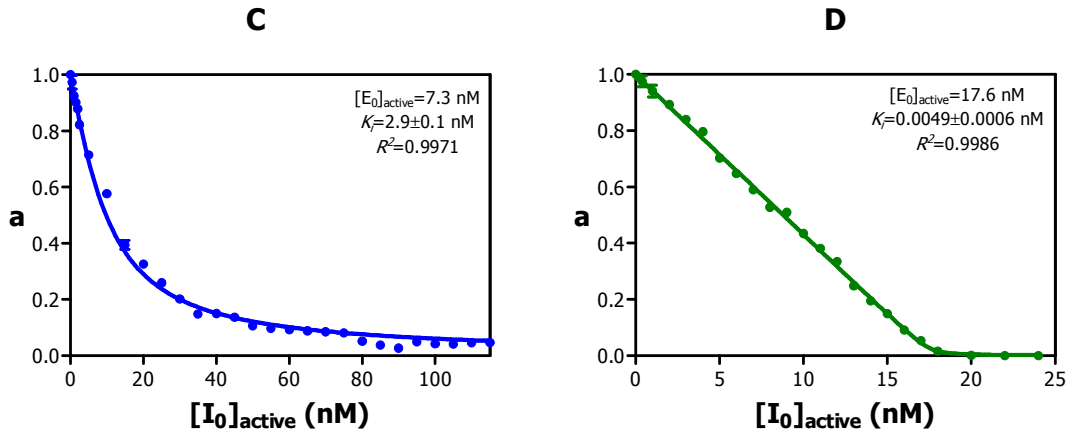


Figure 141. Equilibrium dissociation constant (K_i) of rNvCI against various A-type carboxypeptidases. (A). [bCPA]=2.0 nM. **(B).** [hCPA-1]=4.8 nM. **(C).** [hCPA-2]=7.4 nM. **(D).** [hCPA-4]=22.6 nM. Substrate: AAFP. Preincubation time: 15 min, T=37ºC. Other specifications are described in Materials and Methods. $a=v_i/v_0$: fraction of enzymatic activity in the presence (v_i) and absence of inhibitor (v_0) in terms of initial velocities. The best-fit value of K_i was performed by adjusting the experimental values to the Morrison equation (Eq. 23) using the program GraphPad Prism 5 (GraphPad Software, Inc.) at $p<0.05$. Data are means (n=3) ± S.D.

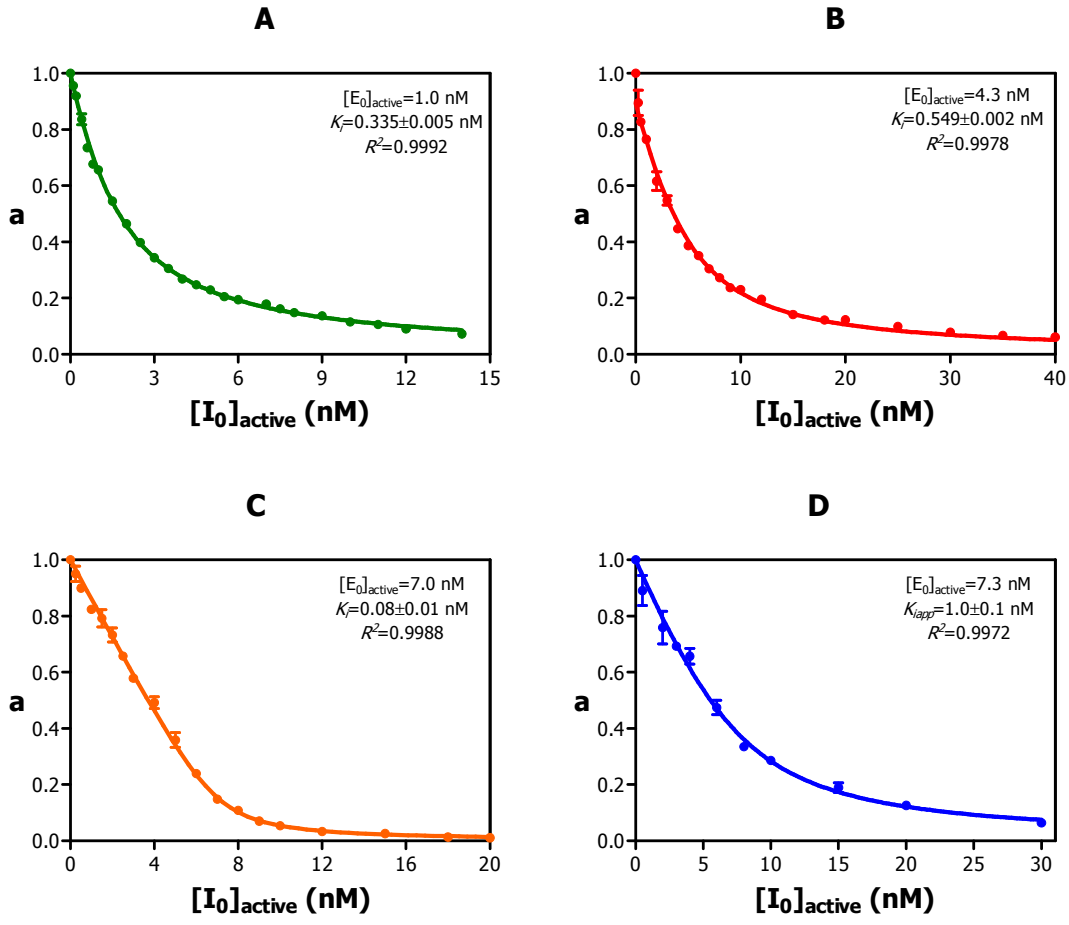


Figure 142. Equilibrium dissociation constant (K_i) of rNvCI against various B-type carboxypeptidases. (A). [pCPB]=1.9 nM. **(B).** [hCPB]=8.9 nM. **(C).** [bTAFI]=23.0 nM. **(D).** [hTAFI]=31.2 nM. Substrate: AAFA. Preincubation time: 15 min, T=37ºC. Other specifications are described in Materials and Methods. $a=v_i/v_0$: fraction of enzymatic activity in the presence (v_i) and absence of inhibitor (v_0) in terms of initial velocities. The best-fit value of K_i was performed by adjusting the experimental values to the Morrison equation (Eq. 23) using the program GraphPad Prism 5 (GraphPad Software, Inc.) at $p<0.05$. Data are means (n=3) ± S.D.

Table 55. Summary of K_i values of natural and recombinant NvCI against various carboxypeptidases

Enzyme	NvCI K_i (pM)	rNvCI K_i (pM)
Bovine CPA	9.6 ± 1.4	5.8 ± 0.2
Human CPA-1	1.8 ± 0.2	1.2 ± 0.1
Human CPA-2	6507.3 ± 68.6	2941.0 ± 132.4
Human CPA-4	9.7 ± 1.0	4.9 ± 0.6
Porcine CPB	649.9 ± 12.2	334.7 ± 4.5
Human CPB	110.9 ± 4.7	549.3 ± 2.4
Bovine TAFI	121.4 ± 20.1	84.1 ± 13.4
Human TAFI	392.9 ± 21.9	950.9 ± 98.8

Table 55 summarizes the K_i values obtained with natural and recombinant NvCI. It is noteworthy the low K_i values obtained against carboxypeptidases such as bCPA, hCPA-1 and hCPA-4, which are in the picomolar range. Other K_i values are in the order of $1x10^{-10}$ M, except for hCPA-2 ($1x10^{-9}$ M). NvCI represents the tightest-binding carboxypeptidase inhibitor from proteinaceous nature currently described against MCPs.

It is important to note that the recombinant inhibitor displays the same kinetic properties of the natural inhibitor, in terms of specificity and K_i values, taking into account the same molar range of these kinetic parameters in both, natural and recombinant NvCI. On the other hand, concave kinetic curves for both inhibitors, as well as their K_i values obtained demonstrate that NvCI is a reversible tight-binding inhibitor and that this marine invertebrate inhibitor represents the strongest MCPs inhibitor described so far.

IV.9. Production monitoring and yield of human ProCarboxypeptidase A4 in *P. pastoris*

Procarboxypeptidase A4 (proCPA4) belongs to the M14A subfamily of carboxypeptidases. The pancreatic members of this subfamily (CPA1, CPA2, and CPB) act in the degradation of dietary proteins in the digestive tract. Other members of this subfamily display a wide spectrum of physiological roles (Tanco *et al.*, 2010). Human PCPA-4 (PhCPA-4) was identified as a gene product, involved in prostate cancer (Huang *et al.* 1999). Imprinting of human CPA-4 (hCPA-4) in adult benign hypertrophic prostate tissue suggests the possibility that mutations or aberrant imprinting in hCPA-4 are related to prostate cancer aggressiveness (Kayashima *et al.*, 2003).

Given the potential importance of this enzyme, ProhCPA-4 was over-expressed and secreted to the extracellular medium using the *Pichia pastoris* heterologous system as described by Pallares *et al.*, 2005. The ProhCPA4/pPIC9 construct transformed into the *Pichia pastoris* KM71 (mut^s phenotype) cells was kindly provided by Msc. S. Tanco (Institut de Biotecnologia, UAB)

Monitoring production of recombinant ProhCPA-4 was carried out by SDS-PAGE as well as by determining the wet cell weight (WCW). As shown in figure 143, from 24 hours of methanol induction a well-defined band appeared at about 50 kDa corresponding to the procarboxypeptidase form. This band is also observed at 48 hours of methanol induction.

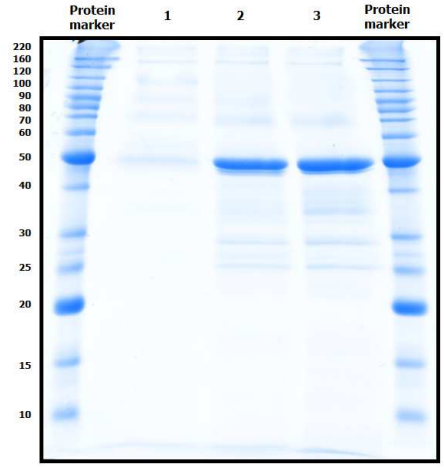


Figure 143. SDS-PAGE of the fermentation supernants in the recombinant production of ProhCPA4. Lane 1: 0 h of induction time, lane 2: 24 h of induction time, lane 3: 48 h of induction time

Monitoring data of ProhCPA-4 production are presented in table 56, where after 48 hours of methanol induction the cell mass increased only 1.4 times with respect to the initial cell mass. This result is attributed to the *P. pastoris* strain used, since phenotype Muts (Methanol utilization slow) displays a reduced ability of cells to metabolize methanol, leading to poor cell growth on medium containing methanol (Lin-Cereghino J. and Lin-Cereghino G.P., 2007). An increase in protein concentration for the extracellular fluid of 1.3 times compared to 0 hours of methanol induction was also determined.

Table Nº 56. Monitoring data of ProhCPA-4 production in *P. pastoris*

Induction time (h)	WCW* (g/L)	Total protein (mg/ml)
0	102.2 ± 0.6	1.3 ± 0.1
24	134.6 ± 0.7	1.7 ± 0.2
48	142.8 ± 0.3	1.6 ± 0.1

*WCW: wet cell weight. Data are means (n=3) ± S.D.

IV.10. Purification of recombinant human CPA4

The enzyme purification was performed using a combination of hydrophobic interaction chromatography (HIC) and a weak anion-exchange chromatography (figure 144). Identification of pro-enzyme form and active carboxypeptidase during the purification process was carried out by SDS-PAGE and by measuring enzymatic activity.

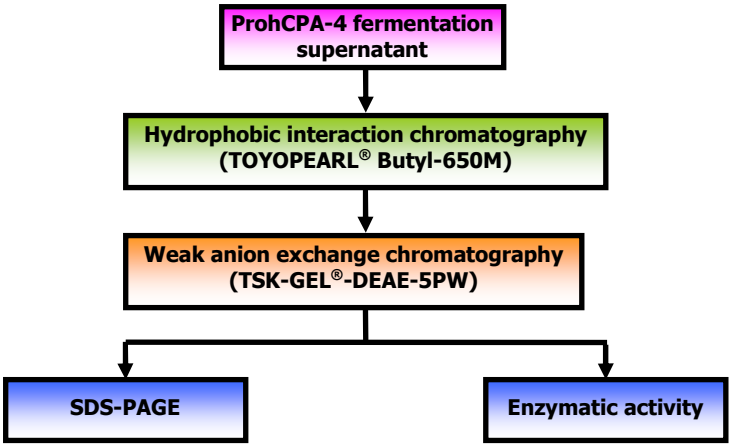


Figure 144. Flowchart process for purification and preliminary characterization of recombinant hCPA4

Figure 145 shows the chromatographic profile obtained from the application of fermentation supernatant to a TOYOPEARL® Butyl-650M column, which was equilibrated at pH 8.0 with 30% ammonium sulfate in order to ensure fixation of protein to the matrix through its superficial hydrophobic groups. Retained proteins were eluted using a linear salt gradient from the same equilibration buffer to 0% ammonium sulfate. It is noteworthy that in the major elution peak a slight enzyme activity was detected, which can be attributed to the presence of a small amount of active enzyme.

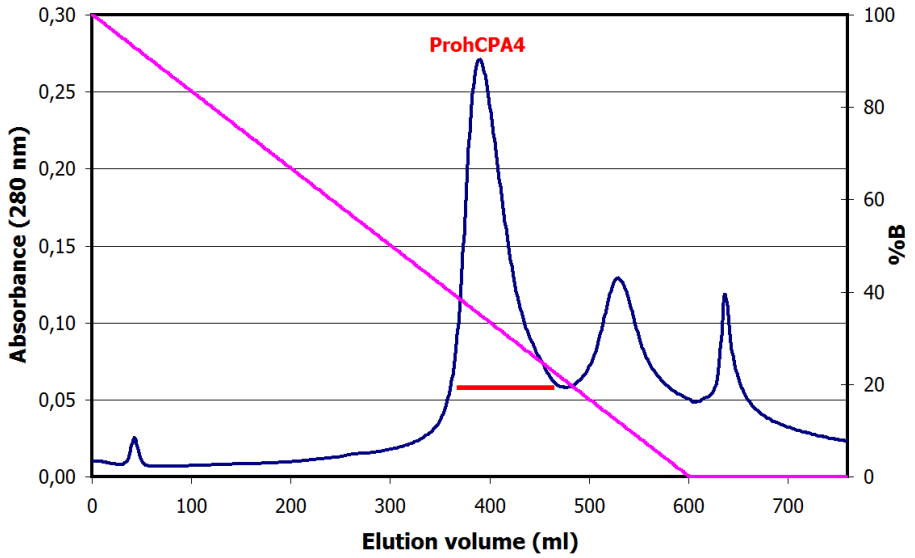


Figure 145. Hydrophobic interaction chromatography of the ProhCPA4 fermentation supernatant on a TOYOPEARL® Butyl-650M column.

Column: 1.6 cm x 20.0 cm. Buffer A: 50 mM Tris-HCl pH 8.0; buffer B: 50 mM Tris-HCl pH 8.0, containing 30% w/v Ammonium sulfate. Column was equilibrated with buffer B over 4 column volumes (CV). Sample loading was performed in buffer B. Molecules non-retained were removed by washing the column with buffer B over 5 CV. Elution was performed using a linear gradient from 100% to 0% B over 20 CV followed by 0% B over 5 CV. The flow rate was 4 ml/min and room temperature. ___Absorbance at 280 nm, ___ProhCPA4 peak, ___ % buffer B

Identification of PhCPA4 in the elution peak was performed by SDS-PAGE, where a more intense band appeared in the central fractions of the peak, corresponding to the molecular mass of zymogen (figure 146).

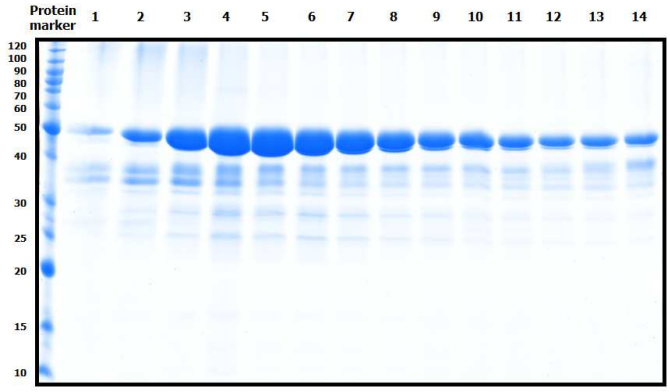


Figure 146. SDS-PAGE of the hydrophobic interaction chromatography of the ProhCPA4 fermentation. Lane 1: elution fraction Nº 1 obtained at elution volume (V_e)=362 ml. Subsequent fractions were collected each 10 ml V_e. Lanes 2 to 14: elution fractions Nº 2 to Nº 14, respectively

The peak containing the Proenzyme form was dialyzed using a cellulose membrane (molecular weight cut-off of 6-8 kDa; Spectrum laboratories Inc., USA), and subsequently subjected to an activation process with trypsin, under the experimental conditions described by Tanco *et al.*, 2010 and detailed in section III.15.1.

The second purification stage of anion-exchange chromatography enabled the final enzyme purification. TSK-GEL® DEAE-5PW column was equilibrated at pH 9, taking into account that the active enzyme has an isoelectric point (pI) of 7.6. The elution was performed using a linear salt gradient by increasing the ionic strength and decreasing pH. Figure 147A displays the chromatographic profile obtained, which indicated the presence of a single peak corresponding to the enzyme and a single band by SDS-PAGE equivalent to its molecular mass (figure 147B).

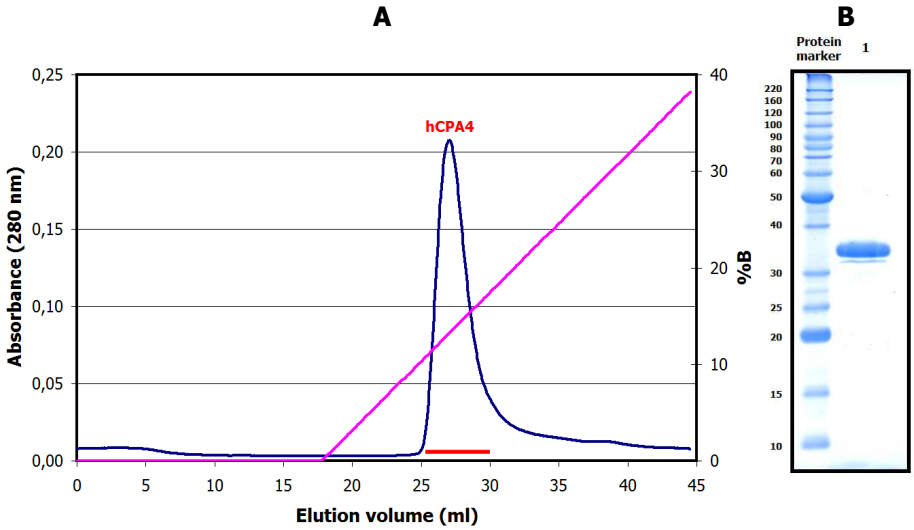


Figure 147. Weak anion exchange chromatography of the elution peak from HIC* on a TSK-GEL® DEAE-5PW column.

(A) Chromatographic profile. Column: 7.5 cm x 7.5 mm. Buffer A: 20 mM Tris-HCl pH 9.0; buffer B: 100 mM Tris-HCl pH 8.6, containing 0.4 M Ammonium acetate. Column was equilibrated with buffer A over 5 column volumes (CV). Sample loading was performed in buffer A. Molecules non-retained were removed by washing the column with buffer A over 5 CV. Elution was performed using a linear gradient from 0% to 100% B over 20 CV followed by 100% B over 10 CV. The flow rate was 1 ml/min and room temperature. —Absorbance at 280 nm, — hCPA4 peak of enzymatic activity, — % buffer B *HIC: hydrophobic interaction chromatography.

(B) SDS-PAGE of purified hCPA4. Lane 1: elution peak of hCPA4 obtained from TSK® DEAE chromatography

The first two stages (fermentation supernatant and hydrophobic interaction chromatography) were not evaluated in terms of yield and purification degree, due to the absence of enzymatic activity in the first case (fermentation supernatant) since the protein is expressed in a pro-enzyme form and detection of slight enzymatic activity in the HIC stage before activation (7.6 units total). Trypsin activation released a total enzymatic activity of 94.9 units and a specific enzymatic activity of 2.8 U/mg, while in the final purification process 68.1 units were recovered; the specific enzymatic activity increased up to 5.6 U/mg, representing a 2-fold purification and 71.8% yield.

Table 57. Summary of the purification procedure of recombinant human CPA4

Step	Total protein (mg)	Enzymatic activity (Ut)	Specific enzymatic activity (U/mg)	Yield (%)	Purification (-fold)
Fermentation supernatant	1054.7 ± 96.1	0	0	-----	-----
Hydrophobic interaction chromatography	32.9 ± 0.6	7.6 ± 1.4	0.23 ± 0.04	-----	-----
Activation with trypsin	33.7 ± 0.6	94.9 ± 7.0	2.8 ± 0.2	100.0 ± 0.0	1.0 ± 0.0
TSK weak anion esxhange	12.3 ± 0.1	68.1 ± 2.6	5.6 ± 0.2	71.8 ± 6.0	2.0 ± 0.2

Data are means (n=3) ± S.D.

IV.11. Analysis of three-dimensional structure of rNvCI in complex with human CPA4

IV.11.1. Formation and purification of rNvCI-hCPA4 complex

Formation and capture of hCPA4 – NvCI complex was performed by preincubation of both proteins for 30 min in 50 mM Tris-HCl pH 8.5 buffer, containing 150 mM NaCl at 37°C. For this purpose, 16 mg of hCPA-4 were incubated with 5.5 mg of rNvCI in a final reaction volume of 70 ml, equivalent to an enzyme: inhibitor ratio of 1:2.

Taking into account that the complex was formed in titration conditions ($[E_0]/K_i \geq 100$) and that titration results of hCPA-4 at pH 8.5 with TCI showed a percentage of active enzyme of 73.4% (data not shown), it was desirable to form the largest possible amount of NvCI-hCPA4 complex by shifting the equilibrium toward the product side. For this reason, a 100% molar excess of rNvCI was used in the complex formation.

The volume of complex solution was reduced to 10 ml by concentration in an Amicon 3000 NMWL. The complex was captured on a molecular exclusion chromatography column (Superdex 75) equilibrated with the same buffer of complex formation. Figure 148 shows the chromatographic profile obtained, which indicated the presence of two minor peaks corresponding to the free enzyme and inhibitor, demonstrated by PAGE using these molecules as markers (figure 149). The major peak obtained in the chromatographic profile corresponded to the hCPA4 – NvCI complex, where through the electrophoresis was evidenced the presence of an intense band with an electrophoretic mobility intermediate between the other two bands.

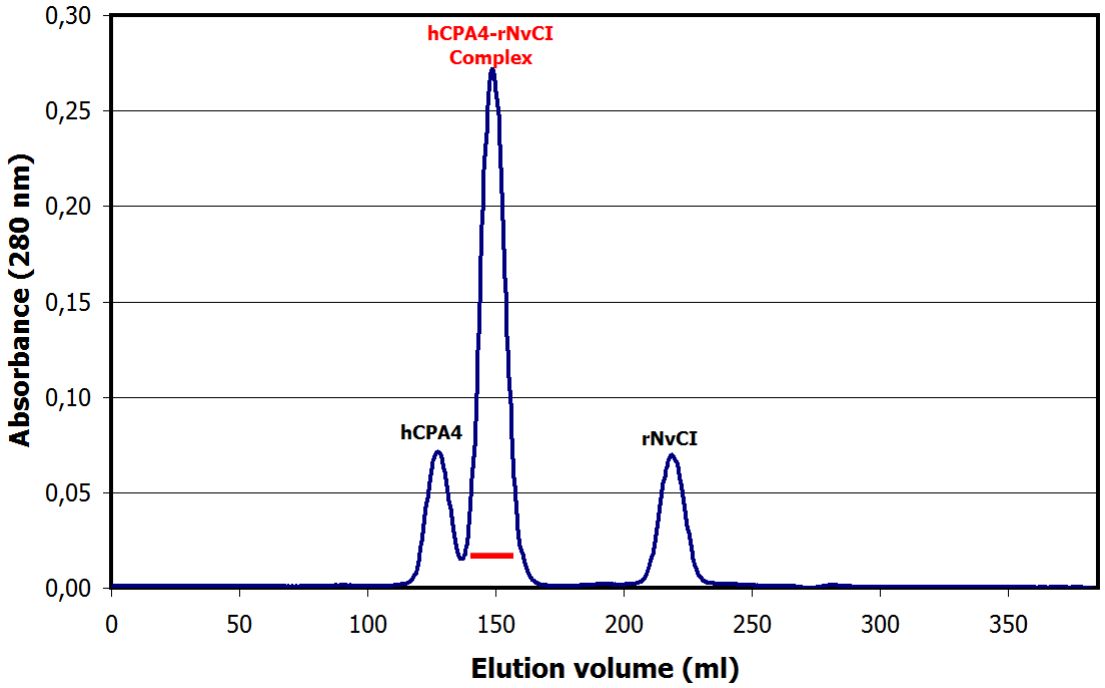


Figure 148. Size exclusion chromatography of hCPA4-rNvCI complex on a HiLoad™ 26/60 Superdex™ 75 prep grad column.

Column: 26 cm x 60 cm. Buffer A: 20 mM Tris-HCl pH 8.5, containing 150 mM NaCl. Column was equilibrated with buffer A over 5 column volumes (CV). Sample loading was performed in buffer A. Elution was performed over 5 CV using buffer A. The flow rate was 2 ml/min and room temperature. —— Absorbance at 280 nm, —— hCPA4-rNvCI complex peak

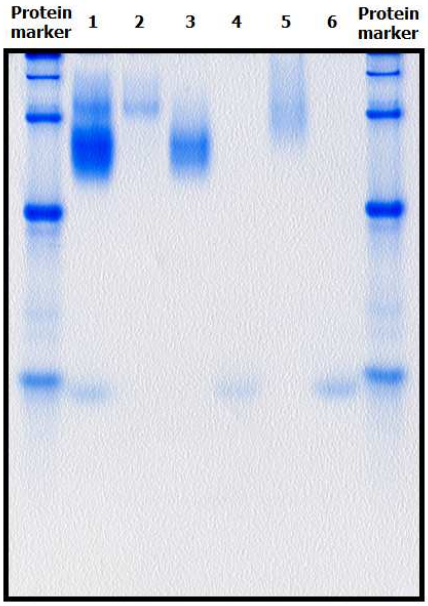


Figure 149. PAGE analysis of size exclusion chromatography of hCPA4-rNvCI complex.
Lane 1: loading sample, lane 2: elution fraction of hCPA-4, lane 3: elution fraction of rNvCI-hCPA-4 complex, lane 4: elution fraction of rNvCI, lane 5: hCPA-4 as a marker, lane 6: rNvCI as a marker

In order to perform crystallization assays, the purified hCPA4–NvCI complex was concentrated to 17.6 mg / ml using an Amicon 3000 NMWLand, and at this time a change of size exclusion chromatography buffer to 5 mM Tris-HCl buffer pH 8.5, containing 50 mM NaCl was carried out. Similarly, hCPA-4 was also concentrated to 11.8 mg/ml and stored in the same final buffer of the complex.

IV.11.2. Three-dimensional structure of NvCI-hCPA4 complex

The polypeptide chain of NvCI in complex with hCPA4 can be clearly and completely traced in the electron density maps (Phe6 to Asn307 for hCPA4, and Val3 to Ala53 for NvCI) (figure 150). The crystal structure of NvCI-hCPA4 displays two complexes in the asymmetric unit. Both complexes are almost identical, showing an rms deviation of 0.8 Å, and displaying similar overlapping for the hCPA4 and the NvCI molecules. Based on the results obtained in gel-filtration chromatography, the biological unit can be considered to be a monomer, formed only by one complex between hCPA4 and NvCI.

The structure of the hCPA4 in complex with NvCI is similar to previous reported structures of hCPA4 alone and in complex with other inhibitors (Pallarès *et al.*, 2005). The structure displays the classical carboxypeptidase fold; with eight α-helices and a mixed eight stranded β-sheet forming a globular α/β motif (figure 150A). The coordination of the zinc atom is conserved when compared with all other metallo-carboxypeptidase structures. In the absence of inhibitor the zinc atom is coordinated to the "catalytic" water molecule and to carboxypeptidase residues His69, His196 and in a bidentate form to Glu72. In the NvCI-hCPA4 complex the catalytic water is not present and it is substituted by a bidentate coordination to the zinc atom by the C-terminal carboxylate group of NvCI, which is buried in the active site groove of the enzyme (figure 150B). As observed in other carboxypeptidase complexes with inhibitors, the most dramatic change observed in the active site residues is the movement of the side chain of Tyr248, almost 180° from the "up" to the "down" position in the complex. There are other minor movements in the active site residues of hCPA4 to accommodate the C-terminal tail of the inhibitor.

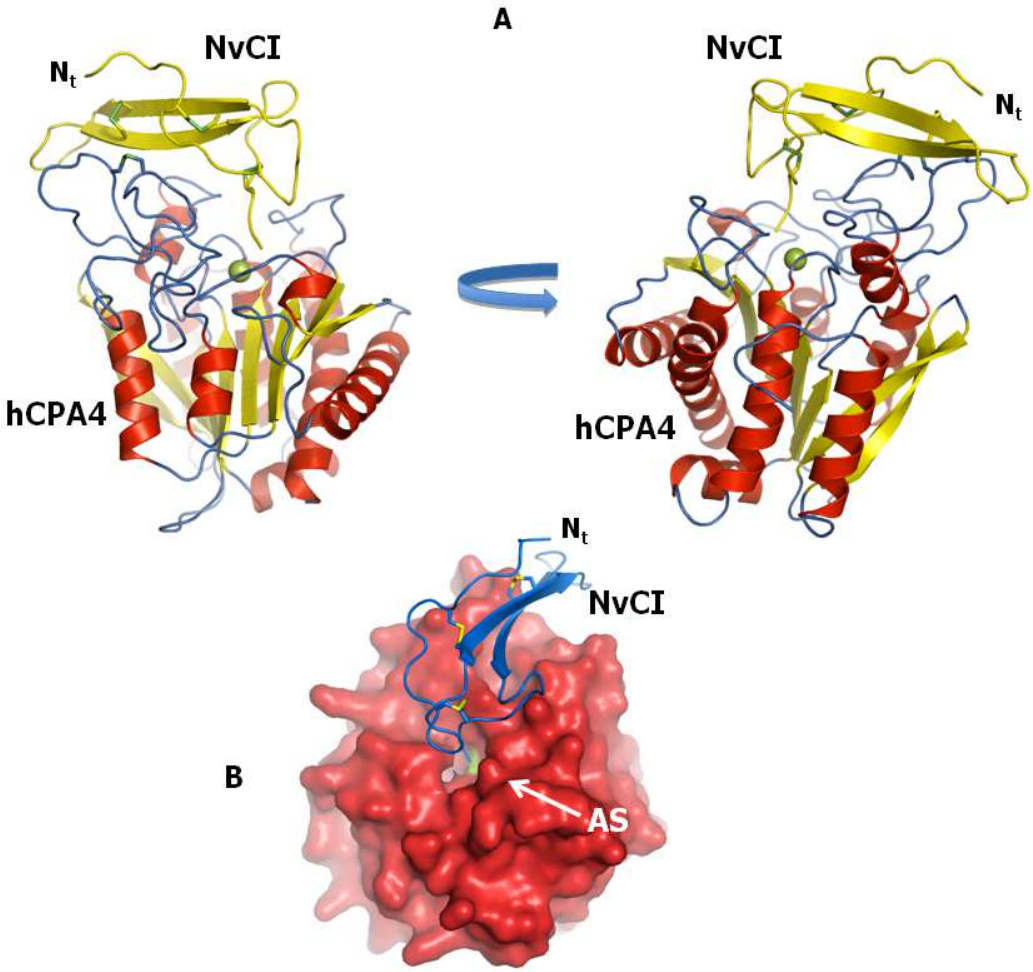


Figure 150. Three-dimensional structure of NvCI in complex with human hCPA4.
(A). Two views of the complex NvCI-hCPA4 shown in ribbon representation. The α-helix, β-strands and coils are highlighted in red, yellow and blue color, respectively. The catalytic zinc atom in the metallo-carboxypeptidase hCPA4 is shown in green. The three disulfide bridges formed in NvCI are shown in stick representation and visualized in green. The N-termini is labeled N_t. **(B)**. Surface representation of hCPA4 in complex with NvCI shown in ribbon representation (blue). AS indicates the position of the active site groove of hCPA4. All figures were prepared with PyMOL (Delano, 2002).

The structure of the NvCI inhibitor displays a new extended small protein-folding motif, which is basically formed by a central anti-parallel two-stranded β-sheet connected by three major loops (figure 151). The β-strands and the three loops are cross connected and stabilized by three disulfide bridges, formed between Cys9 and Cys23, Cys15 and Cys51 and Cys27 and Cys38. In spite of its small size, NvCI has a small hydrophobic core located next to the C-terminal end of the protein formed by non-polar interactions of the side chain of Trp42, which is sandwiched between two disulfide bridges of the inhibitor (figure 151B).

The extended structure of the inhibitor also contains a few bulky exposed hydrophobic residues to the solvent, Phe25, Phe34, and Phe44, which reduces the solubility of the recombinant protein. The C-terminal tail of NvCI is only formed by two residues, Tyr52 and Ala53, but as discussed later in the text, this short extension is sufficient to interact with the active site residues and zinc atom of the carboxypeptidase.

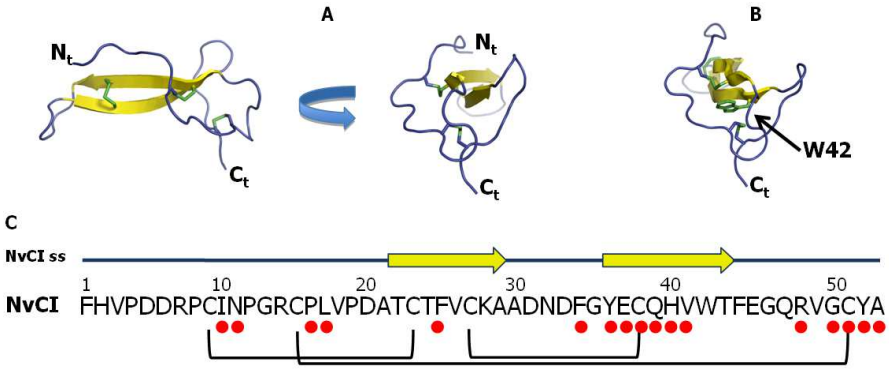


Figure 151. Analysis of secondary structure and disulfide pairing of NvCI.
(A). Two views of the ribbon representation of NvCI. The β-strands and coils are colored in yellow and blue, respectively. The three disulfide bridges formed in NvCI are shown in stick representation (green). The N- and C-termini are labeled N_t or C_t, respectively. **(B)**. Same representation as panel A but depicting in stick representation Trp42. **(C)**. Primary protein sequence of NvCI. Cysteine residues involved into the disulfide bridges are connected by a straight line. Amino acid residues of NvCI interacting with the carboxypeptidase at a distance less than 4 Å are represented by red circles. ss corresponds to the secondary structure elements.

IV.11.3. Primary interaction site in NvCI-hCPA4 complex

Although the small size of NvCI, it interacts extensively with hCPA4, with a total contact area between both proteins of 1875.1 $Å^2$. As observed with the other exogenous carboxypeptidase inhibitors (PCI, LCI, ACI and TCI), inhibition of the enzyme is due to a competitive interaction with the active site of the carboxypeptidase, by occlusion of the active site subsites S1', S1, S2 and S3 (figure 152A). These sites are occupied by the C-terminal tail of NvCI, Tyr52 and Ala53, and constitute the "primary" contact region of the inhibitor. The "secondary" contact region, which is much extended and covers almost one face of the inhibitor, will be discussed later.

The C-terminal tail of NvCI is shorter in comparison to the other known carboxypeptidase inhibitors; it is only composed by two residues, Tyr52 and Ala53, corresponding to P2 and P1 positions, in that order. Three of the four reported exogenous carboxypeptidase inhibitors contain an extra C-terminal residue (position P1') that is cleaved off by the carboxypeptidase and remains trapped in the S1' subsite (glutamic acid and glycine are the P1' for LCI and PCI, respectively) (Reverter *et* al., 2000; Arolas *et al.*, 2005a). NvCI inhibitor was initially isolated from the natural source *N. versicolor* and sequenced by a combination of Edman degradation and *de novo* peptide sequencing by MALDI MS/MS, thus it was unknown whether an extra C-terminal residue is present or not at the C-terminal tail. However, the cDNA sequencing isolated from the body of *N. versicolor* allowed to confirm that the amino acid sequence of NvCI obtained by chemical and proteomic methods is correct and therefore, that this MCPs inhibitor lacks this residue as in the case of tomato MCPs inhibitor (figure 107).

The electron density maps of the complex NvCI-hCPA4 shows that part of the S1' subsite of the carboxypeptidase is occupied by a nitrate molecule from the crystallographic buffer (figure 152A). The nitrate molecule is in contact with the guanidinium group of Arg145 (2.88 and 2.97 Å distance to NH1 and NH2 atoms, respectively) and Asn144 (2.79 Å distance to ND2 atom). These residues belong to the S1' subsite and interact with the carboxylate group of the cleaved residue of a carboxypeptidase substrate. In the PCI and LCI complex structures, the cleaved C-terminal residue occupies a similar position as the nitrate molecule in NvCI.

Sequential and structural comparisons of the C-terminal tails of the different exogenous carboxypeptidase inhibitors indicate an identical conformation of the backbone and side chains for the P1 and P2 residues (Ala53 and Tyr52 for NvCI, respectively) (Figures 152B and 152C). P1 and P2 residues are oriented in a substrate-like manner in all the reported structures of these competitive tight-binding inhibitors. Notably in all of them the chemical character of the residue forming the P1 subsite is aliphatic, whereas for the P2 subsite the preference is for aromatic residues (except a Leu residue for ACI), which perform stacking interactions with the aromatic ring of Tyr248 (figure 152B).

As mentioned before, the C-terminal carboxylate group of Ala53 coordinates the zinc atom in a bidentate form (2.28 Å and 2.44 Å, respectively), whereas the amino group of Ala53 forms a hydrogen bond with the hydroxyl side-chain oxygen of Tyr248 (2.78 Å), which is disposed in the "close" conformation (when bound to substrates). The S1 subsite is also composed by Glu270 and Arg127, both of which participate in the polarization of the carbonyl group of the scissile peptide bond and to proton exchange. In addition to coordinate the zinc atom, each of the carboxylate C-terminal oxygens of Ala53 is at hydrogen bond distance with Glu270 and Arg127 (2.54 Å and 2.79 Å, respectively).

The S2 subsite, which is believed to contribute to the correct orientation of substrates during the catalytic reaction, is mainly formed by backbone hydrogen bonds between the amino and carboxyl group of Tyr52 with the side chains of Glu163 and Arg71 (3.08 and 3.02 Å, respectively) (figure 152A). The side chain of Tyr52 is placed in a hydrophobic pocket created by Tyr248 and Val164 from hCPA4, and internally with Pro16 from the NvCI inhibitor.

Interestingly, the side-chain of Glu163 of hCPA4 is forming a favourable hydrogen bond contacts with either the amino groups of Cys51 and Tyr52 (3.08 and 3.06 Å, respectively). The Glu163 interaction is novel and was not observed in the other structures of complexes with carboxypeptidase inhibitors, basically due to a different backbone orientation of the P3 residue (Cys51 in NvCI). So, it seems that in NvCI the hydrogen bond net created by the C-terminal tail in complex with the active site of hCPA4 is more extended and contain two extra bonds. As mentioned before, NvCI displays the strongest inhibitory constant (in the picomolar range) against all CPA-type forms except hCPA2, where the K_i value is 2.9 nM (table 55).

Interestingly, in hCPA2 a major difference in the contact residues of the "primary" interaction site is the substitution of Glu163 by an aspartic acid, which presumably is too distant to form any hydrogen bond (figure 155). This fact suggests a reason for the nanomolar value of K_i displayed by NvCI against hCPA2 (similar to the other protein carboxypeptidase inhibitors, PCI, LCI, ACI and TCI) instead of the picomolar range K_i displayed in all other the A-type carboxypeptidases with a glutamic acid in that position (table 55).

In summary, most of the active site residues in hCPA4 involved in substrate binding and catalysis are in contact with NvCI. The interaction with the enzyme and the conformation of the C-terminal tail, in NvCI only composed by Tyr52 and Ala53 (P2 and P1 positions, respectively), is similar to the other known exogenous carboxypeptidase inhibitors (LCI, PCI, ACI & TCI), in which the inhibitor tail mimics the substrate binding (figures 152B and 152C).

All these structures of exogenous carboxypeptidase inhibitors from evolutionary distant organisms are unrelated and completely different. However, the only conserved motif in all of them is the structural conformation of the P1 and P2 residues of the C-terminal tail, which can be considered as a general mechanism for carboxypeptidase inhibition and a clear example of convergent evolution.

As mentioned before, the only structural difference in the NvCI-hCPA4 complex is the main-chain conformation of the P3 residue in comparison to the other known CP inhibitors, which favours the formation of two extra hydrogen bonds with Glu163 and presumably induces a reduction of the inhibitory constant by stabilization of the product formation (figure 155).

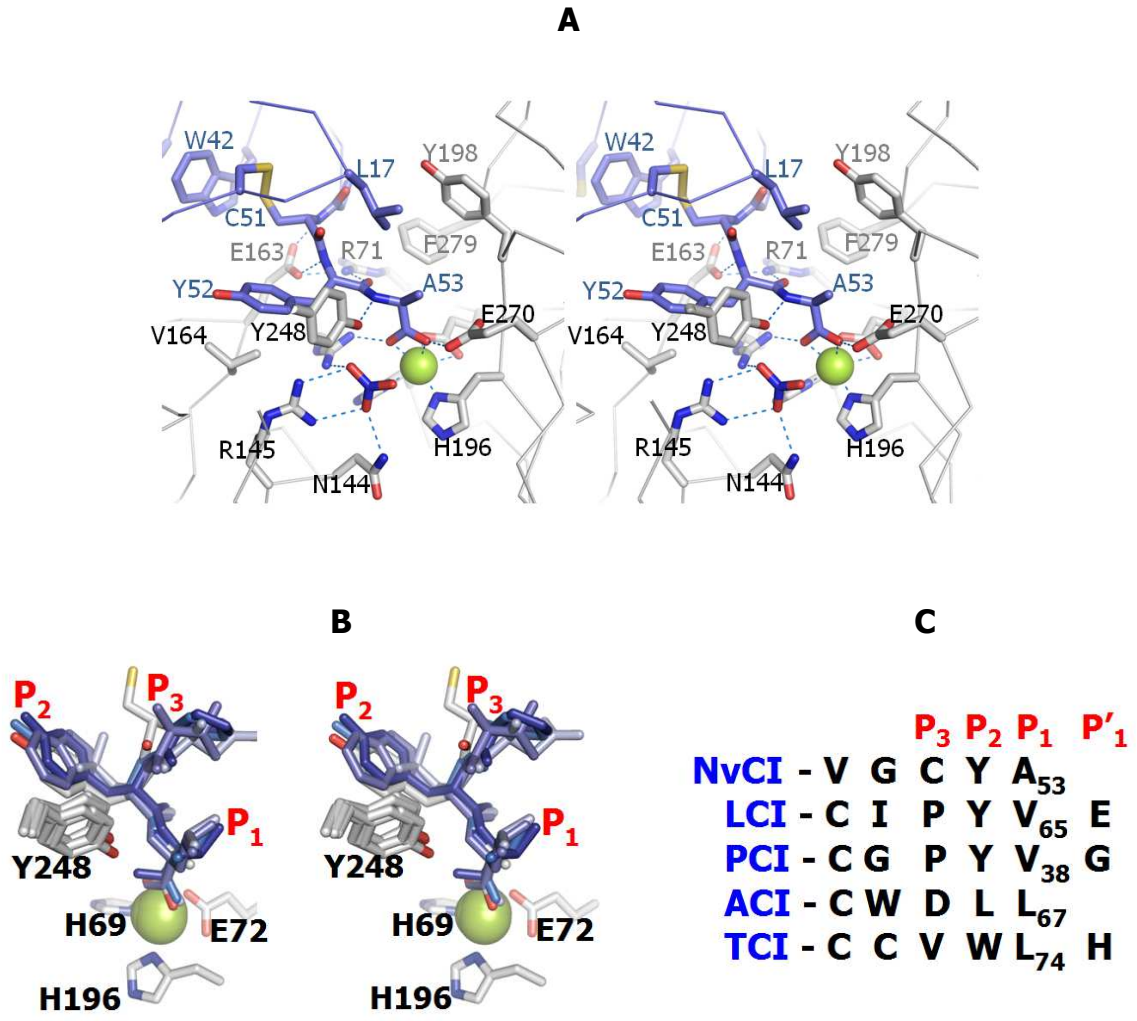


Figure 152. Close-up view of the NvCI "primary" binding region in complex with the active site of hCPA4. (A). Close-up stereo view in stick representation of the C-terminal tail of NvCI with the active site of hCPA4. Amino acid residues corresponding to the C-terminus of NvCI and the active site of hCPA4 are labeled and shown in blue and grey color, respectively. The active-site zinc is visualized as a green sphere. Nitrate molecule is also shown in stick representation (blue). **(B)**. Stereo representation of the structural alignment of the C-terminal tail of the inhibitor from panel B. Amino acid residues corresponding to the C-terminus of the inhibitors are shown in blue. Zinc-binding residues and Tyr248 are shown grey color. The active-site zinc is shown as a green sphere. **(C)**. Sequence alignment of the C-terminal tails with other exogenous proteinaceous carboxypeptidase inhibitors, depicting the P1, P2 and P3 subsite.

IV.11.4. Secondary interaction site in NvCI-hCPA4 complex

The "secondary" interaction site of NvCI is mainly composed by contacts between residues from the two-stranded β-sheet with regions distant from the active site groove of hCPA4 (figure 153). As demonstrated for other carboxypeptidase inhibitors the "secondary" interaction site contributes substantially to the decrease of the inhibitory constant, which in the case of NvCI against hCPA4 is in the picomolar range (table 55).

This strong interaction is reflected in the structure of the complex by a large number of contacts between the inhibitor and the carboxypeptidase domain. Besides several van der Waals interactions, there are major specific polar contacts between backbone and side chains from NvCI with hCPA4. Of special interest are the specific hydrogen bond interactions between the side-chains of Gln39 with Asn123 (2.99 Å), Glu37 with Arg130 (2.86 and 2.86 Å between O1 and O2 with NH1 and NH2 atoms, respectively) and between Arg7 with Asn159 forming two hydrogen bonds (2.77 Å and 2.79 Å, respectively) with a nitrate molecule of the crystallization buffer (figures 153A and 153B).

Four hydrogen bonds involving main-chain atoms are also formed: between carbonyl oxygens of Ile10 and Asp11 with the side-chain of Asn159 (3.30 and 3.71 Å, respectively), between the amide hydrogen of Cys38 with the carbonyl group of Ser137 (3.19 A), and between the amide hydrogen of His40 with the carbonyl group of Cys162 (3.43 Å). A small hydrophobic core can also be distinguished in the "secondary" interface of the NvCI with hCPA4, basically formed by Phe25 of the inhibitor that is buried in a pocket created by two disulfide bridges, Cys27 and Cys38 of NvCI and Cys138 and Cys161 of hCPA4.

Structural comparisons indicate that the conformation of the residues forming the "secondary" interaction region in hCPA4 is highly conserved in all members of type-A and type-B carboxypeptidases (M14A subfamily). In all of them a similar complex interface with NvCI would be expected, in particular considering the conservation and similar orientation of the specific side-chain contacts described before, such as for Glu37 and Gln39. Thus the lower K_i values observed for NvCI in comparison to the other inhibitors can be attributed to both, the "primary" and "secondary" interaction regions, which creates an extended interface with the carboxypeptidase enzyme that minimizes the product release of the catalytic reaction.

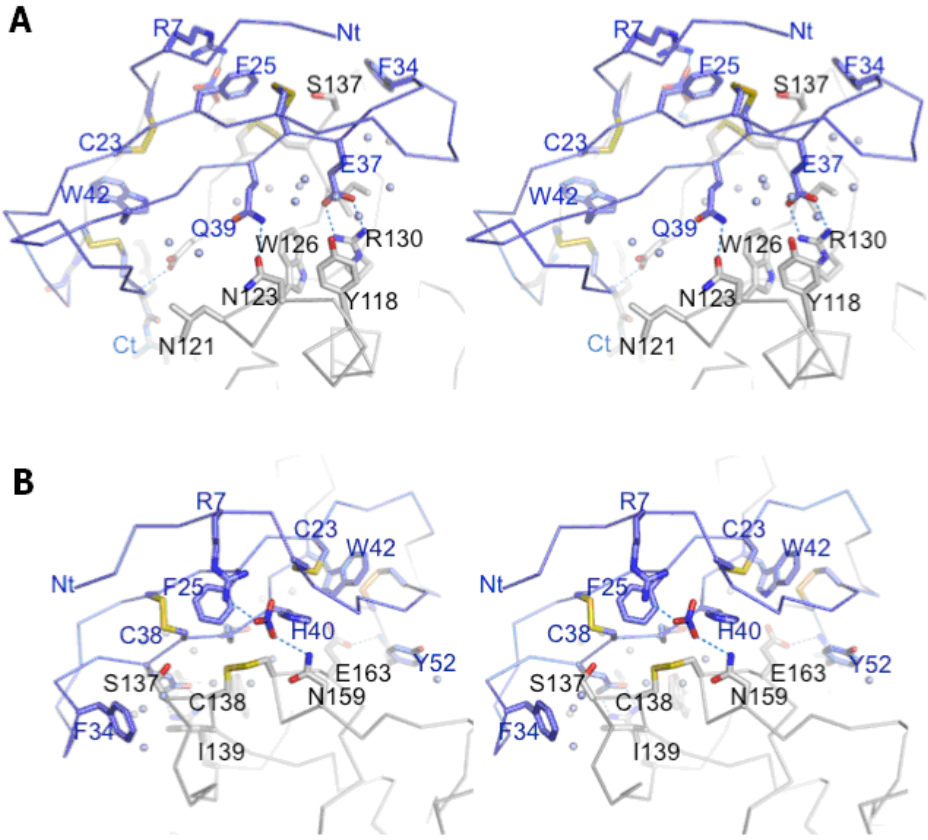


Figure 153. Close-up view of the NvCI "secondary" binding region in complex with hCPA4. **(A).** Close-up stereo view in stick representation of the "secondary binding" region of NvCI with the surface of hCPA4. Amino acid residues corresponding to NvCI and hCPA4 are labeled and shown in blue and grey color, respectively. Nitrate molecule is also shown in stick representation (blue). Water molecules are shown as grey spheres. The N- and C-termini are labeled N_t or C_t, respectively. **(B).** Idem as panel A but with a rotation of the complex of aprox. 180°, depicting the opposite face of the complex.

IV.11.5. Structural comparison of several complexes between exogenous proteinaceous carboxypeptidase inhibitors

All exogenous carboxypeptidase inhibitors possess a similar substrate-like manner mechanism of inhibition, in which the C-terminal tail mimics the interaction of a substrate with the active site of the enzyme. They all behave as tight-binding competitive inhibitors with a very low inhibition constant in all of them, displaying K_i values below the nanomolar range. In addition to this "primary" and more important C-terminal interaction, they all posses a second interface, distant from active site, which covers different regions of the carboxypeptidase and confers stability to the complex formation (figure 154).

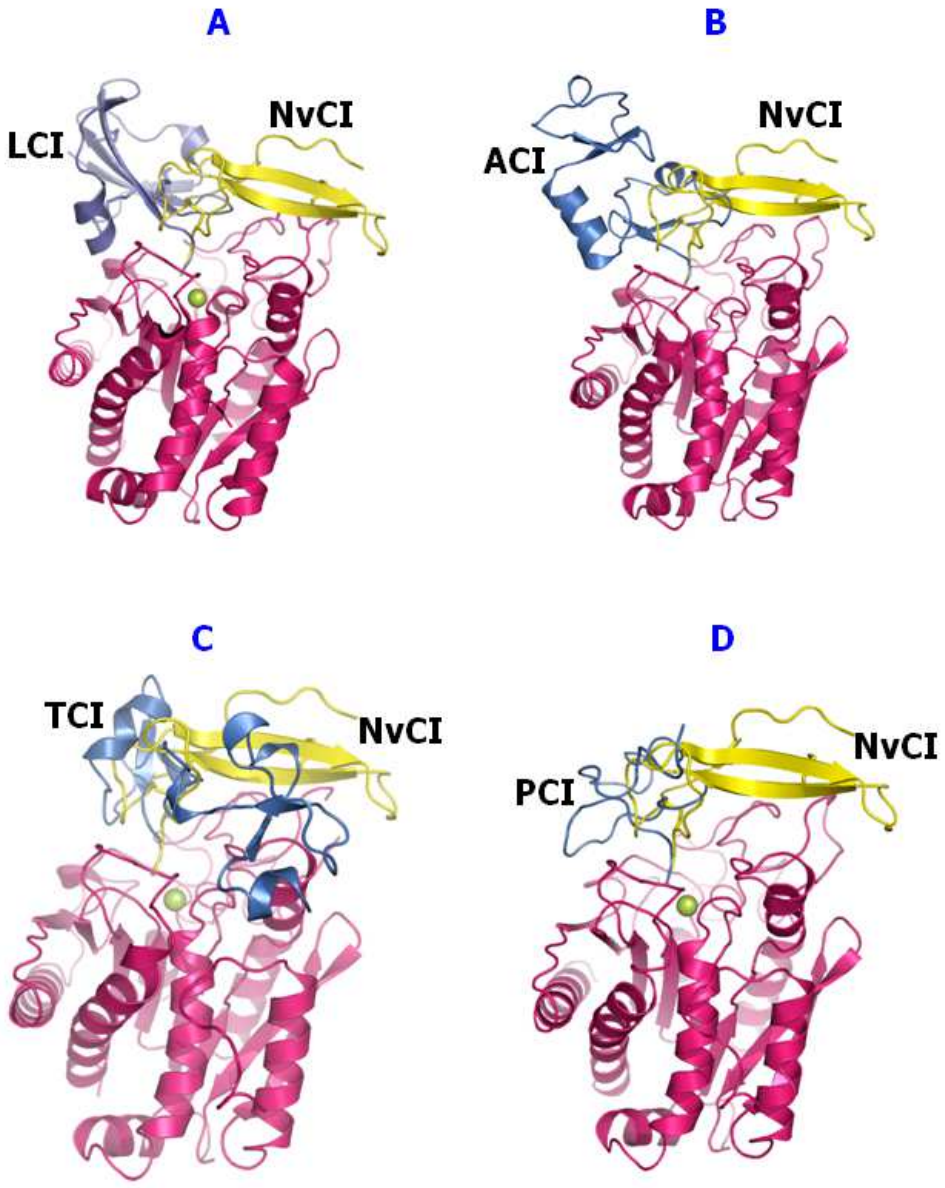


Figure 154. Three-dimensional structure comparison of several complexes between exogenous proteinaceous carboxypeptidase inhibitors. Structures of different carboxypeptidase were overlapped with the NvCI-CPA4 complex. **(A)**. LCI, NvCI and human CPA2 are painted in grey, yellow and magenta, respectively (from PDB code 1DTD). **(B)**. ACI, NvCI and human CPA1 are colored in blue, yellow and magenta, respectively (from PDB code 3FJU). **(C)**. TCI, NvCI and human CPB are represented in blue, yellow and magenta, respectively (from PDB code 1ZLH). **(D)**. PCI, NvCI and bovine CPA1 are shown in blue, yellow and magenta, respectively (from PDB code 4CPA). The active-site zinc is presented as a green sphere.

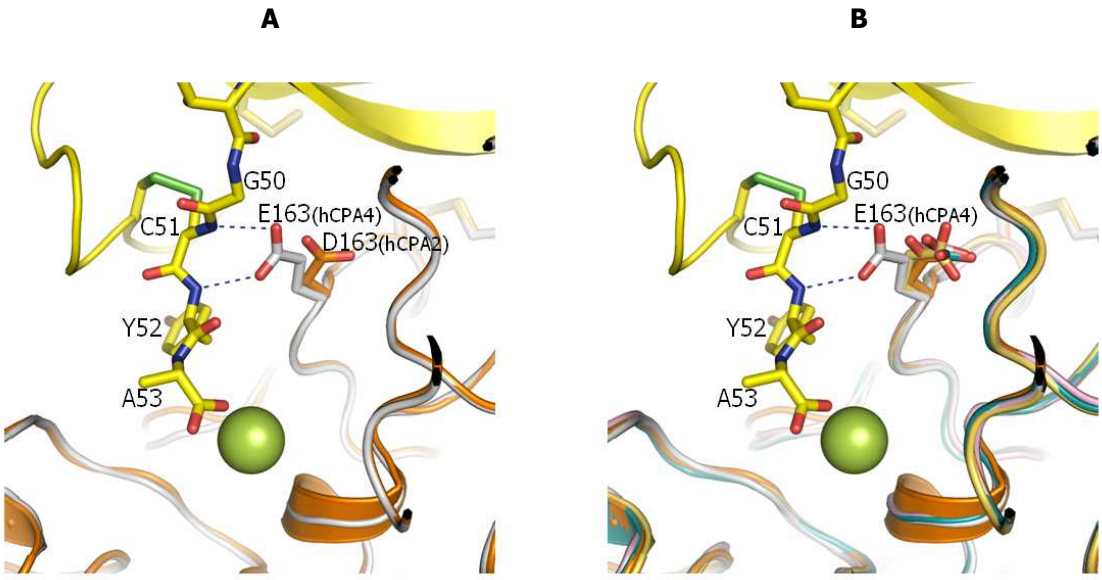


Figure 155. Detailed view of the interaction of Glu163 (hCPA4) with the C-terminal tail of NvCI. (A). Stick representation of the hydrogen bonds between Glu163 and the C-terminal tail of NvCI. hCPA4 structure was overlapped with hCPA2 (PDB code 1DTD). NvCI, hCPA4 and hCPA2 chains are presented in yellow, grey and orange, respectively. The catalytic zinc atom is presented as a green sphere. Selected amino acid residues for both carboxypeptidases and inhibitor are labeled. **(B).** Same representation as panel A but with the CPA4 structure overlapped with hCPA2, hCPA1 (PDB code 3FJU), hCPB (PDB code 1ZLH) and bCPA1 (PDB code 4CPA).

Table 58. Crystallographic data

Data collection	
Space group	$P2_1$
Cell dimensions	
a, b, c (Å)	69.22, 71.98, 79.84
α, β, γ (°)	90.00, 108.84, 90.00
Resolution (Å)	50 - 1.70 (1.79 - 1.70)
Rmerge†	0.061 (0.51)
I⁄σI	13.1 (1.5)
Completeness (%)	97.5 (84.5)
Redundancy	3.3 (2.3)
Refinement	
Resolution (Å)	50 - 1.70
No. reflections	79607
Rwork⁄Rfree ‡	20.54 / 23.46
No. atoms	5903
Protein	702
Water	306
Nitrate	6
Zinc atom	2
Average B-factors	
Protein	
Water	
R.m.s deviations	
Bond lengths (Å)	0.007
Bond angles (°)	1.033

*Statistic for highest resolution shell is shown in parentheses.
†Rmerge 1/4 Σ jIi − < I > j⁄ Σ Ii , where Ii is the ith measurement of the intensity of an individual reflection or itssymmetry-equivalent reflections and <I > is the average intensity of that reflection and its symmetry-equivalent reflections.
‡Rwork 1/4 Σ jjFobsj − jFcalcjj⁄ Σ jFobsj for all reflections and Rfree 1/4 Σ jjFobsj − jFcalcjj⁄ Σ jFobsj, calculated on the 5% of data excluded from refinement.

Table 59. Interactions between primary and secondary regions of NvCI and hCPA4

NvCI	hCPA4	Distance (Å)
Primary interaction region		
Cys-51 N	Glu-163 Oε1	3.06
Tyr-52 N	Glu-163 Oε2	3.08
Tyr-52 Cδ2	Val-164 Cγ1	3.59
Tyr-52 Cδ2	Tyr-248 Cζ	3.59
Tyr-52 O	Arg-71 Nη2	3.02
Tyr-52 O	Arg-127 Nη2	3.23
Tyr-52 O	Phe-279 Cζ	3.20
Ala-53 N	Tyr-248 Oη	3.78
Ala-53 Cβ	Glu-270 Oε2	3.26
Ala-53 O	Glu-270 Oε2	2.54
Ala-53 Oτ	Arg-127 Nη2	2.79
Secondary interaction region		
Ile-10 O	Asn-159 Nδ2	3.30
Asp-11 O	Asn-159 Nδ2	3.71
Pro-16 Cβ	Tyr-249 Cδ2	3.63
Leu-17 Cδ1	Tyr-198 Cε1	3.91
Leu-17 Cδ1	Phe-279 Cζ	3.62
Phe-34 Cε2	Ser-136 O	3.43
Phe-34 Cδ1	Ser-137 Cβ	3.67
Tyr-36 Cε1	Ser-137 Cβ	3.64
Glu-37 Oε1	Arg-130 Nη2	2.81
Glu-37 Oε2	Arg-130 Nη1	2.85
Cys-38 N	Ser-137 O	3.19
Cys-38 O	Cys-138 Cα	3.24
Cys-38 Sγ	Ser-137 Oγ	3.05
Gln-39 Nε2	Asn-123 Oδ1	2.99
Gln-39 Nε2	Trp-126 Cζ3	3.58
His-40 N	Cys-161 O	3.43
His-40 Cε1	Asn-159 Nδ2	3.71
Arg-48	Leu-125 Cδ1	3.86
Gly-50α	Arg-71 Nη1	3.67
Gly-50α	Phe-279 Cε2	3.80

IV.12. Enzymatic proteolysis of rNvCI

The inactivation by proteolysis is one of the most important mechanisms of irreversible inactivation of proteins. Resistance or susceptibility to hydrolysis by proteases is determined by the amino acid sequence of the protein, its folding as well as the protease specificity. Therefore, protein behaviour against proteases contributes to define its stability and structure-stability relationships.

On the other hand, knowledge of susceptibility to hydrolysis by proteases is an important issue for many applications, such as its recombinant expression in systems containing high-proteases levels, purification and biomedical applications. The latter requires knowledge of the protein half-life, its metabolism, etc.

NvCI is the most potent protease inhibitor from proteinaceous nature described against carboxypeptidades, which gives the molecule potential applications in different diseases where carboxypeptidades are involved: hTAFI (cardiovascular diseases), hCPA-4 (prostate cancer), Nna1 (malaria), among others.

In order to study the effect of proteolysis caused by various enzymes on NvCI, this molecule was incubated at different times with proteases of different specificities and the course of proteolysis was monitored by MALDI-TOF MS.

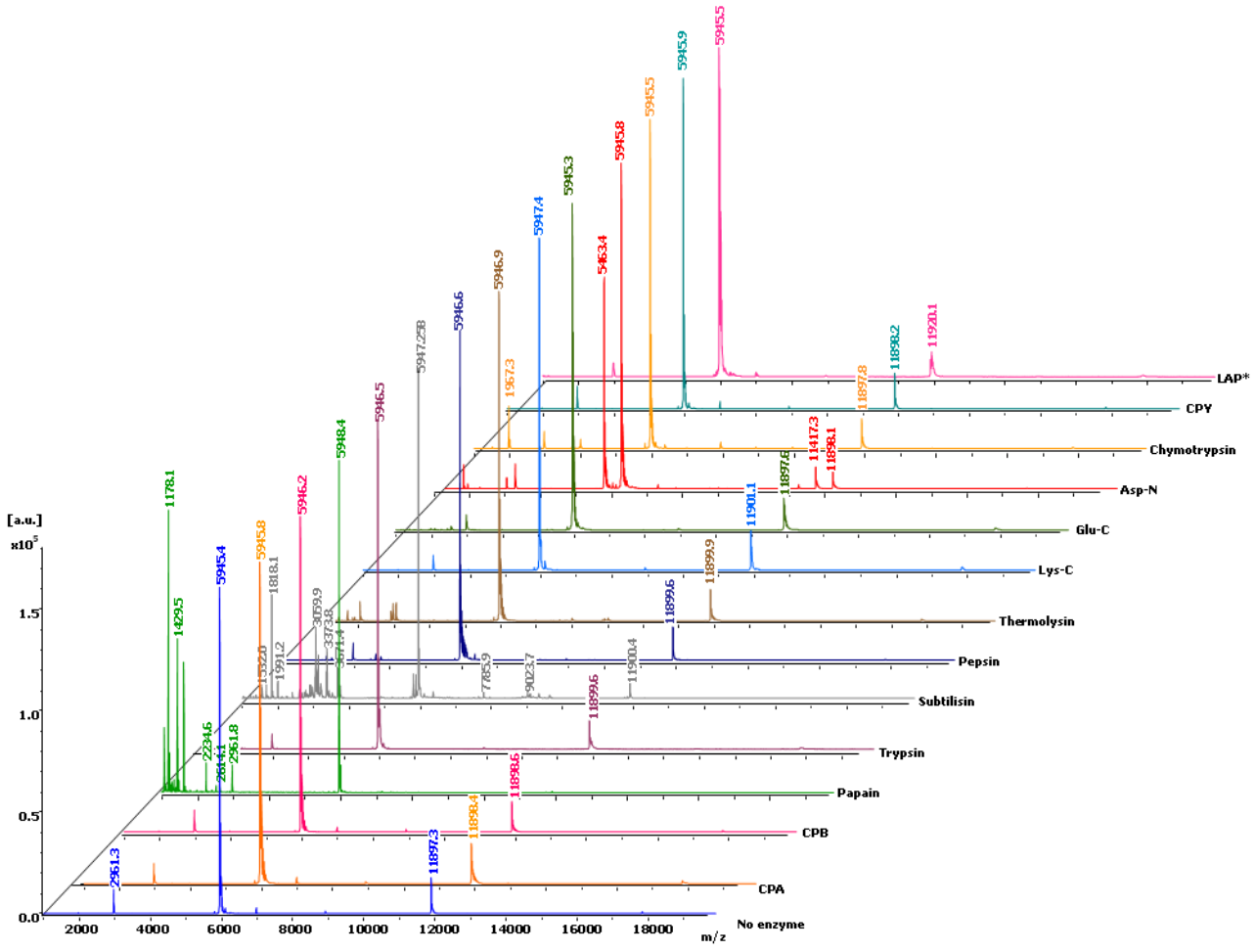


Figure 156. Enzymatic proteolysis of rNvCI at t=1 h.
Reaction conditions: Ratio (w:w) inhibitor:enzyme 10:1**,** T=37 ºC**.** *LAP: Leucyl aminopeptidase

The results obtained after 1 hour of incubation are shown in figure 156. Generally, few changes were observed, except for papain, subtilisin and Asp-N where some hydrolysis products were observed. In the case of Asp-N, a second intense peak close to NvCI was generated.

After 6 hours of enzymatic digestion, an increase in proteolysis with the enzymes mentioned above was displayed and hydrolysis products with trypsin, thermolysin and chymotrypsin were detected at this time (figure 157).

After 18 hours of digestion, proteolysis was enhanced by the enzymes previously mentioned and peptides from hydrolysis with pepsin appeared (figure 158).

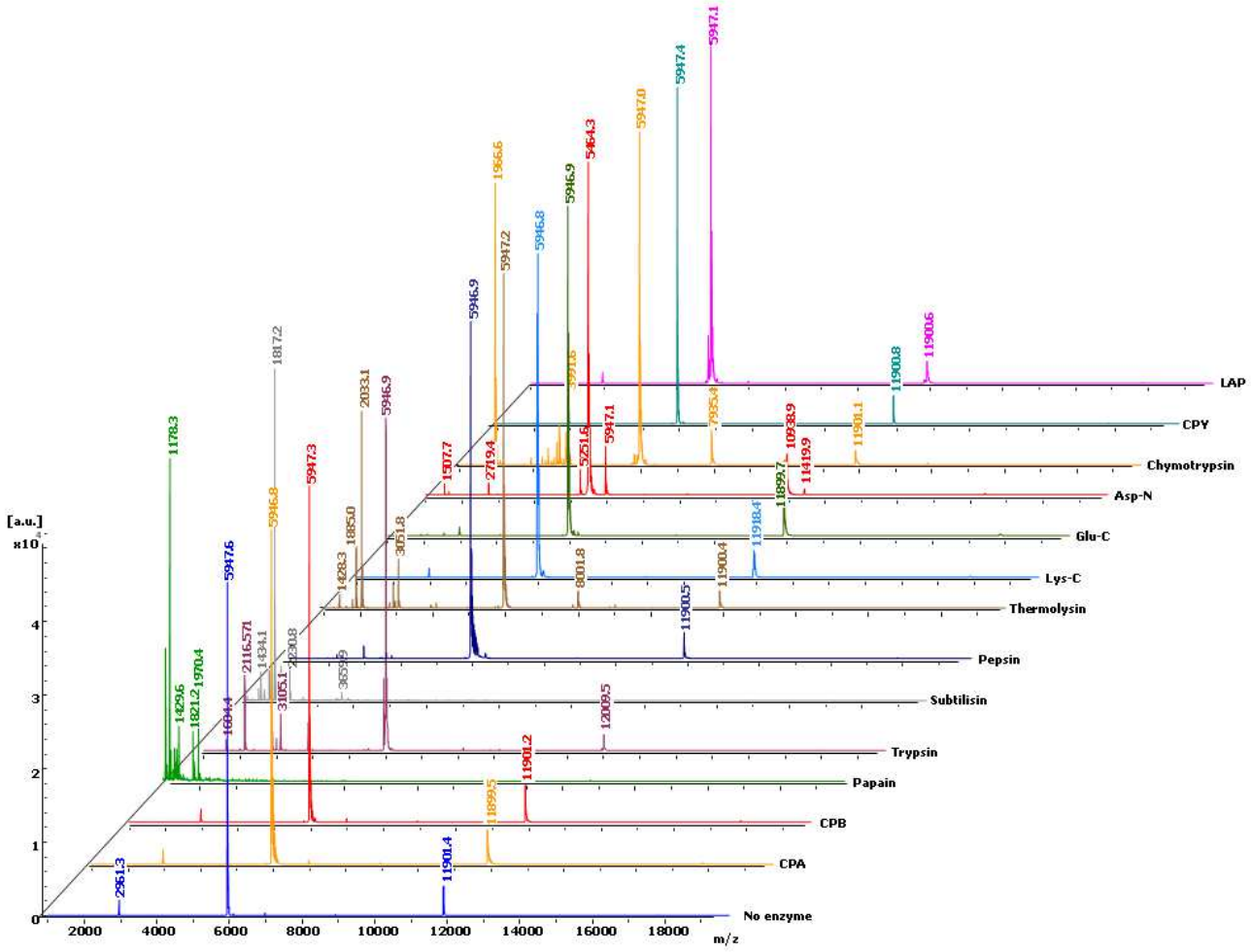


Figure 157. Enzymatic proteolysis of rNvCI at t=6 h.
Reaction conditions: Ratio (w:w) inhibitor:enzyme 10:1**,** T=37 ºC**.** LAP: Leucyl aminopeptidase

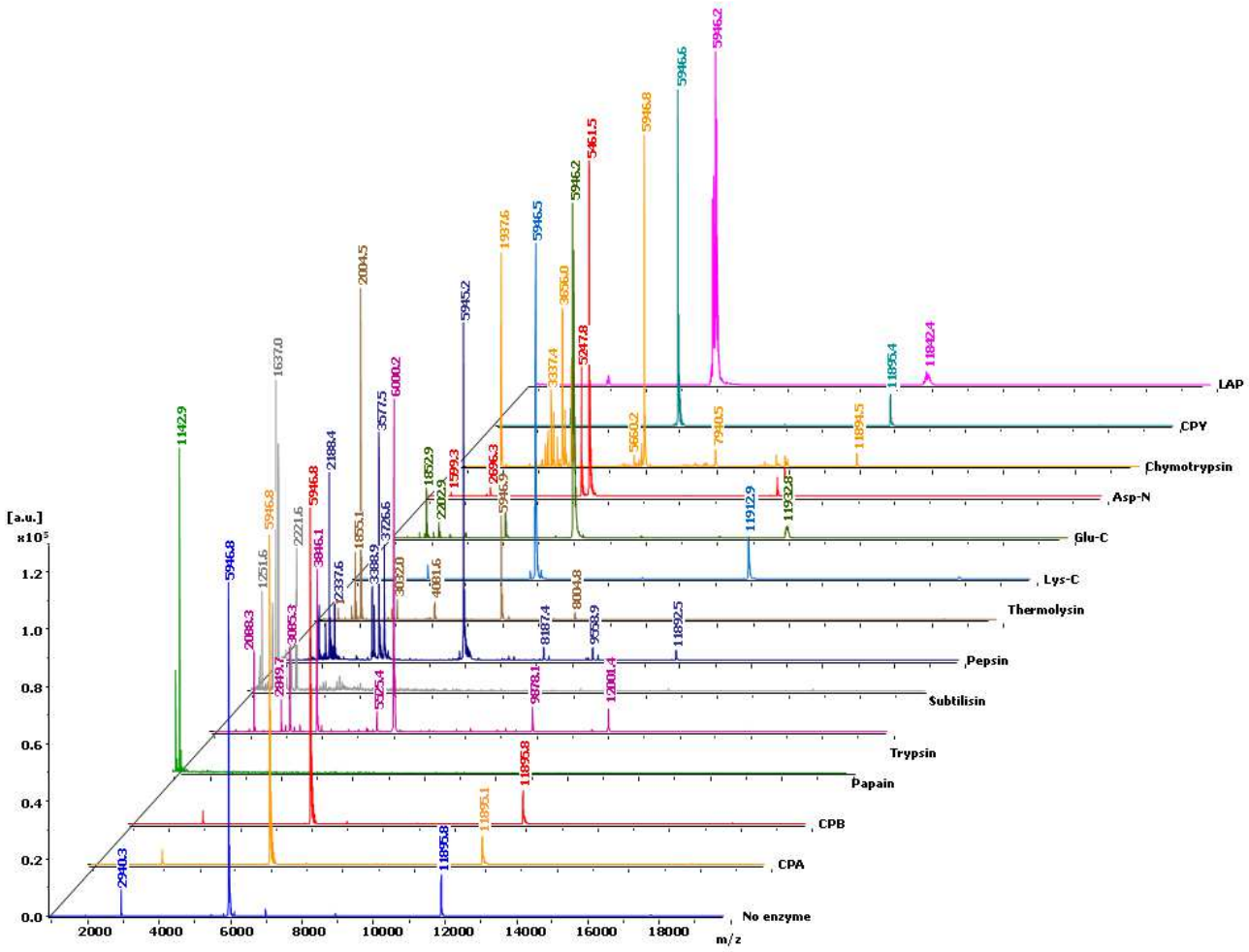


Figure 158. Enzymatic proteolysis of rNvCI at t=18 h.
Reaction conditions: Ratio (w:w) inhibitor:enzyme 10:1**,** T=37 ºC**.** LAP: Leucyl aminopeptidase

Table 60 summarizes the molecular masses and amino acid sequences of the peptides resulting from hydrolysis of rNvCI with proteases as a-chymotrypsin, endoproteinase Asp-N, pepsin and trypsin. For enzymes such as thermolysin, hydrolysis products could not be identified and in the case of papain and subtilisin hydrolysis was non-specific, generating rNvCI digestion products along with autoproteolysis peptides from the enzyme.

Papain displays fairly broad specificity with a preference for bulky hydrophobic or aromatic residues (Ala, Val, Leu, Ile, Phe, Trp, Tyr) in P2 position, Arg/Lys in P1 subsite, and any amino acid residue in P1' subsite with the exception of Val residue at this position (Kimmel and Smith, 1954). Thermolysin preferentially cleaves sites with bulky and aromatic residues (Ile, Leu, Val, Ala, Met, Phe) in P1' position, while pepsin exhibits cleavage specificity for Phe, Tyr Trp and Leu in P1 or P1' subsites (Keil, B., 1992). Regarding Subtilisin, this non-specific endoproteinase presents cleavage specificity for a large uncharged residue in P1 position (Ottessen and Svendsen, 1970; Boyer, P.D., 1970). On the other hand, α-chymotrypsin hydrolyses Tyr, Phe and Trp residues in P1 position, with a lesser extent to Leu residue in this position. A secondary hydrolysis will take place on the C-terminal side of Met, Ile, Ser, Thr, Val, His, Gly or Ala residues (Burrel, M.M., 1993).

Table 60. Summary of molecular masses and amino acid sequence of the resulting peptides from hydrolysis with proteases of rNvCI

Peptide (m/z	Sequence
α-chymotrypsin	
3727.9	FHVPDDRPCINPGRCPLVPDATCTFVCKAADNDF
3947.2	FHVPDDRPCINPGRCPLVPDATCTFVCKAADNDFGY
Endoproteinase Asp-N	
5461.5	DDRPCINPGRCPLVPDATCTFVCKAADNDFGYECQHVWTFEGQRVGCYA
Pepsin	
3728.6	FHVPDDRPCINPGRCPLVPDATCTFVCKAADNDF
Trypsin	
3846.1	CPLVPDATCTFVCKAADNDFGYECQHVWTFEGQR

The key findings are summarized as follows:

- 1 μg rNvCI was used for each enzymatic digestion.
- Proteolysis with Asp-N was found to be rapid and specific, resulting in the release of the first 4 residues from the N terminal of the protein during the first hour of reaction.
- Partial proteolysis of rNvCI with trypsin was observed at 18 hours of reaction. In this case, hydrolysis involved breaking of peptide bonds between R14-C15 and R48-V49.
- In pepsin and chymotrypsin digestions, one peptide at m/z 3727.9 was identified in both. In addition, a peptide at m/z 3947.2 was also observed using the former serine protease.
- Proteases such as papain and subtilisin were quite fast and non-specific. Enzymatic digestion with these enzymes displayed the release of rNvCI peptides at the first hour of reaction until total disappearance of the protein, according to MALDI-TOF MS analysis. These results are related to non-specifc hydrolysis and autoproteolysis of these enzymes, as described in literature.
- Other peptides generated during the enzymatic hydrolysis of rNvCI with thermolysin were observed. However, these could not be matched with peptides from the studied protein. This may be ascribed to non-specific proteolysis for the enzyme used, or as in the case of Glu-C, peptides from autoproteolysis.
- rNvCI hydrolysis using CPA, CPB, Lys-C, Glu-C, CPY and LAP did not release any resulting peptide.

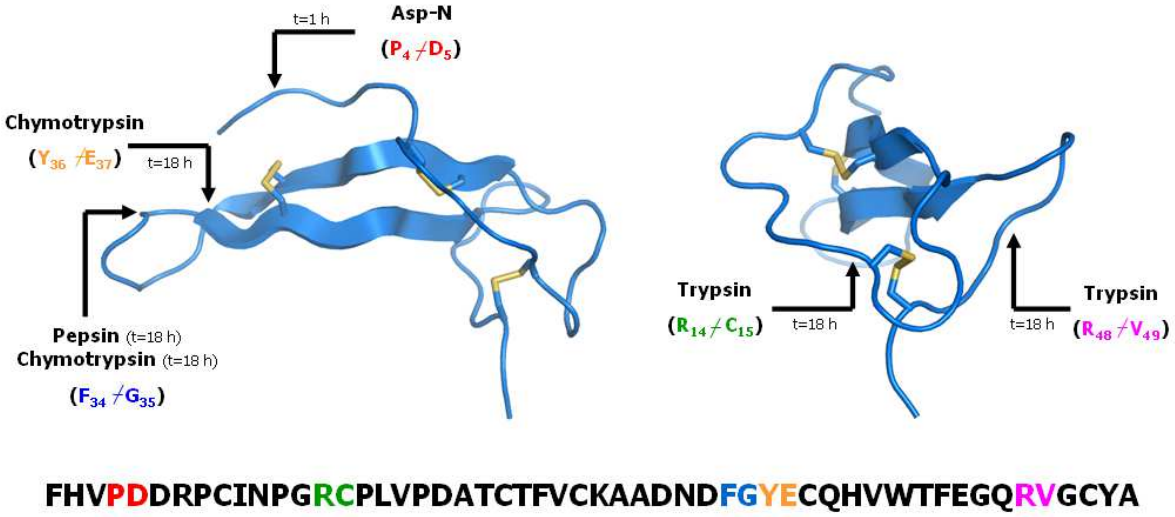


Figure 159. Three-dimensional analysis from hydrolysis with proteases of rNvCI

Analysis of peptides obtained as hydrolysis products related to the three-dimensional structure of rNvCI indicated that Asp-N hydrolyzed from almost the beginning of the hydrolysis (1 h), the flexible N-terminal tail of the inhibitor, which explains the high-speed digestion (figure 159). Chymotrypsin and pepsin digested the same peptide bond between F_{34}-G_{35} in the protein, which appears in an exposed loop. Trypsin also cut two peptide bonds: one between R_{14}-C_{15} and another on R_{48}-V_{49}, which are also located in exposed loops of the inhibitor.

IV.13. *IN VITRO* TESTS FOR ANTIMALARIAL ACTIVITY OF rNvCI

Antimalarial activity studies were performed using *Plasmodium falciparum* Dd2 strain (chloroquine resistant) synchronized by 85% in the ring stage (6-12 hours) and incubated for 48 hours with different concentrations of rNvCI. Subsequently, the percentage of parasite growth in culture was determined by a microfluoromethic method using PicoGreen®. In addition, the percentage of parasitemia and intraerythrocytic stages of the parasite were carried out by microscopic examination of Giemsa-stained parasitized erythrocytes (Radfar *et al.*, 2009).

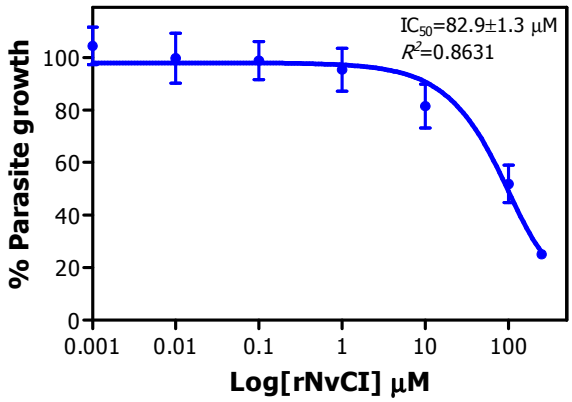


Figure 160. Dose-response relationship for antimalarial activity of rNvCI.
Erythrocytes were infected with *P. falciparum* Dd2 strain, 1% initial parasitemia, 2% hematocrit, 48 hours of treatment, tests by Picogreen. IC_{50} value was calculated by fitting the experimental data to a nonlinear regression using the Graphad prisma 5 software (GraphPad Software, Inc., USA) at $p<0.05$. Data are means (n=3 cultures) ± S.E.M.

As shown in figure 160, rNvCI displayed inhibitory activity within the range of concentrations used, since as inhibitor concentration was increased, percentage of parasite growth decreased. The IC_{50} value of rNvCI was 82.9 µM, which was approximately 4 times lower than the obtained for another carboxypeptidase inhibitor such as PCI (IC_{50}=330 µM, unpublished data). Results suggest that rNvCI behaves as a parasitostatic drug, inhibiting a key metallocarboxypeptidase in the parasite maturation process and erythrocyte invasion.

V. GENERAL DISCUSSION

Proteases and their inhibitors have become into very attractive molecules for Biotechnology and Biomedicine, due to their multiplicity of functions throughout the whole life cycle in all living organisms. Particularly metallocarboxypeptidases, aspartic, cysteine and serine proteases, the families which have been studied in this work constitute potential therapeutic targets for several systemic and infectious diseases. They are involved in cardiovascular, inflammatory disorder, allergic processes, cancer and osteoporosis, among others (Leung *et al.*, 2000; Abbenante and Fairlie, 2005; Turk, B., 2006; Turk *et al*., 2012). In addition, these types of enzymes play important roles in parasitic diseases as malaria, which represents a tropical disease causing more than 2 millions of deaths per year, especially children (Rosenthal, P.J., 1998). The detection of new therapeutic targets as proteases has opened possibilities for developing inhibitors as new antiparasite and antimalarial drugs (Rosenthal *et al*., 2002). This fact is in line with the call of WHO for devoting further efforts to fight such diseases, taking also into account their progressive extension into nearby areas and into developed countries, due to climatic changes, globalization, resistance of the parasites to traditional drugs, etc.

Moreover, it is convenient to emphasize that several of the current drugs have come from natural sources, due to the extraordinary chemical and functional diversity of biomolecules existing in living organisms. At present, advances in proteomics, genomics and bioinformatics have opened up new opportunities for the rapid and efficient identification of these biomolecules in natural complex mixtures, detection against molecular targets and their use as leader structures for the development of biochemically and pharmacologically more efficient molecules. Especially, marine invertebrates are very promising natural sources because of their abundance and variety, poor exploitation and the exceptional properties of their biomolecules, including proteases and protease inhibitors, as a result of their environment and functions of defence and predation (Mebs and Gebauer, 1980; Shiomi *et al*., 1985; Sencic and Macek, 1990; Castañeda *et al*., 1995).

In the present study, biological extracts from 30 different invertebrate species were selected, representing 7 Phyla, from the Caribbean Sea, a region described by its tropical conditions as one of the most attractive sources due to its biodiversity. Moreover, experience of both the Cuban and Spanish groups participating in this project have demonstrated that some of these Phyla show protease inhibition with novel properties (Chávez *et al*., 1998; Delfin *et al.,* 1996; Pascual *et al*., 2004; González *et al*., 2007a; Alonso del Rivero *et al.* 2012).

The strategy developed to identify CPA, CPB, pepsin, papain, trypsin and subtilisin inhibitors into these extracts comprised two types of approaches: traditional kinetic assays and proteomic-based methods such as IF (Intensity Fading) MALDI-TOF MS, which enabled the complementarity of approaches, using methods based on different concepts. Traditional kinetic assays are irreplaceable tools for the identification of these biomolecules, since the principle on which they are based is the assessment of catalytic or inhibitory functions by measuring the velocities of enzymatic reaction (Bieth, J.G., 1995; Copeland, R.A., 2000; Chávez and Gutiérrez, 2012), whilst the principle of the IF MALDI-TOF MS approach used is based on the evaluation of direct molecular interaction between target protease and inhibitor (Villanueva *et al*., 2003; Yanes *et al*., 2005, 2007). An additional advantage of IF MALDI-TOF MS is that the method can be jointly used with MALDI-TDS in order to quickly determine the sequence of resolved and/or single inhibitor peaks obtained in the elution fraction.

Using the first strategy in a high-throughput format (i.e., enzymatic and inhibition assays in microplate reader/dispenser) a group of active extracts for CPA, CPB, pepsin, papain, trypsin and subtilisin inhibitory activities were found. These results were strengthened with the quantitative determination of their corresponding specific inhibitory activities and dose-effect relationships. It is noteworthy that all the identified positive extracts displayed typical dose-effect behaviours (residual enzymatic activity vs. [I]), which supports the qualitative/quantitative identification performed.

The evaluation of the IC_{50} parameter (the amount of protein in the extract needed to produce 50% of inhibition) in selected conditions contributes to have a preliminary assessment of the inhibitory activity present in the extracts. However, this parameter must be analyzed with caution, because it depends on many experimental factors, such as the enzyme and substrate concentrations used in the assay. Here, such concentrations were maintained constant in each type of inhibition assay, which allows the comparison of results obtained for every type of protease. Nevertheless, in the case of crude extracts the IC_{50} values also depend on the ratio inhibitor/protein contents (Bieth, J.G., 1995; Copeland *et al.*, 1995).

The success of these kinetic procedures is essentially based on the experimental conditions selected to allow the detection of a clear inhibition capacity. Among them, high enzyme concentration, low substrate concentration and time to reach equilibrium are the most important. The first enables the detection of strong inhibitory activities, as well as the identification of tight-binding inhibitors. The dose-response curve for tight-binding protease inhibitors can vary dramatically from biphasic ($[E_0]/K_{iapp} \geq 10$) displaying a titrating behaviour typical of these kind of compounds to almost flatted shape ($[E_0]/K_{iapp} \leq 0.1$), which is characteristic of classical inhibitors ($K_i > 10^{-7}$ M). In practical terms, it means that a strong inhibitor can show behaviour identical to that of a weak inhibitor, according to the enzyme concentration used in the assay. Moreover, the low substrate concentrations avoid the dissociation of enzyme-inhibitor complex, and the use of enzyme-inhibitor preincubation time ensures reaction equilibrium, which are factors contributing to the success of such types of kinetic assays.

For the determination of inhibitory activity, the experimental conditions in terms of substrate concentration were kept in an interval of 0.5 to 2.3 K_M (Bieth, J.G., 1995). As expected, it was not possible in all cases to work with high enzyme concentrations, as it is dependent on the sensitivity of the substrate and on the need to guarantee of achieving of a linear reaction progress curve. In this work, this behavior was observed for CPA and CPB.

It is very noteworthy the prevalence of trypsin inhibitory activity (14 extracts) compared to other enzyme inhibitory activities. This result is explained by the greater abundance of serine proteases in these invertebrates due to their broader specificity (Hedstrom, L., 2002), probably contributing not only to the digestive processes of these animals, but also their defence and predation (Mebs y Gebauer, 1980; Shiomi *et al.*, 1985; Sencic y Macek, 1990; Castañeda *et al.*, 1995).

The results obtained for CPA, CPB, papain and subtilisin inhibitory activities were similar in terms of number of positive extract, with 9 extracts in each case, representing 30% of the set of analyzed extracts. Interestingly, CPA and CPB inhibitory activities were detected in the same kind of marine extracts. This fact may indicate that some molecules present in the extract are able to inhibit both enzymes, as have been described for other carboxypeptidase inhibitors from proteinaceous nature (Reverter *et* al., 1998; Arolas *et al.*, 2005; Sanglas *et al.*, 2009) as well as for NvCI, the most powerful inhibitor studied in this work. The lower frequency of finding subtilisin (serine protease) inhibitory activity (in 9 extracts), compared to trypsin inhibitory activity, could be related to the lower abundance of subtilisins in nature, at least as reflected in the number of annotated sequences in the MEROPS database for this family compared to the chymotrypsin/trypsin family (Rawlings *et al.*, 2002).

In contrast, pepsin inhibitory activity was clearly the less frequent (4 positive extracts) into the screening performed. This result is in agreement with the fact that proteinaceous aspartic protease inhibitors are rare in nature and unevenly distributed among classes of organisms in contrast to proteinaceous inhibitors of serine and cysteine proteases (Laing and McManus, 2002). Furthermore, it is not possible to discard that using other substrates more extracts could be found with this activity.

It is important to mention that these results could also be affected by the presence of endogenous and stable protease-inhibitor complexes that do not allow the detection of protease inhibitory activities. This fact was previously confirmed in different marine extracts, using agents allowing the dissociation of complexes (Delfin *et al.*, 1996). The endogenous protease inhibitors are generally tight-binding inhibitors, with values of K_i 10^{-7} - 10^{-8} M or even lower. Thus, they can not be easily dissociated by the presence of the substrate in the assay, particularly at the low substrate concentrations used in the assays (~1 Km) (Bieth, J.G., 1995).

On the other hand, IF MALDI-TOF MS has been shown as a useful method to confirm, by molecular interaction target-ligand (Villanueva *et al.*, 2003; Yanes *et al.*, 2005), the presence of protease inhibitors in the most promising marine invertebrate extracts, previously identified by kinetic enzymatic methods. The extracts were selected based on kinetic characteristics such as IC_{50} values less than 1 mg/ml, specific inhibitory activity greater than 10 mU/mg and bioavailability of the species in nature. This proteomic method was initially developed to detect the formation of complexes between proteins and other biomolecules in solution (Villanueva *et al.*, 2003) and later on modified by using the target protein immobilized on solid matrices (Yanes *et al.*, 2005). In the present work, a screening of protease inhibitors from multiple natural extracts has been developed, where for a first time the IF MALDI-TOF MS has been used as a confirmation criterion and characterization approach at medium-throughput level. In addition, this proteomic method has been extended to other immobilized target proteases such as CPB, subtilisin and pepsin.

The positive identification criterion was based primarily on the retention and subsequent elution/recovery of a MALDI-MS signal corresponding to the presence of a molecule able to selectively interact with the immobilized protease. However, in some positive extracts, previously identified by kinetic methods, the molecular interaction between the immobilized enzyme and biomolecules present in the extract could not be detected in the mass range used (1-20 kDa). This fact may be associated, among others, with the loss of some active components during the desalting process of each fraction previous MALDI analysis. Another reason could be related to the experimental conditions, which do not allow the release of molecules from the affinity matrix, a fact dependent on the characteristics of the inhibitor. These considerations must be taken into account for future determinations. However, the IF MALDI-TOF MS approach is a direct and powerful tool for identifying inhibitors in complex extracts, demonstrated in this work, with the additional advantage, besides its simplicity, high sensitivity and productivity, of providing the molecular mass of ligands and of having a less probability of obtaining false positives if the correct controls are used.

Summarizing, the use of both techniques, based on different principles complements each other, resulting in an overall strategy for the effective identification of inhibitors in complex mixtures of natural products. This overall approach can also be used to confirm molecular interactions when the ligand is an inhibitor and the enzyme is contained into the complex mixture (Yanes *et al.*, 2007; Alonso del Rivero *et al.*, 2009). We expected that our results contribute to the optimization and interplay of traditional and modern proteomic techniques for the identification of biomolecules in natural sources with potential applications in Biomedicine. In addition, all these results help to expand the knowledge on this type of biomolecules in invertebrates, a field much less studied in comparison with vertebrates, although the former represents 95% of the animal kingdom (Brusca and Brusca, 2003).

On the other hand, in the present work it has been demonstrated that the combination of IF MALDI-TOF MS and MALDI-TDS (MALDI-Top-Down Sequencing) allows to identify partial/total sequence of a molecule in the elution fraction of IF MALDI-TOF MS. It should be noted that the application of the MALDI-TDS approach was limited by the complexity of the mass spectrum obtained for the elution fraction of IF MALDI-TOF MS. Thus, CID fragmentation was only used for those spectra showing MALDI signals at low molecular mass range (1000-3000 m/z), as in the case of *H. carunculata* – pepsin NHS activated Sepharose; whereas ISD

fragmentation was employed in samples whose mass spectrum displayed a single signal or a major MALDI peak with other few components, as in the case of *S. helianthus* – trypsin glyoxal Sepharose. So, fragmentation strategies to be fitted to our IF MALDI-TOF MS approach are limited, by now, to certain conditions.

Immobilization of target proteases plays an important role not only in IF MALDI-TOF MS but also in the purification of the identified inhibitors. The glyoxal Sepharose supports used allows multipoint interaction of the ligand with the matrix, which leads to higher efficiency of immobilization, greater operational stability, long storage times and low cost. The introduction of a short spacer arm reduces steric hindrances and mass transfer limitations, which contributes to the greater efficiency of immobilization process, as well as their subsequent application as a matrix in affinity chromatography. Other advantages of this method are its simplicity and lack of toxicity (Guisan *et al.*, 1992). The only requirement of this method that could be a disadvantage for some enzymes is related to the alkaline pH values needed to guarantee the interaction of the ε-NH_2 groups of lysine residues with the aldehyde groups of the matrix. For these reasons, pepsin (labile in such conditions) cannot be immobilized through this method. However, the matrix NHS-Sepharose proved to be a suitable alternative support for the immobilization of this enzyme.

Immobilization yields in terms of protein and enzyme activity, measured by differential methods, in general were high for the 6 assayed proteases. However, the percentages of retention of functional activity, directly measured in the immobilized derivatives, generally were found to be low. However, these results are in agreement with the values described for other immobilized enzymes, taking into account that the enzymatic activity, directly determined on the immobilized derivative, is limited for external and/or internal mass transfer limitations (Suau *et al.*, 2009; Illanes *et al.*, 2010). External mass transfer limitations are easier to remove using a proper mixing system and other experimental conditions such as low substrate and high enzyme concentrations, as demonstrated in this work. However, internal mass transfer limitations are more complex and dependent on a higher number of factors, such as the matrix porosity, molecular mass and structure of the substrate, length of the spacer arm, low substrate and high enzyme concentrations, among others.

In most cases, a suitable agitation demonstrated to be enough to alleviate the external mass transfer limitations. However, the presence of internal mass transfer limitations were evidenced in all cases, with the exception of the immobilized trypsin derivatives. For instance, in the case of pepsin, the low percentages of retention of functional activity obtained are likely related to internal mass transfer limitations, as demonstrated by the low values of effectiveness factor assessed. Here, the greater length of the spacer arm used in the matrix may have favored internal mass transfer limitations. However, the immobilization degrees attained, in terms of protein and enzyme activity, are in the range required to perform an affinity chromatography for use both in IF MALDI-TOF MS and purification of protease inhibitors. This is related to the fact that the most promising inhibitors are tight-binding inhibitors with low K_i values ($<10^{-7}$ M) and therefore, a low immobilization degree must be used to avoid a strong interaction requiring drastic elution conditions. Low immobilization degrees along with high velocities in the affinity chromatography are the basic requirements to allow the success of this chromatography with this kind of molecules (Chávez *et al.*, 1990; Carlsoon *et al.*, 1998).

Metallocarboxypeptidase (MCP) inhibitors are important biochemical tools for the establishment of structure-function relationship of their molecular counterparts as well as in biomedicine against systemic and infectious diseases. Among the MCP inhibitors identified in this work, the inhibitor from the extract of *N. versicolor* was the most interesting and promising due to its ability to strongly inhibit both CPA and CPB (with K_i at sub-nanomolar levels), evidenced by a biphasic curve, typical of tight binding inhibitors, in comparison with other MCP inhibitors detected. In addition, the bioavailability of *N. versicolor* in nature and its easy collection justify the choice of this species for purification and functional-structural characterization of a MCP inhibitor.

The purification of compounds from natural sources is often complex, generally dependent on the own complexity of the extracts and the presence of small quantities of bioactive molecules (Li and Vederas, 2009). The strategy developed in the purification of molecules from natural sources is in agreement with the requirements proposed by Janson and Ryden, 1998. These requirements are the presence of an initial stage of separation, extraction and clarification as a procedure of high capacity for sample processing and low resolution, followed by a second step of capture, which allows to isolate a fraction enriched in the component of interest and finally a polishing step to recover the component with a high degree of purity.

For the purification of protease inhibitors different initial fractionation strategies of the crude extracts have been used, such as heath treatment, such as acid (HCl, TCA, acetic acid) and basic treatments (KOH), precipitation with ammonium sulfate and organic solvents as ethanol or acetone (Chavez *et al*., 1990), among others. Taking into account the stability of the inhibitors isolated from these sources, due to their functions of defense into aggressive environmental conditions, we have submitted the extract to a heath treatment to at least 60 °C, 30 min. The heat treatment has the advantage of not being an invasive method, eliminating thermo labile components such as proteases and contributing in many cases to the dissociation of endogenous protease-inhibitor complexes. This strategy has been described already as a suitable step in the purification of protease inhibitors from marine invertebrates, such as the CPA inhibitor (SmCI) isolated from *S. magnifica* (Alonso del Rivero *et al*., 2012), and a human neutrophil elastase inhibitor (CmPI-II) from the mollusc *C. muricatus* (González *et al*., 2007a).

The use of affinity chromatography as a capture step is successful due to its high resolution and selectivity, biospecific interaction of ligand molecules of the extract and the immobilized target protease. This is particularly useful when the protein of interest is a minor component in a complex mixture, thus obtaining an elution fraction enriched in the component of interest. Another relevant factor to be considered, specifically in the purification of protease inhibitors is the strength of protease-inhibitor interaction. For NvCI, which is a tight-binding inhibitor, a relatively low immobilization degree of CPA was used (3.5 mg CPA/ml gel) that together with an appropriate flow rate allowed the efficient capture and subsequent elution of the inhibitor with a high yield. It is important to mention that the affinity chromatography has been successfully applied in the purification of CP inhibitors, such as PCI (Hass *et al*., 1979; Homandber, *et al*., 1989), TCI (Arolas *et al*. 2005a), ACI (Sanglas *et al*., 2009) and SmCI (Alonso del Rivero *et al*., 2012).

The reversed-phase chromatography, recommended as a polishing step in a protein purification process (Amersham Pharmacia Biotech, 2001), provided a good degree of purity for NvCI but a low yield. The low yield obtained in this chromatography can be attributed to factors such as low stability of the molecule in the rather drastic conditions of elution or retention of the molecule to the matrix, which has been reported for this type of chromatographic procedure (Janson and Ryden, 1998). However, the yield obtained in the purification process of NvCI was enough to obtain a proper amount of highly pure inhibitor used in further studies of structural and functional characterization.

The simultaneous presence of isoforms of slightly different masses and elution properties was detected in our extracts, as in the case of NvCI and *Nerita versicolor* species, as it has been frequently described for inhibitors isolated from marine invertebrates (Fritz *et al*., 1972; Delfín *et al*., 1996; González *et al*., 2007a). Replacement or loss of one or more amino acid residues is explained on the basis of genetic mutations, and also because of unspecific proteolysis but in some cases the occurrence of specific post-translational chemical modifications cannot be discarded (Collins and Jukes, 1994; Delfin *et al*., 1996; Minagawa *et al*., 1997; Huerta *et al*., 1998; González *et al*., 2007a). The study of inhibitory activity for the 4 purified NvCI isoforms revealed an identical functional behavior, as has been previously described for inhibitor isoforms purified from other species (Friedrich *et al*., 1993; Johansson *et al*., 1994; La Barre *et al*., 1996; Mende *et al*., 1999; Campos *et al*., 2004). Therefore, NvCIa isoform, denominated NvCI in this work and representing one of the two major components, was selected for further structural and functional characterization.

The combination of methods and strategies such as Edman degradation, *de novo* sequencing by MALDI-TOF MS and the synthesis and cloning of cDNA demonstrated that the NvCI major form is a polypeptide of 53 amino acid residues, with 3 intra-chain disulfide bonds and an amino acid sequence different from other MCP inhibitors previously described. As mentioned above, only 6 proteinaceous inhibitors of MCPs have been structurally studied. These inhibitors, with the exception of latexin and SmCI, show sequence homology at the C-terminal tail. NvCI, as the other exogenous protein MCP inhibitors PCI, LCI, TCI and ACI, also displays sequence homology at the C-terminal tail for its P1 and P2 positions. Interestingly, the C-terminus sequence of NvCI, as the MCP inhibitor from tomato and ACI, does not have the last amino acid residue (corresponding to P1' position) that is trimmed in the other MCP inhibitors once the interaction with the carboxypeptidase takes place. In the case of NvCI, as demonstrated by the synthesis and cloning of the cDNA, this amino acid residue is not present in this molecule at the gene level but, as confirmed later, its inhibition mechanism is similar to other known exogenous MCP inhibitors, where the inhibitor tail mimics substrate binding (Rees and Lipscomb, 1982; Reverter *et al.*, 2000; Arolas *et al.*, 2005b). Noteworthy, the isoforms of NvCI differ at the N-terminus, showing very similar functional properties; this fact confirms the nulle involvement of its N-terminus in the inhibitory mechanism.

Heterologous expression of NvCI (its main form) in *P. pastoris* system is an important achievement for the development of this work, since the production of active protein in large quantities (329.7 mg rNvCI/liter of culture broth), greatly facilitated the functional and structural characterization of the protein, the crystallization and elucidation of its 3D structure, as well as its application in preliminary studies of biomedical applicability.

As described for other exogenous MCP inhibitors, natural and recombinant NvCI are able to strongly inhibit A/B-type MCPs but not regulatory MCPs, as is the case of CPD-I. In addition, NvCI is not capable of inhibiting proteases belonging to other mechanistic classes, as the other exogenous MCPs inhibitors, with the exception of SmCI (Alonso del Rivero *et al.*, 2012).

But now, the most relevant characteristics found for NvCI are related to its inhibitory function, sustained by some novel features of it in comparison with other MCP inhibitors of the same type. It is important to note that in the same way as previously emphasized, the importance of proper selection of experimental conditions for the inhibition assays and strategies for the determination of kinetic parameters are crucial for a correct kinetic characterization of tight binding inhibitors (Bieth, J.G., 1995; Copeland, R.A., 2000; Chávez and Gutiérrez, 2012).

The selection of titration conditions ($[E_0]/K_i>10$), as previously used in the inhibitory activity assays, are also required to determine the active enzyme concentration and subsequently the active inhibitor concentration. This is an important initial step for an accurate functional characterization of a tight-binding inhibitor. The use of a tight-binding inhibitor such as TCI for CP titration proved to be an effective strategy for determining the active concentration of CP enzymes, subsequently used for NvCI titration. Biphasic behavior of the curves obtained with NvCI, is characteristic of a tight-biding inhibitor. The knowledge of the active inhibitor concentration is essential for the determination of K_i value.

Unlike the experimental conditions used for the determination of active enzyme and inhibitor concentrations, the conditions for the determination of K_i values require low $[E]_0 / K_i$ ratio ($\leq$ 10), as well as the use of low substrate concentrations, in the order of K_M value, reducing the induction of dissociation protease-inhibitor complex by the substrate. These experimental conditions allows to obtain dose-response curves with a concave behavior, suitable for fitting the data to the Morrison's equation (Morrison, J.F., 1982) for tight-binding inhibitors and the determination of K_i value. The concave curves achieved in v_i/v_0 *vs* [I] for the most enzymes assayed confirm that NvCI behaves as a reversible inhibitor.

Although NvCI is a competitive inhibitor, which has proved to have a substrate-like mechanism of inhibition, the dissociation of the enzyme-inhibitor complex induced by the

substrate could not be detected for some CPs. This result is likely related to the strength of the enzyme-inhibitor interaction. In the case of hCPA2, where the highest K_i values were obtained (nanomolar range), the effect of substrate concentration on the dissociation of enzyme-inhibitor complex was evidenced for both natural and recombinant NvCI, while for other CPs with much lower K_i values, such effect could not be detected. However, other factors such as the fast complex dissociation before starting the test or not enough high substrate concentration to display this effect (50 μM, 100 μM and 150 μM AAFP for CPA-like enzymes and 50 μM, 100 μM and 150 μM AAFA for CPB-like enzymes) may be responsible for the apparent absence of dissociation in some cases.

In addition, the time required to reach the system equilibrium, which is a prerequisite to ensure the determination of K_i values, shows that NvCI behaves as a slow inhibitor against some of the proteases tested. The slow inhibition is not only an attribute of tight-binding inhibitors, as it is described also for classical inhibitors, although more often takes place in the first ones (William and Morrison, 1979).

There is no experimental evidence to explain in detail the mechanism of tight-binding inhibition. However, several hypotheses have been postulated (Morrison, 1982; Copeland, 2000), which suggests two models. The first model justifies the inhibition by the formation of two enzyme-inhibitor complexes: one less active complex, which is rapidly formed, followed by a second more active complex, that take place after a conformational change or a rearrangement of groups in the first complex, lead to the formation of a stronger complex between both molecules. The second model is based on the fact that the inhibitor concentration is so small, due to their tight-binding, that hinders the mass diffusion of the inhibitor and its binding to the enzyme, despite slow classical inhibitors (less frequent) require higher inhibitor concentrations in the inhibition assays (Morrison, J.F., 1982; Copeland, R.A., 2000).

The reported proteinaceous MCPs inhibitors are tight-binding, reversible and competitive (except for latexin) inhibitors against A/B-type CPs (M14A subfamily) with K_i values in the nanomolar range (Vendrell *et al*. 2004; Arolas *et al*., 2007; Sanglas *et al*., 2009).

Table 61. Comparison of K_i values (nM) between various proteinaceous MCPs inhibitors

CP	rNvCI	rPCI [(a)]	rTCI [(b)]	rLCI [(c)]	rACI [(d)]	rhuman Latexin [(e)]
bCPA	0.0058 ± 0.0002	1.6 ± 0.2	1.1 ± 0.3	1.1 ± 0.2	2.4 ± 0.3	1.2 ± 0.2
hCPA1	0.0012 ± 0.0001	1.6 ± 0.3	1.2 ± 0.4	1.3 ± 0.4	1.6 ± 0.2	1.6 ± 0.2
hCPA2	2.9 ± 0.1	8.8 ± 0.7	3.6 ± 0.5	4.5 ± 0.6	2.5 ± 0.2	3.5 ± 0.3
hCPA4	0.0049 ± 0.0006	1.3 ± 0.1[(f)]	0.8 ± 0.3 [f)]	7.3 ± 0.4[(f)]	23.9 ± 3.1	3.0 ± 0.3
pCPB	0.335 ± 0.005	7.2 ± 0.6	1.6 ± 0.3	2.4 ± 0.5	NA	NA
hCPB	0.549 ± 0.002	1.8 ± 0.3	1.3 ± 0.2	1.2 ± 0.3	NA	1.1 ± 0.1
bTAFI	0.08 ± 0.01	2.5 ± 0.4	1.3 ± 0.3	NA	NA	NA
hTAFI	1.0 ± 0.1	5.2 ± 0.5	1.2 ± 0.4	NA	42.0 ± 1.7	1.8 ± 0.3

a, b. K_i values from Arolas *et al*., 2005a. **c**. K_i vakues from Arolas *et al*., 2004. **d**. K_i values from Sanglas *et al*., 2009. **e**. K_i values from Pallarès *et al*., 2005. **f**. K_i values from Tanco *et al*., 2010. NA: data not available

Comparison of K_i values obtained for both natural and recombinant NvCI with other proteinaceous MCP inhibitors (table 61) clearly demonstrates the major strength of NvCI inhibition against most of the carboxypeptidases assayed. The lower K_i values of NvCI obtained for bCPA, hCPA-1, and hCPA4, compared with other MCP inhibitors, indicates that this inhibitor is approximately 1000 times more potent against these three carboxypeptidases. The K_i values of NvCI against enzymes such as pCPB, hCPB, bTAFI and hTAFI, in the range of $1x10^{-10}$ M, show to be approximately 10 times lower than those reported for MCP inhibitors described in table 61. However, in the case of the hCPA2, the K_i values obtained for NvCI were in the nanomolar range, like the other inhibitors described. This result, as explained above in the analysis of the three-dimensional structure of NvCI, may be associated with the absence of two hydrogen bonds formed between NvCI and this enzyme, which were observed in the amino groups of Cys51 and Tyr52 in NvCI and the Glu63 residue of hCPA4.

Moreover, comparison of the NvCI sequence with other MCPs inhibitors, based on their inhibition mechanisms, led to analyze some aspects related to the functional activity of NvCI. The inhibition mechanism described for most MCPs inhibitors involves the C-terminal tail of the molecule, which protrudes from the core of the protein and is inserted into the active-site of the protease, similar to the substrate binding (Vendrell *et al*., 2000, Vendrell *et al*., 2004). These inhibitors show sequence similarity in the region comprising the C-terminus of the protein and the importance of this region in the interaction of the inhibitors PCI, LCI, TCI, and ACI with different MCPs has been demonstrated (Rees and Lipscomb, 1982 ; Reverter *et al*., 2000; Arolas *et al*, 2005a; Sanglas *et al*., 2009). Furthermore, by site-directed mutagenesis studies in PCI has been shown that the sequence Pro-Tyr-Val plays a critical role in the MCPs inhibition. Thus, the proline residue gives rigidity to the C-terminus of the inhibitor, allowing adequate guidance towards the active-site of the protease (Marino- Buslje *et al*., 2000).

The few reports that appeared in the last decade on the structure-function relationships of MCP inhibitors of exogenous origin (Hass *et al*., 1975; Reverter *et al*., 1998; Arolas *et al*., 2005a; Sanglas *et al*., 2009), after the initial one from potatoes (Rees and Lipscomb, 1982; Marino- Buslje *et al*., 2000), indicated that they share a similar substrate-like mechanism of inhibition. Thus, they suffer a trimming of their C termini, behaving as competitive tight-binding inhibitors, mimicking the interaction of a peptide substrate with the active site of the enzyme, and requiring only the S1 and S2 subsites to be covered to fully inhibit the enzyme. Thus, even though these exogenous MCP inhibitors are isolated from evolutionarily distant organisms, this is a good example of convergent evolution dictated by the architecture of the active site of the enzyme. However, the structure of NvCI with hCPA4 indicates that the trimming action is absent (*i.e.* a shorter two-residue tail is sufficient) and that the S3 subsite is also implicated, in addition to S1 and S2, in promoting stronger inhibition in the picomolar range.

In addition to this primary and essential C-terminal interaction, MCP inhibitors possess a second interface (distant from active site) that covers different regions of the carboxypeptidase and confers stability to the complex formation (figure 154). The NvCI interface with hCPA4 is quite extended in comparison with the other known inhibitors, covering a total of 1875 Å2. In other known cases, for the protein inhibitors from leech (LCI), potato (PCI), and the *Ascaris* worm (ACI), the interface of CPA-like enzymes is clearly smaller: 1509, 1241, and 1426Å2, respectively. An exception would be the two-domain (exosite-behaving) tick carboxypeptidase inhibitor (TCI), which is the largest and covers a total interface of 2108 Å2. Although the secondary interaction surface provided by all inhibitors is distinct, there is only a single way of positioning the C-terminal tail in the active site of the enzyme.

As mentioned above, the main body of each of the exogenous carboxypeptidase inhibitors has a completely different three three-dimensional structure, with the C-terminal tail as the only similar structural motif (figure 152, *B* and *C*, and figure 154). The short C-terminal tail of NvCI, considered from the third disulfide bridge and formed by only two residues, Tyr-52 and Ala-53, is sufficient for a tight-binding inhibition of several types of carboxypeptidases. Interestingly, as mentioned above, in NvCI, the P3 position (Cys-51) also participates in the binding with a double main chain hydrogen bond with Glu-163 of hCPA4, presumably increasing the affinity and lowering the *Ki* value for NvCI, which is in the picomolar range, 3 orders of magnitude lower than the other inhibitors. This fact suggests a reason for the strongest inhibitor being NvCI in comparison with the other described proteinaceous inhibitors of MCPs.

The different ways nature utilizes to inhibit carboxypeptidases can be used as a valuable tool to elucidate the determinants for a general mechanism of inhibition of carboxypeptidases. Carboxypeptidases are proteolytic enzymes involved in several functions of the organisms, from food digestion to blood coagulation, inflammation, hormone/neuropeptide processing, insect/vegetal attack defence strategies, among others (Reynolds *et al.,* 1989; Bayés *et al*., 2005; Sanglas *et al*., 2008; Sanglas *et al*., 2010). In general, these proteases are secreted and their enzymatic action takes place normally in the extracellular space, except in the case of a novel subfamily of cytosolic carboxypeptidases, which have been recently described and are presumably involved in tubulin processing (Rodríguez de la Vega *et al*., 2007; Kalinina *et al*.,

2007). Knowledge of the control of interference with those mechanisms, by rational structure-based or other drug design approaches, can be potentially of great interest, especially for biotechnical and biomedical industrial purposes. In addition, the expansion of this strategy to the variety of forms found in the very diverse phyla of invertebrates, including marine organisms such as *N. versicolor*, is revealing a rich source of discovery for novel scaffolds and lead compounds.

Finally, the stability of NvCI against proteolytic enzymes of different mechanistic classes was performed, as preliminar knowledge for its pharmacological study as antimalarial agent. The initial study of the antimalarial activity of NvCI demonstrated that this molecule is able to decrease the percentage of parasite growth (IC_{50} value of 82.9 μM), suggesting that NvCI behaves as a parasitostatic drug, inhibiting a key metallocarboxypeptidase in the parasite maturation process and erythrocyte invasion.

In summary, in one of the most focused parts of this work was carried out the structural and functional characterization of the first MCP inhibitor isolated from the marine snail *N. versicolor*, called NvCI. The protein consists of 53 amino acid residues with a molecular mass of 5944 Da, and three disulfide bridges formed between Cys23 and Cys9, Cys15 and Cys51 and Cys27 and Cys38. The three-dimensional structure of NvCI reveals a new extended protein-folded small motif, which is basically formed by a central anti-parallel β-two-stranded sheet connected by three major loops. Regarding functional activity, NvCI is able to tightly inhibit A/B-type carboxypeptidases (M14A subfamily), but not other MCPs subfamilies, nor proteases from other mechanistic classes. The strength of NvCI inhibition against most of the proteases assayed compared to the K_i values for other MCPs reveals that NvCI is the strongest proteinaceous MCP inhibitor described so far.

VI. CONCLUSIONS

1. A new and successful strategy for the identification of protease inhibitors in complex biological samples was designed, that combines kinetic determinations of inhibitory activity and dose-response relationships in medium/high-throughput formats, with the proteomic approach IF MALDI-TOF MS, based on molecular interactions between the target proteases and selective inhibitors from the extract.

2. The strategy applied to the screening of 30 marine invertebrate extracts revealed that they represent a rich source of trypsin inhibitors (14 positive extracts) in terms of specific inhibitory activity and IC_{50} value, as well as of inhibitors of M14 metallocarboxypeptidases (MCPs, of the CPA and CPB type), papain and subtilisin inhibitors (with 9 positive extracts for each enzyme), but only a few number of extracts with pepsin inhibitory activity (4 positive extracts). These results are in agreement with the relative abundance in nature of these proteases.

3. The immobilization of the six target proteases on glyoxal and NHS Sepharose, after an extensive experimental study of immobilization conditions, led to the optimization of the process in terms of degree of capture and retention of functional activity, allowing its successful use as affinity matrices in IF MALDI-TOF MS and purification procedures.

4. The Proteomic approach IF MALDI-TOF MS quickly and specifically identified the presence of interacting (inhibitory) molecules of trypsin and CPA proteases in the marine crude extracts among the most promising, through molecular interaction with the immobilized target enzymes. This approach was positively extended to other immobilized proteases such as CPB, pepsin, papain and subtilisin. In addition, MALDI-TDS, including CID and ISD fragmentations, was successfully applied in two cases, enabling the partial or total sequencing of the protein species corresponding to selected elution peaks from IF MALDI-TOF MS, as a product of the specific interactions between the target protease and the biological extract. This strategy allows the early and quick characterization of protease inhibitors, and facilitates their recombinant production.

5. NvCI, the most potent MCP inhibitor, identified by IF MALDI-TOF MS from the marine mollusc *Nerita versicolor*, was successfully purified by the combination of a thermal clarification step, affinity chromatography on CPA-Glyoxal Sepharose and RP-HPLC. This inhibitor is a polypeptide of 5944 Da, with 53 amino acid residues and three disulfide bridges. The alignment of the amino acid sequence of NvCI did not allow the assignment to any family of MCP inhibitors described, although the comparison of its C-terminal tail with other MCP inhibitors revealed some similarities between them, confirmed later by structural studies.

6. The recombinant production of NvCI allowed to obtain large quantities of pure and active protein, at the level of around 330 mg/l of *Pichia pastoris* culture medium, for its functional and structural characterization. Both natural and recombinant NvCI show a very strong inhibitory activity against A/B-type M14 carboxypeptidases, some of them with K_i in the picomolar range. The strength of the interaction shown between NvCI and M14 carboxypeptidases assayed reveals that the former molecule is the most potent tight-binding inhibitor of MCPs from proteinaceous nature described so far.

7. The three-dimensional structure of NvCI in complex with human CPA4, determined by crystallization and X-ray diffraction (1.7 Å), allowed to establish that this inhibitor contains a central anti-parallel two-stranded β-sheet connected by three major loops and stabilized by three disulfide bridges. The inhibition mechanism of NvCI is due to a competitive interaction in a substrate-like manner through its C-terminal tail with the active site of the carboxypeptidase, by occlusion of the active-site subsites S1', S1, S2 and S3, as it has been observed for most of the reported proteinaceous MCP inhibitors. The "primary" and "secondary" interaction regions create an extended interface with the CP enzyme that minimizes the product release from the catalytic reaction, explaining the lower K_i values observed for NvCI in comparison to the other MCP inhibitors.

8. All of the three-dimensional structures of exogenous proteinaceous inhibitors of MCPs from evolutionarily distant organisms are unrelated and completely different. The only conserved motif in all of them is the conformation adopted by the P1 and P2 residues of the C-terminal tail, when bound to the M14 carboxypeptidase, which can be considered as a general feature, derived from a common functional mechanism for MCP inhibitors, and as a clear example of convergent evolution.

VII. REFERENCES

1. Abe, K., Y. Emori, H. Kondo, S. Arai, and Suzuki, K. The NH2-Terminal 21 Amino Acid Residues Are Not Essential for the Papain-Inhibitory Activity of Oryzacystatin, a Member of the Cystatin Superfamily. *J. Biol. Chem.* 1988; 263: 7655-7659.

2. Abbenante, G., and Fairlie, D.P. Protease inhibitors in the clinic. *Med. Chem.* 2005; 1: 71-104

3. Abrahamson, M., Nathanson, C.M., and Álvarez-Fernández, M. Human cystatins – similarities, diversity and classification. In: Zerovnik E, Kopitar Jerala N, eds. Human stefins and cystatins. 2006; 1-21, New York: Nova Biomedical Books

4. Adams, P.D, Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J. J.; Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W. Oeffner, R. Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H.. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Cryst. D*. 2010; 66: 213-221.

5. Adler, M., Buckman, B., Bryant, J., Chang, Z., Chu, K., Emayan, K., Hrvatin, P., Islam, I., Morser, J., Sukovich, D., West, C., Yuan, S., and Whitlow, M. Structures of potent selective peptide mimetics bound to carboxypeptidase B. *Acta Crystallogr D Biol Crystallogr*. 2008; 64(2): 149-57.

6. Agarwala, K.L.; Kawabata, S.; Hirata, M.; Miyagi, M.; Tsunasawa, S., and Iwanaga, S. A cysteine protease inhibitor stored in the large granules of horseshoe Crab hemocytes: purification, characterization, cDNA cloning and tissue localization. *J. Biochem.* (Tokyo). 1996; 119: 85-94.

7. Aghajanian, C.; Soignet, S.; Dizon, D. S.; Pien, C. S.; Adams, J.; Elliott, P. J.; Sabbatini, P.; Miller, V.; Hensley, M. L.; Pezzulli, S.; Canales, C.; Daud, A.; Spriggs, D. R. *Clinical Cancer Research*. 2002; 8: 2505-2511.

8. Alonso del Rivero, M., Trejo, S., Rodríguez de la Vega, M., González, Y., Bronsoms, S., Canals, F., Delfín, J., Díaz, J., Avilés, F.X., and Chávez, M.A. A novel metallocarboxypeptidase-like enzyme from the marine annelid *Sabellastarte magnifica*– a step into the invertebrate world of proteases. *FEBS Journal*. 2009; 276(17): 4875-4890.

9. Alonso-del-Rivero, M., Trejo, S.A., Reytor, M.L., Rodriguez-de-la-Vega, M., Delfin, J., Diaz, J., González-González, Y., Canals, F., Chávez, M.A., Avilés, F.X. Tri-domain bifunctional inhibitor of metallocarboxypeptidases A and serine proteases isolated from marine annelid *Sabellastarte magnifica*. *J Biol Chem*. 2012; 287(19): 15427-38.

10. Aloy, P., Companys, V., Vendrell, J., Aviles, F.X., Fricker, L.D., Coll, M., and Gomis-Rüth FX. The crystal structure of the inhibitor-complexed carboxypeptidase D domain II and the modeling of regulatory carboxypeptidases. *J Biol Chem*. 2001; 276(19): 16177-84

11. Álvarez, C.; Tejuca, M.; Morera, V.; Besada, V.; Pazos, F.; Lanio, M. E. and Padrón, G. Novel primary structure of sticholysin and its interactions with membranes. *Toxicon*. 1996; 34: 301-301.

12. Álvarez-Santos, S., González-Lafont, A., Lluch, J.M., Oliva, B., and Avilés, F.X. Theoretical study of the role of arginine 127 in the water-promoted mechanism of peptide clevage by carboxypeptidase A. *New J Chem*. 1998; 22: 319-25.

13. Anastasi, A., Brown, M.A., Kembhavi, A.A., Nicklin, M.J., Sayers, C.A., Sunter, D.C., and Barrett, A.J.. Cystatin, a protein inhibitor of cysteine proteinases. Improved purification from egg white, characterization, and detection in chicken serum. *Biochem. J*. 1983; 211: 129 – 138.

14. Andersson, H.O., Fridborg, K., Löwgren, S., Alterman, M., Mühlman, A., Björsne, M., Garg, N., Kvarnström, I., Schaal, W., Classon, B., Karlén, A., Danielsson, U.H., Ahlsén, G., Nillroth, U., Vrang, L., Oberg, B., Samuelsson, B., Hallberg, A., and Unge, T. Optimization of P1-P3 groups in symmetric and asymmetric HIV-1 protease inhibitors. *Eur J Biochem*. 2003; 270(8): 1746-58.

15. Antuch, W.; Berndt, K. D.; Chávez, M. A.; Delfín, J. and Wuthrich, K. The NMR solution structure of a Kunitz- type proteinase inhibitor from the sea anemone *Stichodactyla helianthus*. *Eur J. Biochem*. 1993; 212: 675-684.

16. Aoyagi, T., Morishima, H., Nishizawa, R., Kunimoto, S., and Takeuchi, T. Biological activity of pepstatins, pepstanone A and partial peptides on pepsin, cathepsin D and renin. *J Antibiot (Tokyo)*. 1972; 25(12):689-94.

17. Aoyagi, T., Kunimoto, S., Morishima, H., Takeuchi, T., and Umezawa, H. Effect of pepstatin on acid proteases. *J Antibiot (Tokyo)*. 1971; 24(10): 687-94.

18. Ardelt, W., and Laskowski, M.Jr. Effect of single amino acid replacements on the thermodynamics of the reactive site peptide bond hydrolysis in ovomucoid third domain. *J. Mol. Biol*. 1991; 220: 1041–1053.

19. Arolas, J.L., Bronsoms, S., Lorenzo, J., Avilés, F.X., Chang, J.Y., and Ventura, S. Role of kinetic intermediates in the folding of leech carboxypeptidase inhibitor. *J Biol Chem*. 2004; 279(36): 37261-70.

20. Arolas, J.L.; Lorenzo, J.; Rovira, A.; Castella, J.; Aviles F.X. and Sommerhoff, C.P. A carboxypeptidase inhibitor from the tick *Rhipicephalus bursa;* isolation, cDNA cloning, recombinant expression, and characterization. *J. Biol. Chem*. 2005; 280: 3441-8. **(a)**.

21. Arolas, J.L.; Popowicz, G.M.; Lorenzo, J.; Sommerhoff, C.P.; Huber, R. and Aviles F.X. The three-dimensional structures of tick carboxypeptidase inhibitor in complex with A/B carboxypeptidases reveal a novel double-headed binding mode. *J Mol Biol*. 2005; 350: 489-98. **(b)**.

22. Arolas, J.L.; Vendrell, J.; Aviles, F.X. and Fricker, L.D. Metallocarboxypeptidases: emerging drug targets in biomedicina. *Curr Pharm Des*. 2007; 13(3):347-64.

23. Avilés F.X. and Vendrell J. Carboxypeptidase B. In: *Handbook of Proteolytic Enzymes*. 1998; pp. 1333–1335, Barrett A. J., Rawlings N. D. and Woessner J. F. (eds), Academic Press, San Diego (CA), USA.

24. Azzolini S.A, S., Sasaki, S.D., Torquato, R.J., Andreotti, R., Andreotti, E., Tanaka, A.S. *Rhipicephalus sanguineus* trypsin inhibitors present in the tick larvae: isolation, characterization, and partial primary structure determination. *Arch Biochem Biophys.* 2003; 417: 176-182.

25. Bachmair, A., Finley, D., and Varshavsky, A. In vivo half-life of a protein is a function of its amino-terminal residue. *Science*. 1986; 234: 179-186.

26. Bajzar L, Manuel R and Nesheim ME. Purification and characterization of TAFI, a thrombin activable fibrinolysis inhibitor. *J Biol Chem*, 1995; 270: 14477-14484.

27. Barrett, A.J. Cathepsin D: the lysosomal aspartic proteinase. *Ciba Found Symp*. 1979; 75: 37-50.

28. Barrett, A.J. Cystatin, the egg white inhibitor of cysteine proteinases. *Methods Enzymol*. 1981; 80: 771–778.

29. Barrett, A.J., Kembhavi, A.A., Brown, M.A., Kirschke, H., Knight, C.G., Tamai, M., and Hanada, K. L-trans-Epoxysuccinyl-leucylamido(4-guanidino)butane (E-64) and its analogues as inhibitors of cysteine proteinases including cathepsins B, H and L. *Biochem J*. 1982; 201(1): 189-98.

30. Barrett, A. J., N. Rawlings, M. Davies, W. Machleidt, G.Salvesen, and V. Turk: Cysteine Proteinase Inhibitors of the Cystatin Superfamily. In: *Proteinase Inhibitors*. 1986; pp. 515-569. (A. Barrett and G. Salvesen, Eds.) Elsevier, Amsterdam.

31. Barrett, A.J. Classification of peptidases. *Methods Enzymol*. 1994. 244: 1–15.

32. Barrett, A.J., Rawlings, N.D., and Woessner, J.F. *Handbook of Proteolytic Enzymes.* 1998. New York Academic Press (CD_ROM)

33. Barrett, A.J.; Rawlings, N.D. and Woessner, J.F. eds "*Handbook of Proteolytic Enzymes"* (2004) (2nd Ed.), Elsevier Academic Press, London, UK.

34. Bayés, A., Comellas-Bigler, M., Rodríguez de la Vega, M., Maskos, K., Bode, W., Avilés, F.X., Jongsma, M.A., Beekwilder, J., and Vendrell, J. Structural basis of the resistance of an insect carboxypeptidase to plant protease inhibitors. *Proc Natl Acad Sci USA*. 2005; 102(46): 16602-7.

35. Berti, P.J., and Storer, A.C. Alignment/phylogeny of the papain superfamily of cysteine proteases. J Mol Biol. 1995; 246(2): 273-83.

36. Bertonati, C.; Honig, B. and Alexov, E. Poisson-Boltzmann Calculations of Nonspecific Salt Effects on Protein-Protein Binding Free Energies. *Biophys J*. 2007; 92(6): 1891–1899.

37. Bertsch, A., Leinenbach, A., Pervukhin, A., Lubeck, M., Hartmer, R., Baessmann, C., *et al*. De novo peptide sequencing by tandem MS using complementary CID and electron transfer dissociation. *Electrophoresis*. 2009; 30(21): 3736-47.

38. Biedermann, K., Montali, U., Martin, B., Svendsen, I., and Ottesen, M. The amino acid sequence of proteinase A inhibitor 3 from baker's yeast. *Carlsberg Res. Commun*. 1980; 45: 225-235.

39. Bieth, J.G. Theoretical and practical aspects of proteinase inhibition kinetics. *Methods Enzymol*. 1995; 248: 59–84.

40. Birk, J. Plant protease inhibitors: Significance in nutrition, plant protection, cancer prevention, and genetic engineering. Published by Springer. 2003; 111-120.

41. Birnbaum, M.J., Clem, R.J., and Miller, L.K. An apoptosis-inhibiting gene from a nuclear polyhedrosis virus encoding a polypeptide with Cys/His sequence motifs. *J Virol*. 1994; 68(4): 2521-8.

42. Blanco, R.M., and Guisán, J.M. Stabilization of enzymes by multipoint covalent attachment to agarose-aldehyde gels. Borohydre reduction of trypsinagarose derivatives. *Enzyme Microb Technol.* 1998; 11, 360–366.

43. Blanco-Aparicio, C., Molina, M.A., Fernández-Salas, E., Frazier, M.L., Mas, J.M., Querol, E., Avilés, F.X., and De Llorens, R. Potato carboxypeptidase inhibitor, a T-knot protein, is an epidermal growth factor antagonist that inhibits tumor cell growth. *J Biol Chem*. 1998; 273(20): 12370-7.

44. Bode W and Huber R. Structural basis of the endoproteinase-protein inhibitor interaction. *Biochim Biophys* Acta. 2000; 1477: 241-252.

45. Boja, E.S., and Fales, H.M. Overalkylation of a Protein Digest with Iodoacetamide. *Anal. Chem*. 2001; 73: 3576-3582.

46. Bouma, B.N., and Mosnier, L.O. Thrombin activable fibrinolysis inhibitor (TAFI) at the interfase between coagulation and fibrinolysis. *Pathophysiol. Haemost. Thromb.* 2003; 33: 375-381.

47. Boyer, P.D. (1970). The enzymes. 3rd ed. 1970; pp 564. Ed., Academic Press, NY, USA.

48. Braddock, M.; Quinn, A. *Nature Reviews Drug Discovery*. 2004; 3: 1-10.

49. Brown, W.M., and Dziegielewska, K.M. Friends and relations of the cystatin superfamily-new members and their evolution *Protein Sci*. 1997; 6(1): 5-12.

50. Brunger, A.T. Version 1.2 of the Crystallography and NMR System, *Nature Protocols*. 2007; 2: 2728-2733.

51. Brzin, J., A. Ritonja, T. Popovic, and V. Turk: Low Molecular Mass Protein Inhibitor of Cysteine Proteinases from Soybean. *Biol. Chem. Hoppe-Seyler* 1990; 371(Suppl.):167-170.

52. Brusca R.C and Brusca G.J. *Invertebrates*. 2003; Sinauer Associates Inc. Ed. 2, Publishers Sunderland, Massashusetts 93.

53. Bruker Daltonics GmbH. Tutorials for Biotools version 3.2. October 2008.

54. Burrell, M.M. Enzymes of molecular biology. Vol. 16. 1993; pp 278-281. Ed., Humana Press, Totowa, NJ, USA.

55. Burnette, W.N. Western blotting: electrophoretic transfer of proteins from sodium dodecyl sulfate — polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. *Analytical Biochemistry*. 1981; 112 (2): 195-203. United States: Academic Press.

56. Campos, I.T.N., Tanaka-Azevedo, A.M. and Tanaka, A.S. Identification and characterization of a novel factor XIIa inhibitor in the hematophagous insect *Triatoma infestans* (*Hemiptera Reduviidae*). *FEBS Letters*. 2004; 577: 512–516

57. Carlsoon, J., Janson, J.C., and Sparrman, M. Affinity chromatography. In: *Protein purification*. 2da edicion. (Ed. Janson, J-C y Rydén, L). 1998; Wiley-VHC. p. 375.

58. Carroll, A. R., Pierens, G. K., Fechner, G., De Almeida Leone, P., Ngo, A., Simpson, M., Hyde, E., Hooper, J. N., Boström, S. L., Musil, D. *et al*. Dysinosin A: a novel inhibitor of Factor VIIa and thrombin from a new genus and species of Australian sponge of the family Dysideidae. *J. Am. Chem. Soc.* 2002; 124: 13340-13341

59. Carroll, A. R., Buchanan, M. S., Edser, A., Hyde, E., Simpson, M., and Quinn, R.J. Dysinosins B-D, inhibitors of factor VIIa and thrombin from the Australian sponge *Lamellodysidea chlorea*. *Journal of Natural Products.* 2004; 67: 1291-1294

60. Castañeda, O.; Sotolongo, V.; Amor, A. M.; Stocklin, R.; Anderson, A.; Harvey, A. L. and Karlsson, E. Characterization of a potassium channel blocker from Caribbean sea anemone *Stichodactyla helianthus*. *Toxicon*. 1995; 33: 603-613.

61. Cereghino J.L., and Cregg, J.M. Heterologous protein expression in the‐methylotrophic yeast *Pichia pastoris. FEMS Microbiol Rev*. 2001; 24: 45-66.

62. Cereghino, G.P., Cereghino, J.L., Ilgen, C., Cregg, J.M. Production of recombinant proteins in fermenter cultures of the yeast Pichia pastoris. *Curr Opin Biotechnol*. 2002; 13(4): 329-32.
63. Cereghino, J.L., Wong, W.W., Xiong, S., Giang, W., T. Luong, L.T., Vu, J., Johnson, S.D., and Cereghino, G.P. Condensed protocol for competent cell preparation and transformation of the methylotrophic yeast *Pichia pastoris*. *BioTechniques*. 2005; 38:44-48
64. Cereghino, J.L., and Cereghino, G.P. Vectors and strains for expression. *Methods Mol Biol*. 2007; 389: 11-26.
65. Chaurand, P., Luetzenkirchen, F., and Spengler, B. Peptide and Protein Identification by Matrix Assisted Laser Desorption Ionization (MALDI) and MALDI-Post-Source Decay Time-of-Flight Mass Spectrometry. J Am Soc Mass Spectrom. 1999; 10: 91–103.
66. Chávez, M.A., Díaz, J., Delfín, J., and Pérez, U. *Temas de Enzimología.* ENPES. 1990; La Habana, Cuba.
67. Chávez, M.A.; Gil, Sh.; Fernández, A.; Huerta, V.; Pascual, I.; Abreu, L.; Morera, V.; Saroyán, A.; Delfín, J.; Padrón, G.; Cisneros, M.; Joseph, P.; Charli, J. L. and Díaz, J. Purification and partial characterization of a proteinase inhibitor from sea anemone *Condylactis gigantea*. Toxicon. 1998; 36: 1275-1276.
68. Chávez, M.A., Pascual, I., García-Fernández, R., Delfín, J. and Díaz, J. Capítulo Inmovilización de Proteínas. In Nanotecnología. 2011; Ed. UH, Cuba (In press)
69. Chávez, M.A. and Gutiérrez, O. Cinética de inhibición de unión fuerte. Algunas consideraciones teóricas y prácticas. In: *Enzymología Biotecnológica*. 2012; Ed. ELFOS, La Habana, Cuba (In press).
70. Chávez-Gutiérrez, L., Bourdais, J., Aranda, G., Vargas, M.A., Matta-Camacho, E., Ducancel, F., Segovia, L., Joseph-Bravo, P. and Charli, J.L. (2005) A truncated isoform of pyroglutamyl aminopeptidase II produced by exon extension has dominant-negative activity. *J Neurochem*. 92(4): 807-17.
71. Chen Y.H., Yang Y.T., and Martinez H.H. Determination of the helix and beta form of proteins in aqueous solution by circular dichroism. *Biochemistry*. 1972; 13: 3350-3359.
72. Chen, C.C.; Wang SS, Chen TW, Jap TS, Chen SJ, Jeng FS, et al. Serum procarboxypeptidase B, amylase and lipase in chronic renal failure. *J. Gastroenterology Hepatol*. 1996; 11: 496-9.
73. Chen, Z. and Mayer, L.M. High concentrations of complexed metals in the guts of deposit feeders. *Limnol. Oceanogr.* 2000; 45(6): 1358–1367.
74. Cheng, C., and Gross, M.L. Applications and mechanisms of charge-remote fragmentation. *Mass Spec Rev*. 2000; 19: 398-420.
75. Christeller, J.T., Farley, P.C., Ramsay, R.J., Sullivan, P.A., and Laing, W.A. Purification, characterization and cloning of an aspartic proteinase inhibitor from squash phloem exudate. *Eur J Biochem*. 1998; 254(1): 160-7.
76. Christianson, D.W., and Lipscomb, W.N. *Acc Chem Re*. 1989; 22: 62-9.
77. Coll, M., Guasch, A., Avilés, F.X., and Huber, R. Three-dimensional structure of porcine procarboxypeptidase B: a structural basis of its inactivity. *EMBO J*. 1991; 10: 1-9.
78. Collins, D.W. and Jukes, T.H. Rates of transition and transversion in coding sequences since the human-rodent divergence. *Genomics*. 1994; 20(3): 386-96.
79. Cool, D.R., Normant, E., Shen, F., Chen, H.C., Pannell, L., Zhang, Y., and Loh, Y.P. Carboxypeptidase E is a regulated secretory pathway sorting receptor: Genetic obliteration leads to endocrine disorders in Cpefat mice. *Cell*. 1997; 88: 73-83.
80. Coombs, G.H., Goldberg, D.E., Klemba, M., Berry, C., Kay, J., Mottram, J.C. Aspartic proteases of *Plasmodium falciparum* and other parasitic protozoa as drug targets. *Trends in Parasitology*. 2001; 17: 532-7.
81. Cooks, R. G.; Ast, T.; Beynon, J. H. Anomalous Metastable Peaks. *Int. J. Mass Spectrom. Ion Phys.* 1975; 16: 348–352.
82. Cooper, H.J., Hakansson, K., and Marshall, A.G. *Mass Spectrom. Rev*. 2005; 24: 201–222.
83. Copeland R.A; Lombardo D.; Giannaras J. and Decicco C**.** Estimating Ki values for tight binding inhibitors from dose response plots. *Bioorganic and Medical Chemistry Letters*. 1995; 5(17): 1947-1952.
84. Copeland, R.A. (2000) *Enzymes: A practical introduction to structure, mechanism and data analysis.* Wiley-VCH, Inc.

85. Coughlin, S.R. Thrombin signalling and protease-activated receptors. *Nature*. 2000; 407: 258-264.
86. Coussens, L.M., and Werb, Z. Matrix metalloproteinases and the development of cancer. *Chem Biol*. 1996; 3(11): 895-904.
87. Cregg, J.M., and Madden, K.R. Development of the methylotrophic yeast, *Pichia pastoris*, as a host system for the production of foreign proteins. *Dev Ind Microbiol*. 1988; 29: 33-41.
88. Cregg, J.M., Cereghino J.L, Shi, J., Higgins, D.R. Recombinant protein expression in *Pichia pastoris. Mol Biotech*. 2000; 16: 23-52.
89. Crook, N.E., Clem, R.J., and Miller, L.K. An apoptosis-inhibiting baculovirus gene with a zinc finger-like motif. *J Virol*. 1993; 67(4): 2168-74.
90. Cui, W., Rohrs, H.W., and Gross, M.L. Top-down mass spectrometry: recent developments, applications and perspectives. *Analyst*. 2011; 136(19): 3854-64.
91. Curless, C., Baclaski, J., and Sachdev, R. Phosphate glass as a phosphate source in high cell density *Escherichia coli* fermentations. *Biotechnol. Prog.* 1996; 12: 22-25.
92. Davies, D.R. The structure and function of the aspartic proteinases. *Annu. Rev. Biophys. Biophys. Chem*. 1990; 19: 189–215.
93. Damdinsuren Boldbaatar, Chummy Sikalizyo Sikasunge, Badgar Battsetseg, Xuenan Xuan, Kozo Fujisaki. Molecular cloning and functional characterization of an aspartic protease from the hard tick *Haemaphysalis longicornis*. *Insect Biochemistry and Molecular Biology*. 2006; 36(1): 25-36.
94. Deddish, P.A., Skidgel, R.A., Kriho, V.B., Li, X.Y., Becker, R.P., and Erdös, E.G. Carboxypeptidase M in Madin-Darby canine kidney cells. Evidence that carboxypeptidase M has a phosphatidylinositol glycan anchor. *J Biol Chem*. 1990; 265(25): 15083-9.
95. Delano, W. L. PyMOL, Version 0_99rc6, DeLano Scientific. 2002; San Carlos (CA), USA.
96. Delfín, J.; González, Y.; Díaz, J. and Chávez, M. Proteinase inhibitor from Stichodactyla helianthus: Purification, characterization and inmobilization. *Arch. of Med. Res*. 1994; 25: 199-204.
97. Delfín, J.; Martínez, I.; Antuch, W.; Morera, V.; González, Y.; Rodríguez, R.; Márquez, M.; Saroyán, A.; Larionova, N.; Díaz, J.; Padrón, G. and Chávez, M. Purification, characterization and immobilization of proteinase inhibitors from *Stichodactyla helianthus. Toxicon*. 1996; 34 (11/12): 1367-1376.
98. Deveraux, Q.L., and Reed, J.C. IAP family proteins-suppressors of apoptosis. Genes Dev. 1999; 13: 239– 252.
99. Díaz J., Morera V., Delfin J., Huerta V., Lima G., Rodriguez de la Vega M., Garcia B., Padron G., Assfalg-Machleidt I., Machleidt W., Chavez M**.** Purification and partial characterization of a novel proteinase inhibitor from the sea anemone Stichodactyla helianthus**.** Toxicon. 1998; 36:1275-1276.
100. Dickinson, D.P. Cysteine peptidases of mammals: their biological roles and potential effects in the oral cavity and other tissues in health and disease. *Crit Rev Oral Biol Med*. 2002; 13(3): 238-75.
101. Dongre, A.R., Jones, J.L., Somogyi, A. and Wysocki, V.H. *J. Am. Chem. Soc*. 1996; 118: 8365–8374.
102. Doran, P. (1995). *Bioprocess Engineering Principles*. Academic Press, London, 297-332.
103. Dubin, G.; Stec-Niemczyk, J.; Dylag, T.; Silberring, J.; Dubin, A. and Potempa, J. Characterisation of a highly specific, endogenous inhibitor of cysteine protease from Staphylococcus epidermidis, a new member of the staphostatin family. *Biological Chemistry* . 2004; 385(6): 543.
104. Dubin G. Proteinaceous cysteine protease inhibitors. *Cell Mol Life Sci*. 2005; 62(6): 653-69.
105. Dubois, K.N.; Abodeely, M.; Sajid, M.; Engel, J.C. and McKerrow, J.H. *Giardia lamblia* cysteine proteases. Parasitol Res. 2006; 99: 313–316.
106. Dunbar, R.C. BIRD (blackbody infrared radiative dissociation): evolution, principles, and applications. *Mass Spec Rev*. 2004; 23: 127–158.
107. Dunn, B.M. Structure and mechanism of the pepsin-like family of aspartic peptidases. *Chem. Rev*. 2002; 102: 4431–4458.
108. Dunn, B.M., and Rao, M.B. Human immunodeficiency virus 1 retropepsin: Handbook of Proteolytic Enzymes. 2004. Elsevier (London), Barrett A. J, Rawlings N. D, Woessner J.F (Editors), p.144-154.

109. Eckart, K., Schwarz, H., Tomer, K. B., and Gross, M. L. *J. Am. Chem. Soc*. 1985; 107: 6765–6769.
110. Edman P. A method for the determination of the amino acid sequence in peptides. *Arch. Biochem*. 1949; 22: 475–476.
111. Edman P., and Begg G.S. A protein sequenator. *Eur. J. Biochem*. 1967, 1: 80–91.
112. Eaton D.L., Mallory B.E., Tsai S.P., Henzel W., and Drayna D. Isolation, molecular cloning, and partial characterization of a novel carboxypeptidase B from human plasma. *J Biol Chem*. 1991; 266: 21833-21838.
113. Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. Features and Development of Coot. *Acta Cryst D*. 2010; 66: 486-501.
114. Eng, F.J., Novikova, E.G., Kuroki, K., Ganem, D., and Fricker, L.D. gp180, a protein that binds duck hepatitis B virus particles, has metallocarboxypeptidase D-like enzymatic activity. *J Biol Chem*. 1998; 273(14): 8382-8.
115. Eriksson, H.; Wahlander, K.; Gustafsson, D.; Welin, L. T.; Frison, L.; Schulman, S.; Investigators, T. *Journal of Thrombosis and Haemostasis*, 2003; 1: 41-47.
116. Erlanger, B.F., Kokowsky, N., and Cohen, W. The preparation and properties of two new chromogenic substrates of trypsin. *Arch. Biochem. Biophys*. 1961; 95, 271-278.
117. Ettari, R., Bova, F., Zappala, M., Grasso, S., and Micale, N. Falcipain-2 inhibitors. *Medicinal research reviews*. 2009. Wiley Periodicals Inc.
118. Falick, A.M., Hines, W.M., Medzihradszky, K.F. Baldwin, M.A. Gibson, B.W. *J. Am. Soc. Mass Spectrom*. 1993; 4: 882–893.
119. Fernández-Lafuente, R., Cowan, D. A., and Wood, N. P. Hyperstabilization of a thermophilic esterase by multipoint covalent attachment. *Enzyme Microb. Technol*. 1995; 17: 366-372.
120. Fernández, D., Testero, S., Vendrell, J., Avilés, F.X., and Mobashery, S. The X-ray structure of carboxypeptidase A inhibited by a thiirane mechanism-based inhibitor. *Chem Biol Drug Des*. 2010; 75(1): 29-34.
121. Fife, T.H., and Przystas, T.J. *J Am Chem Soc*. 1986; 108: 4631-6.
122. Filippova, I. Yu., Lysogorskaya, E. N., Oksenoit, E. S., Rudenskaya, G. N., and Stepanov, V. M. *Anal. Biochem*. 1984; 143: 293–297.
123. Folk J. E. (1971) Carboxypeptidase B. In: The Enzymes, pp. 57–79, Boyer P. D. (ed.), Academic Press, New York.
124. Fontenele-Neto, J.D., Kalinina, E., Feng, Y., and Fricker, L.D. Identification and distribution of mouse carboxypeptidase A-6. *Mol Brain Res*. 2005; 137: 132-42.
125. Fossum, K. J.R. and Whitaker. Ficin and papain inhibitor from chicken egg white. *Arch Biochem. Biophys*, 1968; 125: 367-375.
126. Francis, S.E., Sullivan, D.J.Jr., and Goldberg, D.E. Hemoglobin metabolism in the malaria parasite Plasmodium falciparum. *Annu Rev Microbiol*. 1997; 51:97-123.
127. Fricker, L.D. In Barrett, A.J. Rawlings, N.D., and, Woessner, J.F. Handbook of Proteolytic Enzymes. 1998; pp 1341-4. Academic Press, San Diego (CA), USA.
128. Fricker, L.D., and Leiter, E.H. Peptides, enzymes and obesity: new insights from a 'dead' enzyme. *Trends Biochem Sci*. 1999; 24(10): 390-3.
129. Fritz, H.; Brey, B. and Béress, L. Polyvalent isoinhibitoren für trypsin, chymotrypsin, plasmin und kallikreine aus seeanemonen (Anemonia sulcata). Isolierung, Hemmverhalten und aminosäurezusammensetzung. Hoppe-Seyler's Z. *Physiol. Chem*. 1972; 353: 19-30.
130. Friedrich, T., Kroeger, B., Bialojan, S., Lemaire, H.G., Hoeffken, H.W., Reuschenbach, P., Otte, M. and Dodt, J. (1993) A Kazal-type inhibitor with thrombin specificity from *Rhodnius prolixus*. *J. Biol. Chem*. 268: 16216-16222
131. Fritz H. and Wunderer G. Biochemistry and application of aprotinin the kallikrein inhibitor from bovine organs, *Arzneimittelforschung*.1983; 33: 479-479.
132. Frohman, M.A., Dush, M.K. and Martin, G.R. Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc Natl Acad Sci USA*. 1988; 85(23): 8998-9002.
133. Fujinaga,M., Chernaia,M.M., Tarasova,N.I., Mosimann,S.C. & James,M.N. Crystal structure of human pepsin and its complex with pepstatin. *Protein Sci*. 1995; 4: 960-972.
134. Fujita, M., Nakao, Y., Matsunaga, S., Seiki, M., Itoh, Y., Van Soest, R.W.M., and Fusetani, N. Ancorinosides B–D, inhibitors of membrane type 1 matrix metalloproteinase (MT1-MMP), from the marine sponge Penares sollasi Thiele. *Tetrahedron*. 2001; 57(7): 1229-1234.

135. Fujita, M., Nakao, Y., Matsunaga, S., Van Soest, R.W., Itoh, Y., Seiki, M., and Fusetani, N. Callysponginol sulfate A, an MT1-MMP inhibitor isolated from the marine sponge Callyspongia truncata. *J Nat Prod*. 2003; 66(4): 569-71.

136. Fukal, L.; Kasafirek, E.; Strejcek, F. and Kas, J. (1987) selected *p*-nitroanilides as substrates for sensitive papain assay. *Journal of Food Biochemistry*. 1987; 11(2): 99–107

137. Fusetani, N., Matsunaga, S., Matsumoto, H., and Takebayashi, H. Cyclotheonamides, Potent Thrombin Inhibitors, from a Marine Sponge Theonella sp. *J. Am. Chem. Soc.* 1990; 112: 7053-7054.

138. Fusetani, N., Fujita, M., Nakao, Y., Matsunaga, S., and Van Soest, R. W. Tokaramide A, a new cathepsin B inhibitor from the marine sponge Theonella aff. mirabilis. *Bioorg. Med. Chem. Lett*. 1999; 9: 3397-3402.

139. García-Fernández, R., Redecke, L., Pons, T., Perbandt, M., Talavera, A., Gil, D., Gonzalez, Y., de los Angeles Chavez, M.A., and Betzel, C. Structural basis for serine protease inhibition by the recombinant Kunitz type inhibitor-1 from the caribbean sea anemone Stichodactyla helianthus: crystal structures of free and trypsin-complexed inhibitor. 2011. Submitted to the PDB data bank.

140. Gardell, S.J., Craik, C.S., Clauser, E., Goldsmith, E.J., Stewart, C.B., Graf, M. and Rutter, W.J. A novel rat carboxypeptidase, CPA2: characterization, molecular cloning, and evolutionary implications on substrate specificity in the carboxypeptidase gene family. *J Biol Chem*. 1988; 263(33): 17828-36.

141. Gardner, M.J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R.W., et al. Genome sequence of the human malaria parasite Plasmodium falciparum. *Nature*. 2002; 419: 498-511.

142. Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A. *Protein Identification and Analysis Tools on the ExPASy Server* (In) John M. Walker (ed): The Proteomics Protocols Handbook, Humana Press (2005).

143. Gettins P. G. Serpin structure, mechanism, and function. Chem. Rev. 2002; 102: 4751–4804.

144. Gill, S.C., and Von Hippel, P.H. Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* 1989; 182: 319-326.

145. Gil, L.A, Valiente, P.A., Pascutti, P.G., and Pons, T. Computational perspectives into plasmepsins structure-function relationship: implications to inhibitors design. *J Trop Med*. 2011; 2011:657483.

146. Goldberg, A.F. Slater, R. Beavis, B. Chait, A. Cerami and G.B. Henderson, Hemoglobin degradation in the human malaria pathogen *Plasmodium falciparum*: a catabolic pathway initiated by a specific aspartic protease. *Exp. Med.* 1991; 173: 961–969.

147. Gomis-Rüth, F.X., Companys, V., Qian, Y., Fricker, L.D., Vendrell, J., Avilés, F.X., and Coll, M. Crystal structure of avian carboxypeptidase D domain II: a prototype for the regulatory metallocarboxypeptidase subfamily. *EMBO J*. 1999; 18(21): 5817-26.

148. Gonda, D.K., Bachamair, A., Wunning, I., Tobias, J.W., Lane, W.S., and Varshavsky, A. Universality and structure of the N-end rule. *J. Biol. Chem.* 1989; 264: 16700-16712.

149. González-González, Y.; Alonso del Rivero, M.; García, B.; Hernández-Zardui, A.; Araujo, M. S.; Chávez, M.; Oliva, M. L. V. and Sampaio, C.A.M. Purification and preliminary characterization of a plasma kallikrein inhibidor isolated from sea hares Aplysia dactylomela. 2001; Revista Cubana de Química. XIII (2) (A-141).

150. González, C., Neira, J.L., Ventura, S., Bronsoms, S., Rico, M., Avilés, F.X. Structure and dynamics of the potato carboxypeptidase inhibitor by 1H and 15N NMR. *Proteins*. 2003; 50(3): 410-22.

151. González, Y., Araujo, M.S., Oliva, M.L.V., Sampaio, C.A.M., Chávez, M.A. Purification and preliminary characterization of a plasma kallikrein inhibitor isolated from sea hares *Aplysia dactylomela* Rang, 1828. *Toxicon*. 2004; 43: 219-223.

152. González, Y., Tanaka, Y. S., Hirata, I. Y., Alonso del Rivero, M., Oliva, M. L. V., Araujo, M. S. and Chavez, M.A. Purification and partial preliminar characterization of human neutrophile elastase from the marine snail *Cenchritis muricatus* (mollusca). *Comp Biochem Physiol A Mol Integr Physiol*. 2007; 146: 506-513 **(a)**.

153. González, Y., Pons, T., Gil, J., Besada, V., Alonso-del-Rivero, M., Tanaka, A. S., Araujo, M. S. and Chávez, M.A. Characterization and comparative 3D modeling of CmPI-II, a novel

'non-classical' Kazal-type inhibitor from the marine snail Cenchritis muricatus (Mollusca). *Biological Chemistry*. 2007; 388: 1183–1194 **(b)**.

154. Grathwohl, P. Difussion in natural porous media: contaminant transport, sorption/desorption and dissolution kinetics. 1958. Press in Kluwer Academic Publisher, Norwell, MA, USA.

155. Grazu, V., Betancor, L., Montes, T., López-Gallego, F., Guisán, J.M., and Fernández-Lafuente, R. Glyoxyl agarose as a new chromatographic matrix. *Enzyme Microb Technol.* 2006; 38, 960–966.

156. Groneberg, R. D.; Burns, C. J.; Morrissette, M. M.; Ullrich, J. W.; Morris, R. L.; Darnbrough, S.; Djuric, S. W.; Condon, S. M.; McGeehan, G. M.; Labaudiniere, R.; Neuenschwander, K.; Scotese, A. C.; Kline, J. A. *J. Med. Chem.*. 1999; 42: 541-4.

157. Grubb, A. Cystatin C — properties and use as diagnostic marker. Adv. Clin. Chem. 2000. 35: 63–99.

158. Guasch, C.A.; Coll, M.; Aviles, F.X. and Huber, R. Three-dimensional structure of porcine pancreatic procarboxypeptidase A. A comparison of the A and B zymogens and their determinants for inhibition and activation. *J. Mol. Biol*. 1992; 224: 141–157.

159. Guerra, Y., Ramírez, A. Reytor, M.L., Robles, R., Berry, C., Mendiola, J., Hernández-Zanui, A., and Chavez, M.A. Natural inhibitor of plasmepsin II from the gorgonian Plexaura homomalla: partial purification and characterization. *FEBS Journal*. 2008; 275 (S1): 450.

160. Guin, J.A. and Xiaofeng, Y. Hindered diffusion of asphaltenes at evaluated temperature and pressure. 1997; Chemical Engineering Department, Auburn University, Auburn AL, USA.

161. Gunasekera, S.P., McCarthy, P.J., Longley, R.E., Pomponi, S.A., Wright, A.E., Lobkovsky, E., and Clardy, J. Discorhabdin P, a New Enzyme Inhibitor from a Deep-Water Caribbean Sponge of the Genus Batzella. *J. Nat. Prod*. 1999; 62: 173-175 **(a)**.

162. Gunasekera, S.P., McCarthy, P.J., Longley, R.E., Pomponi, S.A., and Wright, A.E. Secobatzellines A and B, two new enzyme inhibitors from a deep-water Caribbean sponge of the genus Batzella. *J Nat Prod*. 1999; 62(8): 1208-11 **(b)**.

163. Guisán, J. M. Agarose- aldehyde gels as supports for immobilization- stabilization of enzymes. *Enzyme Microb. Technol.* 1988; 10, 375- 382.

164. Guisán, J. M., Fernández- Lafuente, V., Rodriguez, A., Bastida, R., Blanco, M., and Alvaro, G. Enzyme stabilization by multipoint covalent attachment to activated pre- existing, supports. In: Stability and stabilization of enzymes. (van den Tweel, W. J. J.; Harder, A. and Buitelaar, R. M. Eds.). *Proceedings of an International Symposium held in Maastricht*. 1992; 55- 62. The Netherlands.

165. Guruprasad, K., Reddy, B.V.B., and Pandit, M.W. Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. *Protein Eng.* 1990; 4: 155-161

166. Gustafson, G.L., Finn, D.J., and Moin, K. Dissociation of proteinase- inhibitor complexes by trichloroacetate. *Anal. Biochem*. 1988; 169: 185- 188.

167. Haiyan, G.; Jinlin, Z.; Min, L.; Takeshi, H.; Thasaneeya, H.; Rika, U.; Noboru I.; Xuenan, X. and Kozo, F. Characterization of a carboxypeptidase inhibitor from the tick Hemaphysalis longicornis. *Journal of Insect Physiology*. 2007; 53: 1079–1087.

168. Hamuro, Y., Coales, S.J., Molnar, K.S., Tuske, S.J. and Morrow, J.A. Specificity of immobilized porcine pepsin in H/D exchange compatible conditions. *Rapid Commun Mass Spectrom*. 2008; 22(7): 1041-6.

169. Hass, G.M.; Nau. H.; Biemann, K.; Grahn, D.T.; Ericsson, L.H. and Neurath, H. The amino acid sequence of a carboxypeptidase inhibitor from potatoes. *Biochemistry*. 1975; 14: 1334-42.

170. Hass, M. G., Derr, J. E. and Makus, D. J. Distribution of carboxypeptidase isoinhibitors in the potato plant. *Plant. Physiol*. 1979; 64: 1029- 1031.

171. Hass, G.M. and Hermodson, M.A. Amino acid sequence of a carboxypeptidase inhibitor from tomato fruit. *Biochemistry*. 1981; 20: 2256-60.

172. Hardouin, J. Protein sequence information by matrix-assisted laser desorption/ionization in-source decay mass spectrometry. Mass Spec Rev. 2007; 26: 672–682.

173. Harrison, M.J., Burton, N.A., and Hillier, I.H. Catalytic mechanism of the enzyme papain: predictions with a hybrid quantum mechanical/molecular mechanical potential. *J Am Chem Soc*. 1997; 119: 12285–91.

174. Hartley, B.S. Proteolytic enzymes. *Annu. Rev. Biochem*. 1960. 9: 45–72.

175. Hayashi, M.; Hamada, A.; Okaya, Y.; Wakitani, K.; Aisaka, K. *European Journal of Pharmacology*. 2001; 428: 163-168.
176. Hedstrom, L. Introduction: Proteases, Chem. Rev. 2002; 102(12): 4429-4430.
177. Hedstrom, L. Serine Protease Mechanism and Specificity. *Chem. Rev*. 2002; 102: 4501-4523.
178. Hendriks, D.F. Carboxypeptidase U. In: Barrett A.J., Rawlings, N.D., Woesner, J.F. Jr. eds. *Handbook of proteolytic enzymes*, 2nd edn. London: Elsevier. 2004; 825-828.
179. Henskens, M.C.; Veerman, E.C.I. and Amerongen, A.V.N. Cystatins in health and disease. *Biol. Chem*. 1996. 377: 71–86.
180. Higashiyama, S., Ohkubo, I.; Ishiguro, H.; Kunimatsu, M.; Sawaki,K. and Sasaki M. Human high molecular weight kininogen as a thiol proteinase inhibitor: presence of the entire inhibition capacity in the native form of heavy chain. Biochemistry. 1986; 25: 1669-75.
181. Higgins, D.R., Busser, K., Comiskey, J., Whitier, P.S., Purcell, T.J., and Hoeffler, J.P. Small vectors for protein expression in *Pichia pastoris* based on dominant drug resistante with direct multicopy selection and tags for heterologous protein detection and purification. *Pichia* Protocols. *Methods in Molecular Biology*. 1998; 103: 41-53.
182. Homandberg, G.A.; Litwiller, R.D. and Peanasky, R.J. Carboxypeptidase inhibitors from *Ascaris suum*: the primary structure. *Arch Biochem Biophys*. 1989; 270: 153-61.
183. Hourdou, M.L., Guinand, M., Vacheron, M. J., Michel, G., Denoroy, L., Duez, C., Englebert, S., Joris, B., Weber, G., and Ghuysen, J. M. Characterization of the sporulationrelated gamma-d-glutamyl-(l)meso-diaminopimelic-acid-hydrolysing peptidase I of Bacillus sphaericus NCTC 9602 as a member of the metallo(zinc) carboxypeptidase A family. Modular design of the protein. *Biochem. J.* 1993; 292: 563–570.
184. Hu, J.F., Schetz, J.A., Kelly, M., Peng, J.N., Ang, K.K., Flotow, H., Leong, C.Y., Ng, S.B., Buss, A.D., Wilkins, S.P., and Hamann, M.T. New antiinfective and human 5-HT2 receptor binding natural and semisynthetic compounds from the Jamaican sponge Smenospongia aurea. *J Nat Prod*. 2002; 65(4): 476-80.
185. Huang, H., Reed, C.P., Zhang, J.S., Shridhar, V., Wang, L., and Smith, D.I. Carboxypeptidase A3 (CPA3): a novel gene highly induced by histone deacetylase inhibitors during differentiation of prostate epithelial cancer cells. *Cancer Res*. 1999; 59, 2981–2988
186. Huber, R., Kukla, D., Ruehlmann, A., Epp, O., and Formanek, H. The Basic Trypsin Inhibitor of Bovine Pancreas. I. Structure Analysis and Conformation of the Polypeptide Chain. Naturwissenschaften. 1970; 57: 389-392.
187. Huerta, V., Morera, V., López, N., Betancourt, L., Besada, V., Padrón, G., Lima, G., Chávez, M.A., Delfín, J., and Díaz, J. Characterization and 3D model of a new proteinase inhibitor isolated from *Stichodactyla helianthus*. *Biotecnología Aplicada*. 1998; 15: 108-110.
188. Hugli, T. E. Protease inhibitors: novel therapeutic application and development. *Trends Biotechnol.* 1996; 14: 409-12.
189. Hunter, A.M., LaCasse, E.C., and Korneluk, R.G. The inhibitors of apoptosis (IAPs) as cancer targets. *Apoptosis*. 2007; 12(9): 1543-68.
190. Huntington, J.A., Read, R.J., and Carrell, R. W. Structure of a serpin-protease complex shows inhibition by deformation. Nature. 2000; 407: 923–926.
191. Ikai, A.J. Thermostability and aliphatic index of globular proteins. *J. Biochem.* 1980; 88: 1895-1898.
192. Ikegami, S., Kobayashi, H., Myotoishi, Y., Ohta, S. and Kato K. H. Selective inhibition of exoplasmic membrane fusion in echinoderm gametes with jaspisin, a novel antihatching substance isolated from a marine sponge. J. Biol. Chem. 1994; 269(37): 23262-7.
193. Illanes, A., González, J.M., Gómez, J.M., Valencia, P. and Wilson L. Diffusional restrictions in glyoxyl-agarose immobilized penicillin G acylase of different particle size and protein loading. *Electron. J. Biotech*. 2010; 13(1), Valparaiso (Chile)
194. Invitrogen Corp. (2000). *Pichia* fermentation process guidelines. *http://www.invitrogen.com/.* Invitrogen Corp., USA.
195. Ishikawa, T., Murakami, K., Kido, Y., Ohnishi, S., Yazaki, Y., Harada, F., and Kuroki, K. Cloning, functional expression, and chromosomal localization of the human and mouse gp180-carboxypeptidase D-like enzyme. *Gene*. 1998; 215(2): 361-70.
196. Janin, J., and Chothia, C. The structure of protein-protein recognition sites. *J. Biol. Chem*. 1990; 265: 16027–16030.

197. Janson, J.C., and Rydén, L. Protein Purification principles, high resolution methods and applications. 2nd. Edition Jonh Wiley and Sons, Inc. 1998; pp.3-40.
198. Jenko, K.S., Guncar, G., Stern, I., Morgan, G., Rabzelj, S., Kenig, M., Staniforth, R.A., Waltho, J.P., Zerovnik, E., and Turk, D. Essential role of proline isomerization in stefin B tetramer formation. *J Mol Biol.* 2007; 366(5): 1569-79.
199. Jennings, K.R. Int. J. Mass Spectrom. Ion Phys., 1968, 1, 227–235.
200. Johansson, M.W., Keyser, P. and Soderhall, K. Purification and cDNA cloning of a four-domain Kazal proteinase inhibitor from crayfish blood cells. *Eur. J. Biochem*. 1994; 223: 389-394
201. Johnson, R.S., Martin, S.A., Biemann, K., Stults, J.T., and Watson, J.T. Novel fragmentation process of peptides by collision-induced decomposition in a tandem mass spectrometer: differentiation of leucine and isoleucine. *Anal Chem*. 1987; 59(21): 2621-5.
202. Johnson, R.S., Martin, S.A., and Biemann, K. Int. J. Mass Spectrom. Ion Process. 1988; 86: 137–154.
203. Joint Center for Structural Genomics (JCSG). Crystal structure of Putative carboxypeptidase A (YP_562911.1) from *Shewanella denitrificans os-217* at 2.39 A resolution. Protein Data Bank (PDB); 2009.
204. Jones, A.W., and Cooper, H.J. Dissociation techniques in mass spectrometry-based proteomics. *Analyst*. 2011; 136; 3419–3429.
205. Juvvadi S.; Fan, X; Nagle G. T. and Fricker, D.L. (1997) *FEBS Letters*. 1997; 408: 195-200.
206. Kabsch, W. XDS. *Acta Cryst. D.* 2010; 66: 125-132.
207. Kageyama T. Molecular cloning, expression and characterization of an Ascaris inhibitor for pepsin and cathepsin E. *Eur J Biochem*. 1998; 253(3): 804-9.
208. Kalinina, E., Biswas, R., Berezniuk, I., Hermoso, A., Avilés, F.X., and Fricker, L.D. A novel subfamily of mouse cytosolic carboxypeptidases. *FASEB J*. 2007; 21(3): 836-50.
209. Kamphuis, I.G., Kalk, K.H., Swarte, M.B., and Drenth, J. Structure of papain refined at 1.65 A resolution. *J Mol Biol*. 1984; 179(2): 233-56.
210. Karas, M. and Bahr, U. Laser desorption Ionization Mass-Spectrometry of Large Biomolecules. *Trends in Anal. Chem*. 1990; 9: 321-325.
211. Kayashima, T., Yamasaki, K., Yamada, T., Sakai, H., Miwa, N., Ohta, T., Yoshiura, K., Matsumoto, N., Nakane, Y., Kanetake, H., Ishino, F., Niikawa, N., and Kishino, T. The novel imprinted carboxypeptidase A4 gene (CPA4) in the 7q32 imprinting domain. *Hum. Genet.* 2003; 112: 220–226
212. Keil, B. Specificity of proteolysis. 1992; Ed. Springer-Verlag, Berlin-Heidelberg-New York
213. Keil, C., Maskos, K., Than, M., Hoopes, J.T., Huber, R., Tan, F., Deddish, P.A., Erdös, E.G., Skidgel, R.A., and Bode, W. Crystal structure of the human carboxypeptidase N (kininase I) catalytic domain. *J Mol Biol*. 2007; 366(2): 504-16.
214. Keilova, H. and Tomasek, V. Further characteristics of cathepsin D inhibitor from potatoes. *Collect Czech Chem Commun*. 1976; 41**:** 2440–2447.
215. Kelleher, N.L. Top-down proteomics. Anal Chem. 2004; 76(11): 197A-203ª.
216. Kennedy, J.F., and Cabral, J.M.S. Immobilized enzymes. In *Solid Phase Biochemistry*. (W.H. Scouten, Ed.), 1983; p. 253. Wiley (Interscience), USA.
217. Kimmel, J.R. and Smith, E.L. Crystalline papain. I. Preparation, specificity, and activation. *J Biol Chem*. 1954; 207(2): 515-31
218. Kirschke, H.; Barrett, A.J. and Rawlings, N.D. Proteinases 1: Lysosomal cysteine proteinases; in Proteine Profile 2 (Sheterline, P., ed.) Oxford University Press. 1995. 1587–1643.
219. Kishimura, H. and Hayashi, K. *Comp. Biochem Physiol Biochem Mol Biol*. 2002; 133(2): 183-9.
220. Klenø, T.G., Andreasen, C.M., Kjeldal, H.Ø., Leonardsen, L.R., Krogh, T.N., Nielsen, P.F., Sørensen, M.V., and Jensen, O.N. MALDI MS peptide mapping performance by in-gel digestion on a probe with prestructured sample supports. *Anal Chem*. 2004; 76(13): 3576-83.
221. Knight, C.G. Active-site titration of peptidases. *Meth. Enzymol.* 1995; 248: 85-101
222. Koehn, F.E, and Carter, G.T. The evolving role of natural products in drug discovery. *Nature Reviews*. 2005; 4: 206-220

223. Koester, C., Holle, A. Proceedings of the 47th ASMS Conference on Mass Spectrometry and Allied Topics. Dallas, TX, June 13-17, 1999, MPA-003.
224. Kolkenbrock, H., and Tschesche, H. A new inhibitor of elastase from the sea anemone (*Anemonia sulcata*). *Biol. Chem. Hoppe Zeyler*. 1987; 368: 93-99.
225. Komiyama, T., Ray, C.A., Pickup, D.J., Howard, A.D., Thornberry, N.A., Peterson, E.P. *et al*. Inhibition of interleukin-1 beta converting enzyme by the cowpox virus serpin CrmA: an example of cross-class inhibition. J. Biol. Chem. 1994; 269: 19331–19337.
226. Krizkova, S., Zitka, O., Masarik, M., Adam, V., Stiborova, M., Eckschlager, T., Hubalek, J., and Kizek, R. Clinical importance of matrix metalloproteinases. *Bratisl Lek Listy*. 2011; 112(8): 435-40.
227. Krowarsch, D., Cierpicki, T., Jelen, F., and Otlewski, J. Canonical protein inhibitors of serine proteases. *Cell Mol Life Sci*. 2003; 60(11): 2427-44.
228. Kumazaki, T., Hoshiba, N., Yokosawa, H., and Ishii, S.I Primary structure of ascidian trypsin inhibitors in the hemolymph of a solitary ascidian *Halocynthia roretzi*. *J. Biochem*. 1990; 107 : 409-413.
229. Kuroki, K., Eng, F., Ishikawa, T., Turck, C., Harada, F., and Ganem, D. gp180, a host cell glycoprotein that binds duck hepatitis B virus particles, is encoded by a member of the carboxypeptidase gene family. *J Biol Chem*. 1995; 270(25): 15022-8.
230. Kyte, J., and Doolittle, R.F. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* 1982; 157: 105-132
231. La Barre, S., Longeon, A., Barthelemy, M., Guyot, M., Le Caer, J-P. and Bargibant, G. (1996) Characterization of a novel elastase inhibitor from a fan coral. *C. R. Acad. Sci.* 319: 365-370
232. Laber, B., Krieglstein, K., Henschen, A., Kos, J., Turk, V., Huber, R., and Bode, W. The cysteine proteinase inhibitor chicken cystatin is a phosphoprotein. *FEBS Lett.* 1989; 248(1-2): 162-8.
233. Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*. 1970; 227(5259): 680-5.
234. Laing, W. A., and McManus, M. In *Protein Interactions in Plants* (McManus, M. T., Laing, W. A., and Allan, A. C., eds). 2002; pp. 77–119, Sheffield Academic Press, Sheffield, UK.
235. Laskin, J., and Futrell, J.H. Activation of large ions in ft-icr mass spectrometry. *Mass Spec Rev*. 2005; 24: 135–167.
236. Lalmanach, G., Naudin, C., Lecaille, F., and Fritz, H. Kininogens: More than cysteine protease inhibitors and kinin precursors. *Biochimie*. 2010; 92(11): 1568-79.
237. Langer, G., Cohen, S.X., Lamzin, V.S., and Perrakis, A. Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nature Protocols.* 2008; 3: 1171-1179.
238. Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J. and Higgins, D.G. *Bioinformatics*. 2007; 23(21): 2947-2948.
239. Larsen, K. S. and Auld, D. S. Carboxypeptidase A: mechanism of zinc inhibition. *Biochem*. 1989; 28(25): 9620- 9625.
240. Lazdunski, M.; Vincent, J.P.; Schweitz, H.; Peron-Renner, M. and Pudles, J. The mechanism of association of trypsin (or chymotrypsin) with the pancreatic trypsin inhibitors (Kunitz and Kazal). Kinetics and thermodynamics of the interaction. In: H. Fritz, H. Tschesche and L.J. Greene, Editors, *Bayer Symposium V*, Springer Verlag, Berlin. 1974. 420–431.
241. Laskowski, M. Jr. and Kato, I. Protein inhibitors of proteinases. *Annu. Rev. Biochem*. 1980; 49: 593-626.
242. Laskowski, M.Jr., Kato, I., Ardelt, W., Cook, J., Denton, A., Empie, M.W., Kohr, W.J., Park, S.J., Parks, K., *et al*. Ovomucoid third domains from 100 avian species isolation, sequences, and hypervariability of enzyme-inhibitor contact residues. *Biochemistry*. 1987; 26: 202-221.
243. Lecaille, F.; Serveau, C.; Gauthier, F. and Lalmanach, G. Revisiting the S2 specificity of papain by structural analogs of Phe. *FEBS Letters*. 1999; 445: 311-314.
244. Lee, S.G., and Chmielewski, J. Rapid synthesis and in situ screening of potent HIV-1 protease dimerization inhibitors. *Chem Biol*. 2006; 13: 421-426.

245. Lenarčič, B., Ritonja, A., Strukelj, B., Turk, B., Turk, V. Equistatin, a new inhibitor of cysteine proteinases from Actinia equina, is structurally related to thyroglobulin type-1 domain. *J. Biol. Chem*. 1997; 272: 13899-13903.

246. Lenarčič, B., and Turk, V. Thyroglobulin Type-1 Domains in Equistatin Inhibit Both Papain-like Cysteine Proteinases and Cathepsin D. *The Journal of Biological Chemistry*. 1999; 274, 563–566.

247. Leung, D.; Abbenante, G. and Fairlie, D.P. Protease Inhibitors: Current Status and Future Prospects. *Journal of Medicinal Chemistry*. 2000; 3(43): 305-241

248. Li, J.W. and Vederas, J.C. Drug discovery and natural products: end of and era or an endless frontier?. *Science*. 2009; 325(5937):161-5

249. Li, M., Phylip, L.H., Lees, W.E., Winther, J.R., Dunn, B.M., Wlodawer, A., Kay, J., Gustchina, A. The aspartic proteinase from Saccharomyces cerevisiae folds its own inhibitor into a helix. *Nat Struct Biol*. 2000; 7(2): 113-7.

250. Li, X.Y., and Skidgel, R.A. Release of glycosylphosphatidylinositol-anchored carboxypeptidase M by phosphatidylinositol-specific phospholipase C upregulates enzyme synthesis. *Biochem Biophys Res Commun*. 1999; 258(1): 204-10.

251. Lin, K.; Gates, C. A.; Luong, Y. P.; Perni, R. B.; Kwong, A. D. *Hepatology*, 2003, *38*, 222a-222a.

252. Lippstreufisher, D.L., and Gross, M. L. *Anal. Chem*. 1985; 57: 1174–1180.

253. Little, D.P., Speir, J.P., Senko, M.W., O'Connor, P.B., and McLafferty, F.W. Infrared multiphoton dissociation of large multiply charged ions for biomolecule sequencing. *Anal Chem*. 1994; 66(18): 2809-15.

254. Liu, Q.; Yu, L.; Gao, J.; Fu, Q.; Zhang, J. and Zhang, P. Cloning, tissue expression pattern and genomic organization of latexin, a human homologue of rat carboxypeptidase A inhibitor. *Mol Biol. Rep*. 2000; 27: 241-6.

255. Ljunggren, A., Redzynia, I., Alvarez-Fernandez, M., Abrahamson, M., Mort, J.S., Krupa, J.C., Jaskolski, M., and Bujacz, G. Crystal structure of the parasite protease inhibitor chagasin in complex with a host target cysteine protease. *J Mol Biol*. 2007; 371(1): 137-53.

256. Loo, J.A., Edmonds, C.G., and Smith, R.D. Primary sequence information from intact proteins by electrospray ionization tandem mass spectrometry. *Science*. 1990; 248(4952): 201-4.

257. López-Otín, C., and Matrisian, L. M. Emerging roles of proteases in tumour supression. Nature Rev Cancer. (2007); 7(10), 800-808

258. Lorrain, J.; Millet, L.; Lechaire, I.; Lochot, S.; Ferrari, P.; Visconte, C.; Sainte-Marie, M.; Lunven, C.; Berry, C. N.; Schaeffer, P.; Herbert, J. M.; O'Connor, S. E. *Journal of Pharmacology and Experimental Therapeutics*. 2003; 304: 567-574.

259. Lu W, Zhang W, Molloy SS, Thomas G, Ryan K, Chiang Y, Anderson S, Laskowski M Jr. Arg15-Lys17-Arg18 turkey ovomucoid third domain inhibits human furin. *J. Biol. Chem*. 1993; 268: 14583– 14585.

260. Lundequist, A., Tchougounova, E., Abrink, M., and Pejler, G. Cooperation between mast cell carboxypeptidase A and the chymase mouse mast cell protease 4 in the formation and degradation of angiotensin II. *J Biol Chem*. 2004; 279: 32339-44.

261. Lyons, P.J., Fricker, L.D. Substrate specificity of human carboxypeptidase A6. *J Biol Chem*. 2010; 285(49): 38234-42.

262. Lyons, P.J., and Fricker, L.D. Carboxypeptidase O is a glycosylphosphatidylinositol-anchored intestinal peptidase with acidic amino acid specificity. *J Biol Chem*. 2011; Sep 15.

263. Lyublinskaya, L.A., Belyaev, S.V., Strongin, A.Y., Matyash, L.F., Levin, E.D., and Stepanov V.M. A new chromogenic substrate for subtilisin. *Anal. Biochem.* 1974; 62(2), 371-376

264. Mabud, M.D.A., Dekrey, M.J., Cooks, R.G. Surface-Induced Dissociation of Molecular-Ions. *Int. J. Mass Spectrom.* 1985; 67: 285–294.

265. Macauley-Patrick, S., Fazenda, M.L., McNeil, B., and Harvey, L.M. Heterologous protein production using the Pichia pastoris expression system. *Yeast*. 2005; 22(4): 249-70.

266. Mala, B.R.; Aparna, M.T.; Mohini, S.G. and Vasanti, V. Deshpande. Molecular and Biotechnological Aspects of Microbial Proteases. *Microbiology and Molecular Biology Reviews*. 1998; 597–635.

267. Mann, M.; Hendrickson, R.C. and Pandey, A. Analysis of proteins and proteomes by mass spectrometry. *Annu. Rev. Biochem*. 2001; 70: 437–473.

268. Mares, M., Meloun, B., Pavlik, M., Kostka, V., and Baudys, M. Primary structure of cathepsin D inhibitor from potatoes and its structure relationship to soybean trypsin inhibitor family. *FEBS Lett*. 1989; 251(1-2): 94-8.

269. Marino-Buslje, C., Venhudová, G., Molina, M.A., Oliva, B., Jorba, X., Canals, F., Avilés F.X., and Querol, E. Contribution of C-tail residues of potato carboxypeptidase inhibitor to the binding to carboxypeptidase A. A mutagenesis analysis. *Eur. J. Biochem*. 2000; 267: 1502-1509.

270. Martin, P., Raymond, M.N., Bricas, E., and Dumas, B.R. Kinetic studies on the action of *Mucor pusillus*, *Mucor miehei* acid proteases and chymosins A and B on a synthetic chromophoric hexapeptide. *Biochim. Biophys. Acta.* 1980; 612, 410-420.

271. Martzen, M.R., McMullen, B.A., Smith, N.E., Fujikawa, K., and Peanasky, R.J. Primary structure of the major pepsin inhibitor from the intestinal parasitic nematode Ascaris suum. *Biochemistry*. 1990; 29(32): 7366-72.

272. Marx, P.F. Thrombin-activable fibrinolysis inhibitor. *Curr. Med. Chem.* 2004; 11: 2335-2348.

273. Mateo, C., Abian, O., Bernedo, M., Cuenca, E., Fuentes, M., Fernández-Lorente, G., Palomo, J.M., Grazu, V., Pessela, B., Giacomini, C., Irazoqui, G., Villarino, A., Ovsejevi, K., Batista-Viera, F., Fernández-Lafuente, R., and Guisán, J.M. Some special features of glyoxyl supports to immobilize proteins. *Enzyme Microb Technol.* 2005; 37, 456–462

274. Mateo, C., Palomo, J.M., Fuentes, M., Betancor, L., Grazu, V., López-Gallego, F., Pessela, B., Hidalgo, A., Fernández-Lorente, G., Fernández-Lafuente, R., and Guisán, J.M. Glyoxyl agarose: A fully inert and hydrophilic support for immobilization and high stabilization of proteins. *Enzyme Microb Technol.* 2006; 39, 274–280.

275. Mathialagan, N., and Hansen, T.R. Pepsin-inhibitory activity of the uterine serpins. *Proc. Natl. Acad. Sci. USA*. 1996; 93: 13653–13658.

276. Matsunaga, S.; Kamimura, T. and Fusetani, N. Isolation of 1-carboxymethylnicotinic acid from the marine sponge anthosigmella cf. raromicrosclera As a cysteine protease inhibitor1. *J Nat Prod*. 1998; 61: 671-2.

277. Maynes, J.T., Cherney, M.M., Qasim, M.A., Laskowski, M.Jr, and James, M.N. Structure of the subtilisin Carlsberg-OMTKY3 complex reveals two different ovomucoid conformations. *Acta Crystallogr D Biol Crystallogr*. 2005; 61(5): 580-8.

278. McGrath, M.E. The lysosomal cysteine proteases. *Annu. Rev. Biophys. Biomol. Struct*. 1999; 28: 181–204.

279. McGwire, G.B., Becker, R.P., and Skidgel, R.A. Carboxypeptidase M, a glycosylphosphatidylinositol-anchored protein, is localized on both the apical and basolateral domains of polarized Madin-Darby canine kidney cells. *J Biol Chem*. 1999; 274(44): 31632-40.

280. McLafferty, F.W., and Bryce, T.A. *Chem. Commun*. 1967; 1215.

281. McLuckey, S.A., and Goeringer, D.E. *J. Mass Spectrom.* 1997; 32: 461–474.

282. Mebs, D.; Gebauer, E. Isolation of proteinase inhibitory, toxic and hemolytic polypeptides from a sea anemone, *Stoichactis sp*. *Toxicon*. 1980; 18: 97-106

283. Ménard, R., Carrière, J., Laflamme, P., Plouffe, C., Khouri, H.E., Vernet, T., Tessier, D.C., Thomas, D.Y., and Storer, A.C. Contribution of the glutamine 19 side chain to transition-state stabilization in the oxyanion hole of papain. *Biochemistry*. 1991; 30(37): 8924-8.

284. Ménard, R., Plouffe, C., Laflamme, P., Vernet, T., Tessier, D.C., Thomas, D.Y., and Storer, A.C. Modification of the electrostatic environment is tolerated in the oxyanion hole of the cysteine protease papain. *Biochemistry*. 1995; 34(2): 464-71.

285. Mende, K., Petoukhova, O, Koulitchkova, V., Schaub, G.A., Lange, U., Kaufmann, R. and Nowak, G. Dipetalogastin, a potent thrombin inhibitor from the blood-sucking insect *Dipetalogaster maximus.* cDNA cloning, expression and characterization. *Eur. J. Biochem.* 1999; 266: 583-590

286. Metz, M., Piliponsky, A.M., Chen, C.C., Lammel, V., Abrink, M., Pejler, G., Tsai, M., and Galli, S.J. Mast cells can enhance resistance to snake and honeybee venoms. *Science*. 2006; 313(5786): 526-30

287. Minagawa, S., Ishida, M., Shimakura, K., Nagashima, Y., and Shiomi, K. Isolation and amino acid sequences of two Kunitz-type protease inhibitors from the sea anemone *Anthopleura aff. xanthogrammica*. *Comp. Biochem. Physiol.* 1997; 118B: 381-386.

288. Missen, R.W., Mims, C.A. and Saville, B.A. *Introduction to chemical reaction engineering and kinetics*. Chap. 8. 1999; John Wiley & Sons, Inc. New York, USA.

289. Mock, W.L.; Liu, Y. and Stanford, D.J. Arazoformyl peptide surrogates as spectrophotometric kinetic assay substrates for carboxypeptidase A. Anal Biochem. 1996; 239(2): 218-22.

290. Mock, W.L. and Stanford, D.J. Anisylazoformylarginine: a superior assay substrate for carboxypeptidase B type enzymes. Bioorg. Med. Chem. Lett. 2002; 12(8): 1193-4.

291. Moneriz, C., Marín, P., Bautista, J., Diez, A. and Puyet, A. Haemoglobin interference and increased sensitivity of fluorimetric assays for quantification of low-parasitaemia Plasmodium infected erythrocytes. *Malaria Journal*. 2009; 8: 279.

292. Monteiro, A.C., Abrahamson M, Lima, A.P., Vannier-Santos, M.A., and Scharfstein, J. Identification, characterization and localization of chagasin, a tight-binding cysteine protease inhibitor in *Trypanosoma cruzi*. *J Cell Sci*. 2001; 114(21): 3933-42.

293. Montes, T., López-Gallego, F., Fuentes, M., Mateo, C., Grazu, V., Betancor, L., Guisan, J.M. and Fernandez-Lafuente, R. Improved stabilization of chemically aminated enzymes via multipoint covalent attachment on glyoxil supports. In Immobilization of enzymes and cells, 2nd ed., 2006; Humana Press Inc., New Jersey, USA.

294. Morávek, L., and Kostka, V. Complete amino acid sequence of hog pepsin. *FEBS Lett*. 1974; 43(2): 207-11.

295. Morrison, J.F. The slow-binding and slow, tight-binding inhibition of enzyme-catalyzed reactions. *TIBS*. 1982; 7: 102-105

296. Muller, C.A.; Appelros, S., Uhl W, Bucheler MW, Borgstrom A. Serum levels of procarboxypeptidase B and its activation peptide in patients with acute pancreatitis and non pancreatics. *Diseases.* 2002; 51: 229-35.

297. Nagle, G.T., Jong-Brink, M., Painter, S.D., Li, K.W. Structure, localization and potential role of a novel molluscan trypsin inhibitor in Lymnaea. *Eur. J. Biochem*. 2001; 268, 1213-1221.

298. Nakao, Y.; Fujita, M.; Warabi, K.; Matsunaga, S.; Fusetani, N. and Miraziridine, A. A Novel Cysteine Protease Inhibitor from the Marine Sponge *Theonella aff. mirabilis*. *J. Am. Chem. Soc*. 2000; 122: 10462-10463.

299. Nakao, Y., and Fusetani, N. Enzyme inhibitors from marine invertebrates. *Journal of Natural Products*. 2007; 70: 689-710.

300. Nalamachu, S. R.; Song, L. and Fricker, L. D. Regulation of carboxypeptidase E- effect of Ca^{2+} on enzyme activity and stability. *J. Biol. Chem*. 1994; 269: 11192- 11195.

301. Neuhoff, V.; Arold, N.; Taube, D. and Ehrhardt, W. Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using Coomassie Brilliant Blue G-250 and R-250. Electrophoresis. 1988; 9(6): 255-62.

302. Neurath, H. Proteolytic processing and physiological regulation. *Trends Biochem. Sci*. 1989; 14: 268–271.

303. Ng, K.K., Petersen, J.F., Cherney, M.M., Garen, C., Zalatoris, J.J., Rao-Naik, C., Dunn, B.M., Martzen, M.R., Peanasky, R.J., and James, M.N. Structural basis for the inhibition of porcine pepsin by *Ascaris* pepsin inhibitor-3. *Nat. Struct. Biol*. 2000; 7(8): 653–657.

304. Nirmala X.; Kodrik D.; Zurovec M and Sehnal Frantisek. Insect silk contains both a kunitz-type and a unique Kazal-type proteinase inhibitor *Eur J. Biochem,* 2001; 268: 2064-2073.

305. Normant, E.; Martres, M. P.; Schwartz, J. C. and Gros, C. Purification, cDNA cloning, functional expression and characterization of a 26-kDa endogenous mammalian carboxypeptidase inhibitor. *Proc. Natl. Acad. Sci*. 1995; 92: 12225-12229 **(a)**.

306. Normant, E., Gros, C., and Schwartz, J.C. Carboxypeptidase A isoforms produced by distinct genes or alternative splicing in brain and other extrapancreatic tissues. *J Biol Chem*. 1995; 270(35): 20543-9 **(b)**.

307. Novikova, E.G., Eng, F.J., Yan, L., Qian, Y., and Fricker, L.D. Characterization of the enzymatic properties of the first and second domains of metallocarboxypeptidase D. *J Biol Chem*. 1999; 274(41): 28887-92.

308. Novikova, E.G., Reznik, S.E., Varlamov, O., and Fricker, L.D. Carboxypeptidase Z is present in the regulated secretory pathway and extracellular matrix in cultured cells and in human tissues. *J Biol Chem*. 2000; 275(7): 4865-70.

309. Ødum, L., Bundgaard, J.R., and Johnsen, A.H. A Kazal-type trypsin inhibitor from the protochordate *Ciona intestinalis*. *Eur J Biochem*. 1999; 259: 872-876.

310. Oehler, R., Lesnicki, G., and Galleno M. High cell density fermentation of *Pichia pastoris* using nonphosphate precipitate forming sodium hexametaphosphate as a phosphate source. In: *Current topics in gene expression annual meeting*. 1998; San Diego (CA), USA.

311. Oh, H., Breuker, K., Sze, S.K., Ge, Y., Carpenter, B.K., and McLafferty, F.W. Secondary and tertiary structures of gaseous protein ions characterized by electron capture dissociation mass spectrometry and photofragment spectroscopy. *Proc Natl Acad Sci USA*. 2002; 99(25): 15863-8.

312. Otto, H. and Schirmeister, T. Cysteine proteases and their inhibitors. *Chem. Rev*. 1997. 97: 133–171.

313. Pallarès, I.; Bonet, R.; Garcia-Castellanos, R.; Ventura, S.; Aviles, F.X. and Vendrell, J. Structure of human carboxypeptidase A4 with its endogenous protein inhibitor, latexin. *Proc Natl Acad Sci USA*. 2005; 102: 3978-83.

314. Pallarès, I., Fernández, D., Comellas-Bigler, M., Fernández-Recio, J., Ventura, S., Avilés, F.X., Bode, W. and Vendrell J. (2008) Direct interaction between a human digestive protease and the mucoadhesive poly(acrylic acid). *Acta Crystallogr D Biol Crystallogr*. D64(7): 784-91.

315. Pan, S., Gu, S., Bradbury, E.M., and Chen, X. Single peptide-based protein identification in human proteome through MALDI-TOF MS coupled with amino acids coded mass tagging. *Anal Chem*. 2003; 75(6): 1316-24.

316. Pascual, R., Burgos, F.J., Salva, M., Soriano, F., Mendez, E. and Aviles, F.X. (1989) Purification and properties of five different forms of human procarboxypeptidases. *Eur J Biochem*. 179(3): 609-16.

317. Patel, M.; Kayani, I.S.; Mellor, G.W.; Sreedharan, S.; Templeton, W.; Thomas, E.W.; Thomas, M. and Brocklehurst, K. Variation in the P2-S2 stereochemical selectivity towards the enantiomeric N-acetylphenylalanylglycine 4-nitroanilides among the cysteine proteinases papain, ficin and actinidin. *Biochem J*. 1992; 281: 553–559.

318. Pascual, I.; Gil-Parrado Sh.; Cisneros, M.; Joseph-Bravo, P.; Diaz, J.; Possani, L.; Charli, J.L. and Chávez, M.A. Purification of a specific inhibitor of pyroglutamyl aminopeptidase II from the marine annelide *Hermodice carunculata. In vivo* effects in rodent brain. *Int. J. Biochem. Cell Biol*. 2004; 36: 138–152.

319. Pascual, I.; Lopéz, A.; Gómez, H.; Chappé, M.; Saroyán A.; González Y.; Cisneros, M.; Charli, J.L. and Chávez, M.A. Screening of inhibitors of porcine dipeptidyl peptidase IV activity in aqueous extracts from marine organisms. *Enzyme and microbial technology*. 2007; 40(3): 414-419.

320. Patil, A.D., Kokke, W.C., Cochran, S., Francis, T.A., Tomszek, T., and Westley, J.W. Brominated polyacetylenic acids from the marine sponge Xestospongia muta: inhibitors of HIV protease. *J Nat Prod*. 1992; 55(9): 1170-7.

321. Pedroche, J., Yust, M.M., Girón-Calle, J., Vioque, J., Alaiz, M., Mateo, C., Guisán, J.M. and Millán, F. Stabilization–immobilization of carboxypeptidase A to aldehyde–agarose gels: A practical example in the hydrolysis of casein. Enzyme and Microbial Technology. 2002; 31(5): 711-718.

322. Perry, R.H., and Chilton, C.H. Chemical Engineering's Handbook. Mc Graw-Hill, London, 1973.

323. Satterfield, C.N. and Sherwood, T.K. *The role of diffusion in catalysis*. Addison-Wesley, Chap. 1. 1963.

324. Pizzuti, A., Calabrese, G., Bozzali, M., Telvi, L., Morizio, E., Guida, V., Gatta, V., Stuppia, L., Ion, A., Palka, G., and Dallapiccola, B. A peptidase gene in chromosome 8q is disrupted by a balanced translocation in a duane syndrome patient. *Invest Ophthalmol Vis Sci*. 2002; 43: 3609-12.

325. Polgár, L, and Halász, P. Current problems in mechanistic studies of serine and cysteine proteinases. *Biochem J*. 1982; 207(1): 1-10.

326. Polgár, L. The mechanism of action of aspartic proteases involves 'push-pull' catalysis. *FEBS Lett*. 1987; 219(1): 1-4.

327. Potts, B.C.M., Faulkner, D.J., Chan, J.A., Simolike, G.C., Offen, P., and Hemling, M.E., and Francis, T. Didemnaketals A and B, HIV-1 protease inhibitors from the ascidian Didemnum sp. *J. Am. Chem. Soc.*, 1991; 113 (16): 6321–6322.

328. Prashar, V., and Hosur, M.V. 1.8A X-ray structure of C95M/C1095F double mutant of tethered HIV-1 protease dimer complexed with acetyl pepstatin. *Biochem Biophys Res Commun*. 2004; 323: 1229-1235.
329. Quan, M. L.; Lam, P. Y. S.; Han, Q.; Pinto, D. J. P.; *et.al. Journal of Medicinal Chemistry*. 2004; *ASAP Alerts*.
330. Quiton, L., Demeure, K., Dobson, R., Gilles, N., Gabelita, V., and De Pauw, E. New method for characterizing highly disulfide-bridged peptides in complex mixtures: application to toxin identification from crude venoms. *Journal of proteome research*. 2007; 6: 3216-3223.
331. Radfar, A., Méndez, D., Moneriz, C., Linares, M., Marín-García, P., Puyet, A., Diez, A., and Bautista, J.M. Synchronous culture of *Plasmodium falciparum* at high parasitemia levels. *Nature Protocols*. 2009; 4: 1899-1915
332. Raksakulthai, R. and Haard, N.F. *J. Agric. Food Chem*. 2001; 49 (10): 5019-30.
333. Rawlings, N.D. and Barrettt, A.J. Evolutionary families of peptidases. Biochem. J. 1993; 290: 205–218.
334. Ramírez, A.R., Guerra, Y., Otero, A., García, B., Berry, C., Mendiola, J., Hernández-Zanui, A., and Chávez M.A. Generation of an affinity matrix useful in the purification of natural inhibitors of plasmepsin II, an antimalarial-drug target. *Biotechnol. Appl. Biochem.* 2009; 52(2), 149-157.
335. Rawlings, N.D. and Barrett, A.J. MEROPS: The peptidase database. Nucleic Acids Res. 1999; 27: 325–331.
336. Rawlings, N.D. and Barrett, A.J. MEROPS: the protease database. *Nucleic Acids Res.* 2000; 28: 323-325.
337. Rawlings, N.D., O'Brien, E., and Barrett, A.J. (2002) MEROPS: the protease database. *Nucleic Acids Res.* 2002; 30: 343-345.
338. Rawlings**,** N.D., Tolle, D.P., Barret, A.J. Evolutionary families of peptidase inhibitors. *Biochem. J*. 2004; 378: 705-716.
339. Rawlings, N.D., Barrett, A.J. and Bateman, A. (2010) MEROPS: the protease database. *Nucleic Acids Res*. 38:D227-33.
340. Rees, D. C., and Lipscomb, W. N. Refined crystal structure of the potato inhibitor complex of carboxypeptidaseAat 2.5Åresolution. *J. Mol. Biol*. 1982; 160: 475–498
341. Reid, G.E., and McLuckey, S.A. 'Top down' protein characterization via tandem mass spectrometry. *J Mass Spectrom*. 2002; 37(7): 663-75.
342. Reverter, D., García-Sáez, I., Catasús, L., Vendrell, J., Coll, M. and Avilés, F.X. (1997) Characterisation and preliminary X-ray diffraction analysis of human pancreatic procarboxypeptidase A2. *FEBS Lett*. 420(1): 7-10.
343. Reverter, D.; Vendrell, J.; Canals, F.; Horstmann, J.; Avilés, F.X.; Fritz, H. and Sommerhoff, C. P. A carboxypeptidase inhibitor from the medical leech. *Hirudo medicinalis*. *J. Biol. Chem*. 1998; 273(49): 32927-32933.
344. Reverter, D., Fernández-Catalán, C., Baumgartner, R., Pfänder, R., Huber, R., Bode, W., Vendrell, J., Holak, T.A. and Avilés, F.X. Structure of a novel leech carboxypeptidase inhibitor determined free in solution and in complex with human carboxypeptidase A2. *Nat Struct Biol*. 2000; 7(4): 322-8.
345. Reynolds DS, Gurley DS, Stevens RL, Sugarbaker DJ, Austen KF and Serafin WE. Cloning of cDNAs that encode human mast cell carboxypeptidase A, and comparison of the protein with mouse mast cell carboxypeptidase A and rat pancreatic carboxypeptidases. *Proc Natl Acad Sci USA*. 1989; 86: 9480-9484.
346. Reznik, S.E. and Fricker, L.D. Review: Carboxypeptidases from A to Z: implications in embryonic development and Wnt binding. *CMLS, Cell. Mol. Life Sci*. 2001; 58: 1790–1804.
347. Rhazi, N., Charlier, P., Dehareng, D., Engher, D., Vermeire, M., Frere, J.M., Nguyen-Disteche, M., and Fonze, E. Catalytic mechanism of the *Streptomyces* K15 DD-transpeptidase/penicillin-binding protein probed by site-directed mutagenesis and structural analysis. *Biochemistry*. 2003; 42: 2895-2906.
348. Rijken, D.C., and Lijnen, H.R. New insights into the molecular mechanism of the fibrinolysis system. *J. Trhomb. Haemost.* 2009; 7: 4-13.
349. Rodríguez de la Vega, M.; Sevilla, R.G.; Hermoso, A.; Lorenzo, J.; Tanco, S.; Diez, A.; Fricker, L.D.; Bautista, J. M. and Avilés, F.X. Nna1-like proteins are active

metallocarboxypeptidases of a new and diverse M14 subfamily. *FASEB J*. 2007; 20: 0001–0015.

350. Roepstorff, P., and Fohlman, J. Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biomed Mass Spectrom*. 1984; 11(11): 601.

351. Romanos, M. Advances in the use of *Pichia pastoris* for high-level expression. *Curr. Opinion Biotechnol.* 1995; 6: 527-533

352. Rosenthal, P.J. Proteases of Malaria Parasites: New Targets for Chemotherapy, Emerg. Infect. Diseases. 1998; 4: 49–56.

353. Rosenthal, P.J.; Sijwali, P.S.; Singh, A. and Shenai, B.R. Cysteine proteases of malaria parasites: targets for chemotherapy. *Curr Pharm Des*. 2002; 8(18):1659–1672.

354. Ryan, C.A.; Hass, G.M. and Kuhn, R.W. Purification and properties of a Carboxypeptidase Inhibitor from potatoes. *The Journal of Biological Chemistry*. 1974; 249(17): 5495-5499.

355. Ryan, C.A. Proteinase inhibitor gene families: strategies for transformation to improve plant defenses against herbivores. *Bioessays*. 1989; 10(1): 20-4.

356. Ryle, A.P. The porcine pepsins and pepsinogens. *Methods in Enzymol.* 1970; 19, 316-336

357. Rzychon, M.; Sabat, A.; Kosowska, K.; Potempa, J. and Dubin, A. Staphostatins: an expanding new group of proteinase inhibitors with a unique specificity for the regulation of staphopains, *Staphylococcus* spp. cysteine proteinases. *Molecular Microbiology*. 2003; 49(4): 1051–1066

358. Rzychon, M., Chmiel, D., and Stec-Niemczyk, J. Modes of inhibition of cysteine proteases. *Acta Biochim Pol*. 2004; 51(4): 861-73.

359. Sajid, M. and McKerrow, J.H. Cysteine proteases of parasitic organisms. Mol Biochem Parasitol. 2002; 120(1): 1–21.

360. Sakharov I.I. and Prieto G.A. *Mar Biotchnol*. (NY). 2000; 2(3): 259-266.

361. Salas, E., Cabrera, A., Chávez, M.A., and Cazzulo, J.J. (2012) Plextatin, a tight-binding inhibitor of cysteine proteases from the carribean coral *Plexaura homomalla*. Identification, functional characterization and evaluation of its antiparasitic eefect against *Plasmodium falciparum* and *Trypanosoma cruzi*. In press.

362. Sanglas, L., Valnickova, Z., Arolas, J.L., Pallarés, I., Guevara, T., Solà, M., Kristensen, T., Enghild, J.J., Avilés, F.X., Gomis-Rüth, F.X. Structure of activated thrombin-activatable fibrinolysis inhibitor, a molecular link between coagulation and fibrinolysis. *Mol Cell*. 2008; 31(4): 598-606.

363. Sanglas, L., Aviles, F.X., Huber, R., Gomis-Rüth, F.X., and L.Arolas, J.L. Mammalian metallocarboxypeptidase inhibition at the defense barrier of *Ascaris* parasite. *PNAS*. 2009; 106(6): 1743-1747.

364. Sanglas, L., Arolas, J.L., Valnickova, Z., Avilés, F.X., Enghild, J.J., and Gomis-Rüth, FX. Insights into the molecular inactivation mechanism of human activated thrombin-activatable fibrinolysis inhibitor. *J Thromb Haemost*. 2010; 8(5): 1056-65.

365. Schaschke, N. Miraziridine A: natures blueprint towards protease class-spanning inhibitors. *Bioorganic & Medicinal Chemistry Letters*. 2004; 14: 855–857.

366. Schägger, Hermann. Tricine-SDS-PAGE. *Nature Protocols.* 2006; 1(1): 16-22

367. Schechter, I. and Berger, A. On the size of the active site in proteases. I. Papain. *Biochem Biophys Res Común*. 1967; 27(2):157--162.

368. Schiff, E. R.; Pockros, P.; Shiffman, M. L.; McHutchison, J.; Gish, R.; Afdhal, N. H.; Makhviladze, M.; Huyghe, M.; Hecht, D.; Oltersdorf, T.; Shapiro, D. A. *Gastroenterology.* 2004; 126: A668-A668.Semba U, Shibuya Y, Okabe H, Hayashi I and Yamamoto T. Whale high-molecular-weight and low-molecular-weight kininogens. Thromb Res. 2000; 97(6): 481-90.

369. Seidler, J., Zinn, N., Boehm, M.E., and Lehmann, W.D. De novo sequencing of peptides by MS/MS. *Proteomics*. 2010; 10(4): 634-49.

370. Semba, U.; Umeda, Y.; Shibuya, Y.; Okabe, H.; Tanase, S. and Yamamoto, T. Primary structures of guinea pig high- and low-molecular-weight kininogens. *Int. Immunopharmacol*. 2004; 4(10-11): 1391-400.

371. Sencic, L. and Macek, P. New method for isolation of venom from the sea anemone *Actinia cari*. Purification and characterization of cytolytic toxins. *Biochem. Physiol*. 1990; 97B: 687.

372. Senko, M.W., Speir, J.P., and McLafferty, F.W. Collisional activation of large multiply charged ions using Fourier transform mass spectrometry. *Anal Chem*. 1994; 66(18): 2801-8.
373. Serrano, S.M.T., Shannon, J.D., Wang, D., Camargo, A.C.N., and Fox, J.W., *Proteomics*. 2005; 5: 501-510.
374. Shesely, E.; Hu, C.; Alhenc-Gelas, F.; Meneton, P. and Carretero, O. A second expressed kininogen gene in mice. *Physiol. Genomics*. 2006; 26(2): 152-7.
375. Shi, X., Karkut, T., Chamankhah, M., Alting-Mees, M., Hemmingse, S.M., and Hegedus, D. Optimal conditions for the expression of a single-chain antibody (scFv) gene in *Pichia pastoris*. *Protein Exp. Pur.* 2003; 28: 321–330.
376. Shimamori, Y., Kumagai, Y., Watanabe, Y., and Fujimoto, Y. Human placental carboxypeptidase M is anchored by a glycosyl-phosphatidylinositol moiety. Biochem Int. 1990; 20(3): 607-13.
377. Shiomi, K.; Tanaka, E.; Yamanaka, H. and Kikuchi, T. Isolation and characterization of a lethal hemolysin in the sea anemone *Parasicyonis actinostoloide*. *Toxicon*. 1985; 23:865.
378. Sielecki, A.R., Fedorov, A.A., Boodhoo, A., Andreeva, N.S., and James, M.N. Molecular and crystal structures of monoclinic porcine pepsin refined at 1.8 A resolution. *J Mol Biol*. 1990; 214(1): 143-70.
379. Silverman, G.A., Bird, P. I., Carrell, R.W., Church, F.C., Coughlin, P.B., Gettins, P.G. *et al*. The serpins are an expanding superfamily of structurally similar but functionally diverse proteins: evolution, mechanism of inhibition, novel functions, and a revised nomenclature. J. Biol. Chem. 2001; 276: 33293–33296.
380. Skidgel, RA. Basic carboxypeptidases: regulators of peptide hormone activity. *Trends Pharmacol Sci*. 1988; 9(8):299–304.
381. Skidgel, R.A. Structure and function of mammalian zinc carboxypeptidases. In: Hooper NM, editor. Zinc metalloproteases in health and disease. London: Taylor and Francis Ltd. 1996. p. 241–83.
382. Skidgel, R.A., McGwire, G.B., and Li, X.Y. Membrane anchoring and release of carboxypeptidase M: implications for extracellular hydrolysis of peptide hormones. *Immunopharmacology*. 1996; 32(1-3): 48-52.
383. Skidgel, R.A. *Zinc Metalloproteases in Health and Disease*. 1996; (Edited by Nigel M. Hooper. Taylor and Francis. Ltd).
384. Skidgel, R.A., and Erdos, E.G. Lysine carboxypeptidase. In: Barrett AJ, Rawlings ND, Woessner JF Eds. Handbook of Proteolytic Enzymes, 2 edn. Elsevier, London. 2004; 837-40.
385. Standing, K.G. Peptide and protein de novo sequencing by mass spectrometry. *Current Opinion in Structural Biology*. 2003; 13: 595–601.
386. Štrukelj, B., Lenarčič, B., Gruden, K., Pungerčar, J., Rogelj, B., Turk, V., Bosch, D., and Jongsma, MA. Equistatin, a Protease Inhibitor from the Sea Anemone *Actinia equina*, Is Composed of Three Structural and Functional Domains. *Biochemical and Biophysical Research Communications*. 2000; 269: 732-736.
387. Smith, P.K.; Krohn, R.I.; Hermanson, G.T.; Mallia, A.K.; Gartner, F.H.; Provenzano, M.D.; Fujimoto, E.K.; Goeke, N.M.; Olson, B.J. and Klenk, D.C. Measurement of protein using bicinchoninic acid. Anal Biochem. 1985; 150: 76-85.
388. Sober, H. A., Harte, R.A., Sober, E. K., eds (1968): Handbook of Biochemistry. Selected data for Molecular Biology. 2da Edición. The Chemical Rubber Co., Ohio.
389. Sobott F, and Robinson CV. Protein complexes gain momentum. Curr Opin Struct Biol. 2002; 12(6): 729-34.
390. Song, L., and Fricker, L.D. Cloning and expression of human carboxypeptidase Z, a novel metallocarboxypeptidase. *J Biol Chem*. 1997; 272(16): 10543-50.
391. Springman, E.B., Dikov, M.M., and Serafin, W.E. Mast cell procarboxypeptidase A. Molecular modeling and biochemical characterization of its processing within secretory granules. *J Biol Chem*. 1995; 270: 1300-7.
392. Stein, P.E., and Carrell, R.W. What do dysfunctional serpins tell us about molecular mobility and disease? *Nat. Struct. Biol*. 1995; 2: 96–113.
393. Strobl, S., Fernandez-Catalan, C., Braun, M., Huber, R., Masumoto, H., Nakagawa, K., Irie, A., Sorimachi, H., Bourenkow, G., Bartunik, H., Suzuki, K., and Bode, W. The crystal structure of calcium-free human m-calpain suggests an electrostatic switch mechanism for activation by calcium. *Proc Natl Acad Sci USA*. 2000; 97(2): 588-92.

394. Suau, T., Álvaro, G.M., Benaiges, D., López-Santín, J. Performance of an immobilized fuculose-1-phosphate aldolase for stereoselective synthesis. *J. Biocatalysis and Biotransformation*. 2009; 27(2): 136-142.
395. Suckau, D., Resemann, A., Schuerenberg, M., Hufnagel, P., Franzen, J., Holle, A. A novel MALDI LIFT-TOF/TOF mass spectrometer for proteomics. *Anal. Bioanal. Chem.* 2003, 376: 952-965 **(a)**.
396. Suckau, D., Resemann, Anja. T3-sequencing: targeted characterization of the N- and C-termini of undigested proteins by mass spectrometry. *Anal. Chem.* 2003, 75: 5817-5824 **(b)**
397. Summerfield, S.G., Whiting, A., and Gaskell, S.J. *Int. J. Mass Spectrom. Ion Processes*. 1997; 162: 149–161.
398. Takahashi, K., and Chang, W.J. Specific chemical modifications of acid proteases in the presence and absence of pepstatin. *J Biochem*. 1973; 73(3): 675-7.
399. Tan, A.K. and Eaton, D.L. Activation and characterization of procarboxypeptidases B from human plasma. *Biochem*. 1995; 34(17): 5811-5816.
400. Tan, F., Rehli, M., Krause, S.W., and Skidgel, R.A. Sequence of human carboxypeptidase D reveals it to be a member of the regulatory carboxypeptidase family with three tandem active site domains. *Biochem J*. 1997; 327(1): 81-7.
401. Tan, F., Balsitis, S., Black, J.K., Blöchl, A., Mao, J.F., Becker, R.P., Schacht, D., and Skidgel, R.A. Effect of mutation of two critical glutamic acid residues on the activity and stability of human carboxypeptidase M and characterization of its signal for glycosylphosphatidylinositol anchoring. *Biochem J*. 2003; 370(2): 567-78.
402. Tanco, S., Zhang, X., Morano, C., Avilés, F.X., Lorenzo, J., and Fricker, L.D. Characterization of the substrate specificity of human carboxypeptidase A4 and implications for a role in extracellular peptide processing. *J Biol Chem.* 2010; 285(24): 18385-96.
403. Tang, J., Sepulveda, P., Marciniszyn, J.Jr., Chen, K.C., Huang, W.Y., Tao, N., Liu, D., and Lanier, J.P. Amino-acid sequence of porcine pepsin. *Proc Natl Acad Sci USA*. 1973; 70(12): 3437-9.
404. Tardioli, P.W., Fernández-Lafuente, R., Guisán, J.M., and Giordano, R.L.C. Design of New Immobilized-Stabilized Carboxypeptidase A Derivative for Production of Aromatic Free Hydrolysates of Proteins. *Biotechnol. Prog.* 2003; 19: 565-574.
405. Thiede, B., Höhenwarter, W., Krah, A., Mattow, J., Schmid, M., Schmidt, F., Jungblut, P.R. Peptide mass fingerprinting. *Methods*. 2005; 35(3): 237-47.
406. Thompson, M.S., Cui, W., and P.Reilly, J. Abstracts of Papers, 51st ASMS Meeting, Montreal, Canada, June 8–12, 2003.
407. Tobias, J.W., Shrader, T.E., Rocap, G., and Varshavsky, A. The N-end rule in bacteria. *Science*. 1991; 254: 1374-1377
408. Turk, V., J. Brzin, B. Lenarcic, P. Locnikar, T. Popovic, A. Ritonja, J. Babnik, W. Bode, and W. Machleidt: Structure and Function of Lysosomal Cysteine Proteinases and Their Protein Inhibitors. In: *Intracellular Protein Catabolism*. 1985; pp. 91-103. (E. A. Khairallah, J. S. Bond, and J. W. C. Bird, Eds.) Alan R. Liss, New York.
409. Turk, D.; Podobnik, M.; Kuhelj, R.; Dolinar, M. and Turk, V. Crystal structures of human procathepsin B at 3.2 and 3.3 Å resolution reveal an interaction motif between a papain-like cysteine protease and its propeptide. *FEBS Lett*, 1996. 384: 211–214.
410. Turk, B., Turk, du S.A. and Turk, V. Protease signalling: the cutting edge. *EMBO J*. 2012; 31(7): 1630-43.
411. Turk, B. Targeting proteases: successes, failures and future prospects. *Nat Rev Drug Discov*. 2006; 5(9):785-99.
412. Umezawa, H., Aoyagi, T., Morishima, H., Matsuzaki, M., and Hamada, M. Pepstatin, a new pepsin inhibitor produced by Actinomycetes. *J. Antibiot.* 1970; 23(5): 259-62.
413. Umezawa, H. Enzyme Inhibitors of Microbial Origin. 1972; p. 31, University Park Press, Baltimore.
414. Van Noorden, J. F. IBC's 2nd International Conference on Protease Inhibitors: Novel Therapeutic Applications and Development, Washington DC, 24-26 February. *Acta Histochem.* 1997; 249-55.
415. Vannice, M.A. Kinetic of catalytic reactions. Press in Springer Science +business media, Inc. 2005. New York, USA.

416. Vendrell, J.; Querol, E. and Avilés, F. X. Metallocarboxypeptidases and their protein inhibitors: Structure, function and biomedical properties. *Act. Biochem. Biophys*. 2000; 1477: 284-298.

417. Venhudova, G., Canals, F., Querol, E., and Avilés F.X. Mutations in the N- and C-terminal tails of potato carboxypeptidase inhibitor influence its oxidative refolding process at the reshuffling stage. *J Biol Chem*. 2001; 276: 11683-11690.

418. Vendrell J., Aviles F.X. and Fricker L.D. *Handbook of metalloproteins* (edited by Messerschmidt A.; Bode W. and Cygler M. Jonh Wiley and Sons, Ldt,). 2004; 3: 176-189.

419. Villanueva, J., Canals, F., Querol, E., and Avilés, F.X. Monitoring the expression and purification of recombinant proteins by MALDI-TOF mass spectrometry. *Enzyme Microb. Technol*. 2001; 29: 99–103

420. Villanueva, J. Yanes, O.; Querol, E.; Serrano, L. and Aviles, F.X. Identification of Protein Ligands in Complex Biological Samples Using Intensity-Fading MALDI-TOF Mass Spectrometry. *Anal. Chem*. 2003; 75: 3385-3395.

421. Visconte, C.; Sainte-Marie, M.; Lorrain, J.; Millet, L.; O'Connor, S. E.; Schaeffer, P.; Herbert, J. M. *Journal of Thrombosis and Haemostasis*. 2004; 2: 629-636.

422. Weisz, P. B. and Prater C. D., Interpretation of Measurement in Experimental Catalysis. *Adv. Catal*. 1954; 6: 143.

423. Wei, S., Segura, S., Vendrell, J., Avilés, F.X., Lanoue, E., Day, R., Feng, Y., and Fricker, L.D. Identification and characterization of three members of the human metallocarboxypeptidase gene family. *J Biol Chem*. 2002; 277(17): 14954-64.

424. Wei, S., Feng, Y., Kalinina, E., and Fricker, L.D. Neuropeptide-processing carboxypeptidases. *Life Sci*. 2003; 73: 655-62.

425. Werner, R., Guitton, M.C., and Mühlbach, H.P. Nucleotide sequence of a cathepsin D inhibitor protein from tomato. *Plant Physiol*. 1993; 103(4): 1473.

426. Wilke, C. R., and Chang, P. Correlation of diffusion coefficients in dilute solutions. *AIChE Journal*. 1955; 1: 264–270.

427. Willemse, J.L., and Hendriks, D.F. A role of procarboxypeptidase U (TAFI) in thrombosis. *Front. Biosci.* 2007; 12: 1973-1987.

428. Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G., McCoy, A., McNicholas, S.J., Murshudov, G.N., Pannu, N.S., Potterton, E.A., Powell, H.R., Read, R.J., Vagin, A., and Wilson K.S. Overview of the CCP4 suite and current developments. *Acta. Cryst. D*. 2011; 67: 235-242.

429. Wlodawer, A., Walter, J., Huber, R., and Sjölin, L. Structure of bovine pancreatic trypsin inhibitor. Results of joint neutron and X-ray refinement of crystal form II. *J Mol Biol*. 1984; 180(2): 301-29.

430. Wysocki, V.H., Tsaprailis, G., Smith, L. L., and Breci, L.A. *J. Mass Spectrom*. 2000; 35: 1399–1406.

431. Wysocki VH, Resing KA, Zhang Q, Cheng G. Mass spectrometry of peptides and proteins. Methods. 2005; 35: 211–222.

432. Wlodawer, A., and Vondrasek, J. Inhibitors of HIV-1 protease: a major success of structure-assisted drug design. *Annu Rev Biophys Biomol Struct*. 1998; 27: 249-284.

433. Woessner, J.F.Jr. The family of matrix metalloproteinases. *Ann N Y Acad Sci*. 1994; 732: 11-21.

434. Yamane, J., Yao, M., Zhou, Y., Hiramatsu, Y., Fujiwara, K., Yamaguchi, T., Yamaguchi, H., Togame, H., Tsujishita, H., Takemoto, H., and Tanaka, I. In-crystal affinity ranking of fragment hit compounds reveals a relationship with their inhibitory activities. *J.Appl.Crystallogr*. 2011; 44: 798-804.

435. Yanes, O., Villanueva, J., Querol, E., and Avilés, F.X. Intensity-fading MALDI-TOF-MS: novel screening for ligand binding and drug discovery. *Drug Discov. Today: TARGETS*. 2004; 3: Suppl. 1, 23–30

436. Yanes, O.; Villanueva, J.; Querol, E. and Avilés, F.X. Functional screening of serine protease inhibitors in the medical leech *Hirudo medicinalis* monitored by intensity fading MALDI-TOF MS. Mol. Cell Proteomics. 2005; 4: 1602–1613.

437. Yanes, O.; Villanueva, J.; Querol, E. and Avilés, F.X. Detection of non-covalent protein interactions by 'intensity fading' MALDI-TOF mass spectrometry: applications to proteases and protease inhibitors. *Nature Protocols*. 2007; 2: 119-130 **(a)**.

438. Yanes, O., Aviles, F.X., Roepstorff, P. and Jorgensen, T.J. Exploring the "intensity fading" phenomenon in the study of noncovalent interactions by MALDI-TOFmass spectrometry. *J Am Soc Mass Spectrom*. 2007; 18: 359–367 **(b)**.

439. Yates, J.R. 3rd. Database searching using mass spectrometry data. *Electrophoresis*. 1998; 19(6): 893-900.

440. Ye, S., Cech, A.L., Belmares, R., Bergstrom, R.C., Tong, Y., Corey, D.R., Kanost, M.R., and Goldsmith, E.J. The structure of a Michaelis serpin-protease complex. *Nat Struct Biol*. 2001; 8(11): 979-83.

441. Yokosawa, H., Odajima, R., and Ishii, S. Trypsin inhibitor in the hemolymph of *Halocynthia roretzi*. Purification and characterization. *J. Biochem*. 1985; 97: 1621-1630.

442. Zhang, W., Inan, M., and Meagher, M.M. Fermentation strategies for recombinant protein expression in the methylotrophic yeast *Pichia pastoris*. *Biotechnol. Bioprocess Eng*. 2000; 5: 275-287.

443. Zingali, R. B., Jandrot-Perrus, M., Guillin, M.C. and Bon, C. Bothrojaracin, a new thrombin inhibitor isolated from *Bothrops jararaca* venom: characterization and mechanism of thrombin inhibition. *Biochemistry*. 1993; 32**:** 10794-10802.

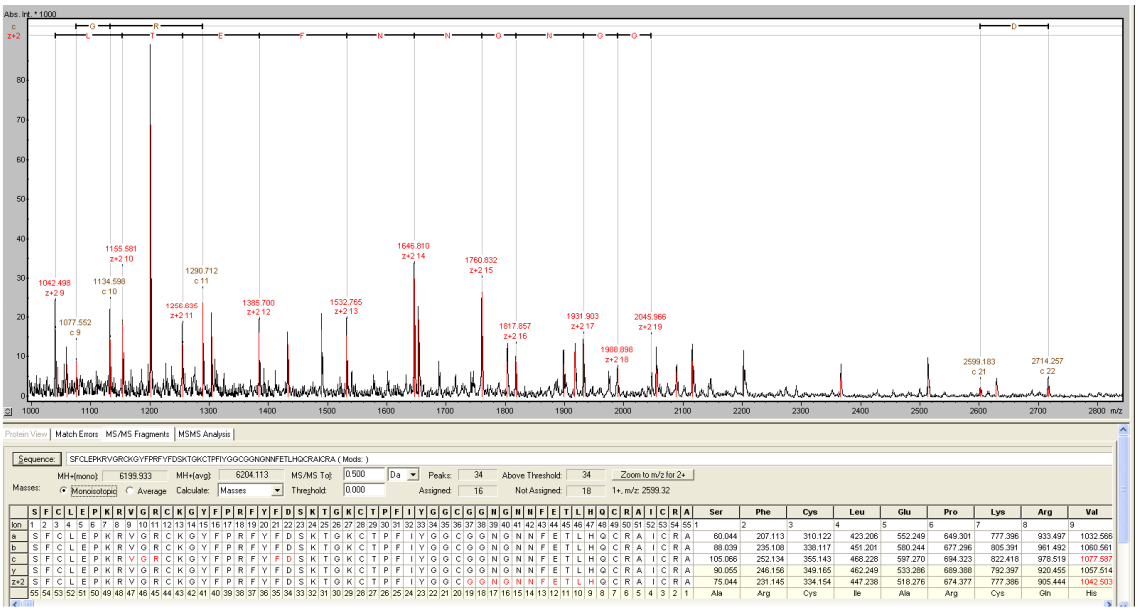
444. Zubarev, R.A., Kelleher, N.L., and McLafferty, F. W. *J. Am. Chem. Soc*. 1998; 120: 3265–3266.

445. Zwilling, R.; Jacob, F.; Bauer, H.; Neurath, H. and Enfield D.L. *Eur J. Biochem*. 1979; 94(1): 223-9.

446. Zykova, T. A.; Monastyirnaia, M. M.; Apalikov, O. V.; Shvets, T. V. and Kzlovskaia, E. P. Low molecular cytolysins and trypsin inhibitors from sea anemone *Radianthus macrodactylus*. Isolation and partial characterization. *Bioorg. Khim*. 1998; 24(7): 509-516.

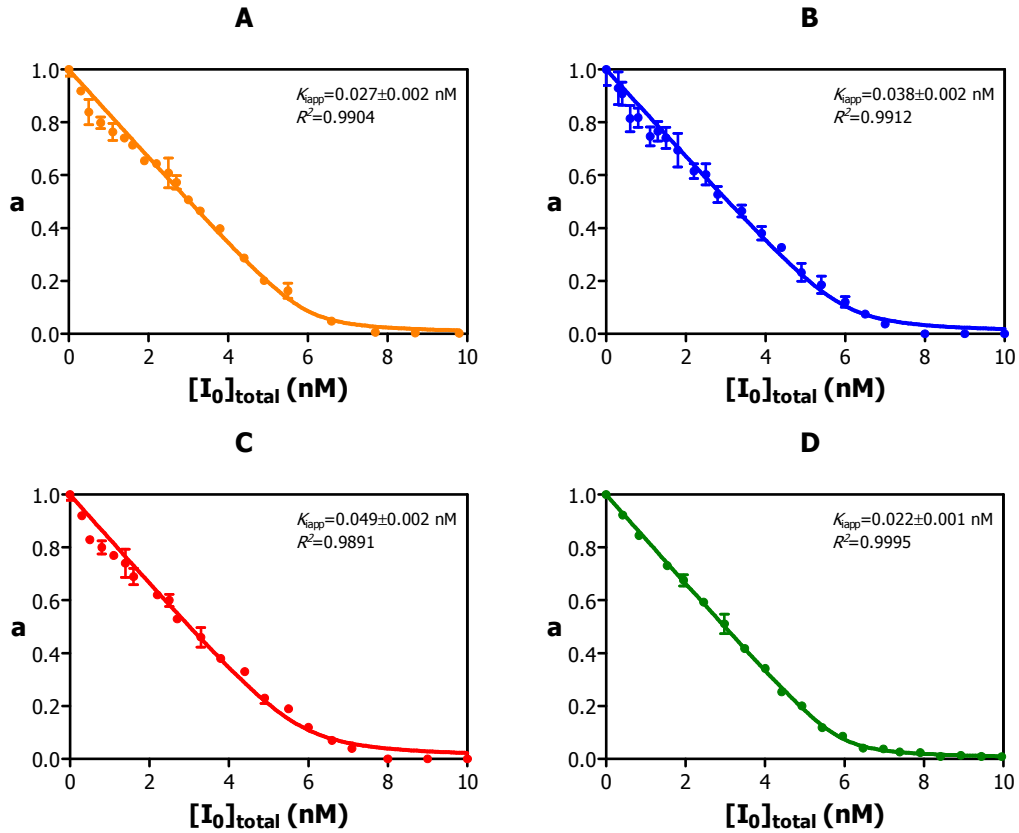
ANNEXES

ANNEX I.

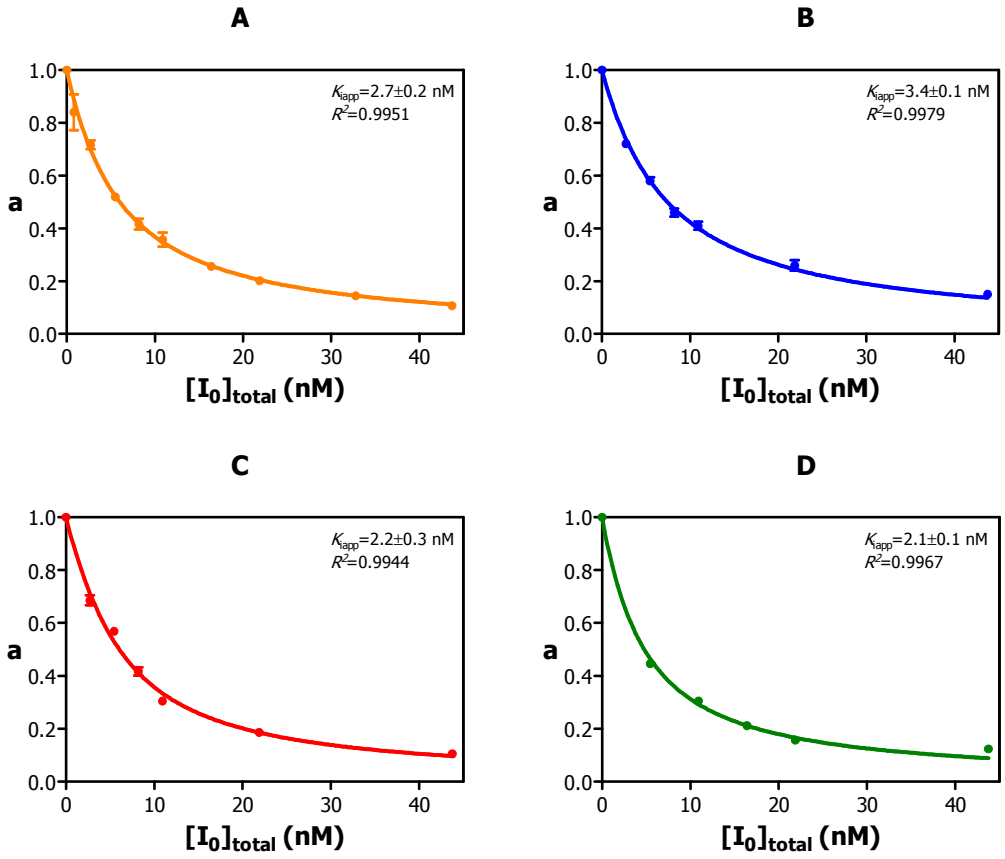


Tags mapping to sequences of ShPI-II in tree view

ANNEX II.

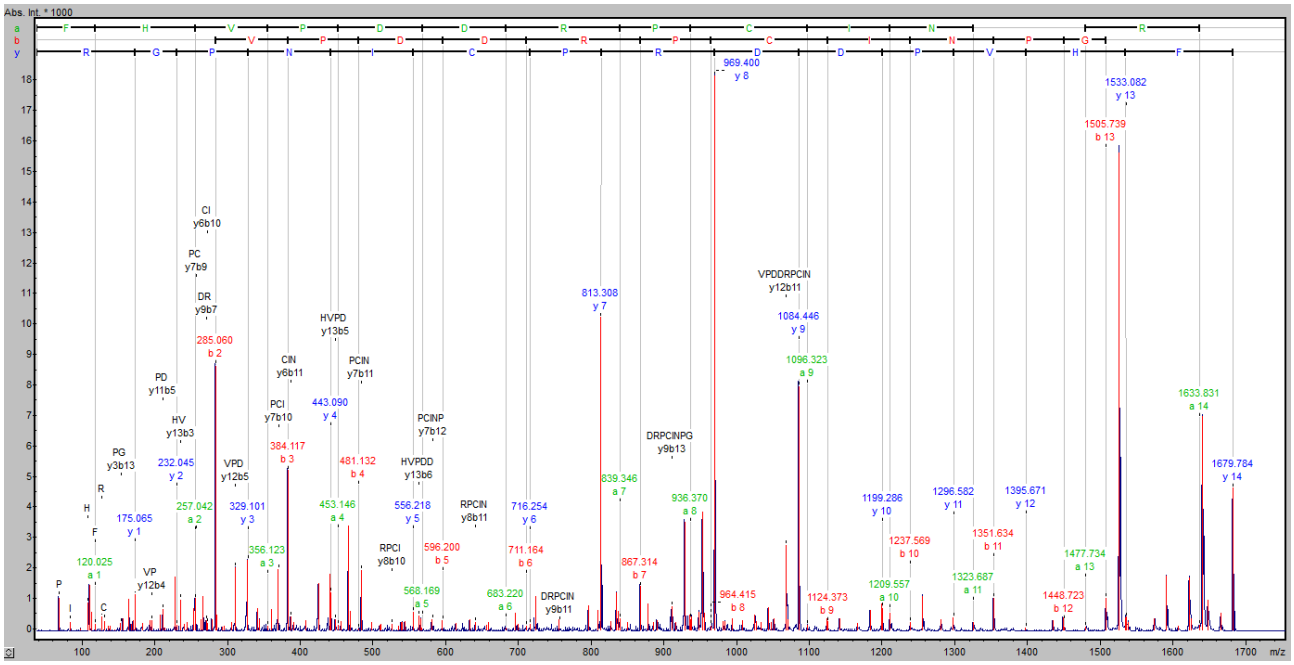


Apparent equilibrium dissociation constant (K_{iapp}) of NvCI isoforms against bovine CPA.
(A). NvCIa. **(B).** NvCIb. **(C).** NvCIc. **(D).** NvCId. [bCPA]=3.5 nM. Substrate: AAFP, [S_0]=100 μM. Preincubation time: 15 min, T=37ºC. Other specifications are described in Materials and Methods. a=v_i/v_0: fraction of enzymatic activity in the presence (v_i) and absence of inhibitor (v_0) in terms of initial velocities. The best-fit value of K_{iapp} was performed by adjusting the experimental values to the Morrison equation (Eq. 23) using the program GraphPad Prism 5 (GraphPad Software, Inc.) at $p<0.05$. Data are means (n=3) ± S.D.

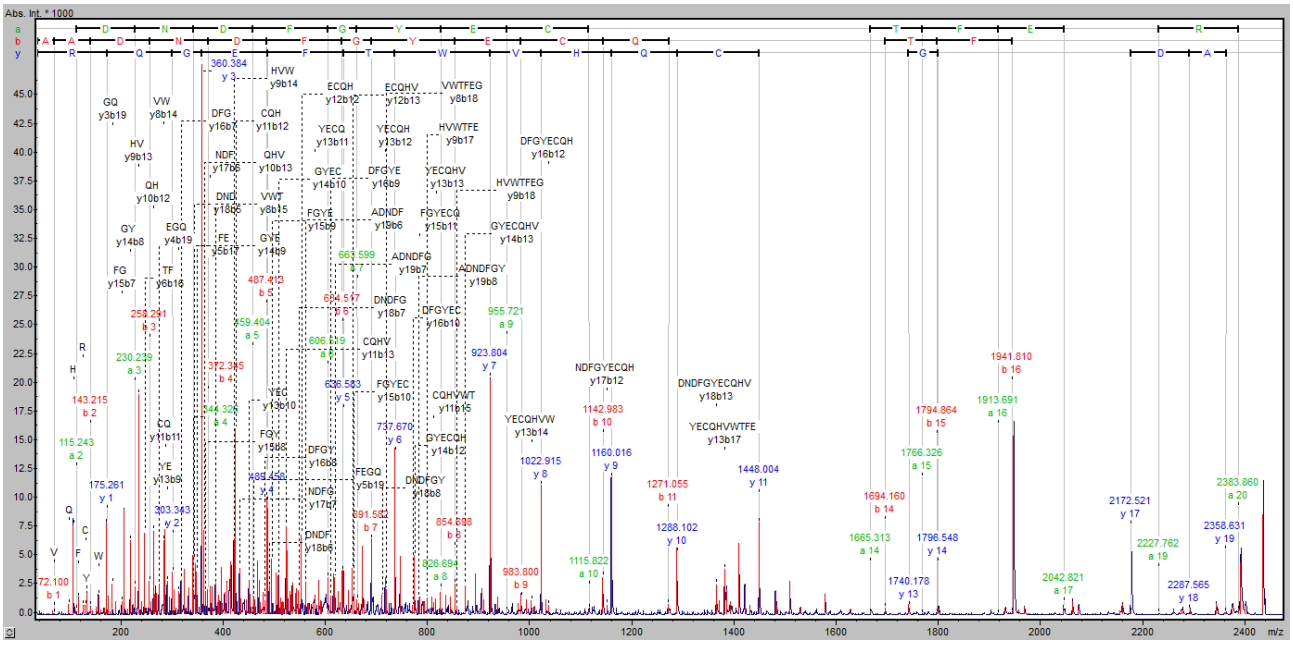


Apparent equilibrium dissociation constant (*K_{iapp}*) of NvCI isoforms against porcine CPB. **(A).** NvCIa. **(B).** NvCIb. **(C).** NvCIc. **(D).** NvCId. [pCPB]=1.9 nM. Substrate: AAFA, [S_0]=100 µM. Preincubation time: 15 min, T=37ºC. Other specifications are described in Materials and Methods. a=v_i/v_0: fraction of enzymatic activity in the presence (v_i) and absence of inhibitor (v_0) in terms of initial velocities. The best-fit value of *K_{iapp}* was performed by adjusting the experimental values to the Morrison equation (Eq. 23) using the program GraphPad Prism 5 (GraphPad Software, Inc.) at $p<0.05$. Data are means (n=3) ± S.D.

ANNEX III.

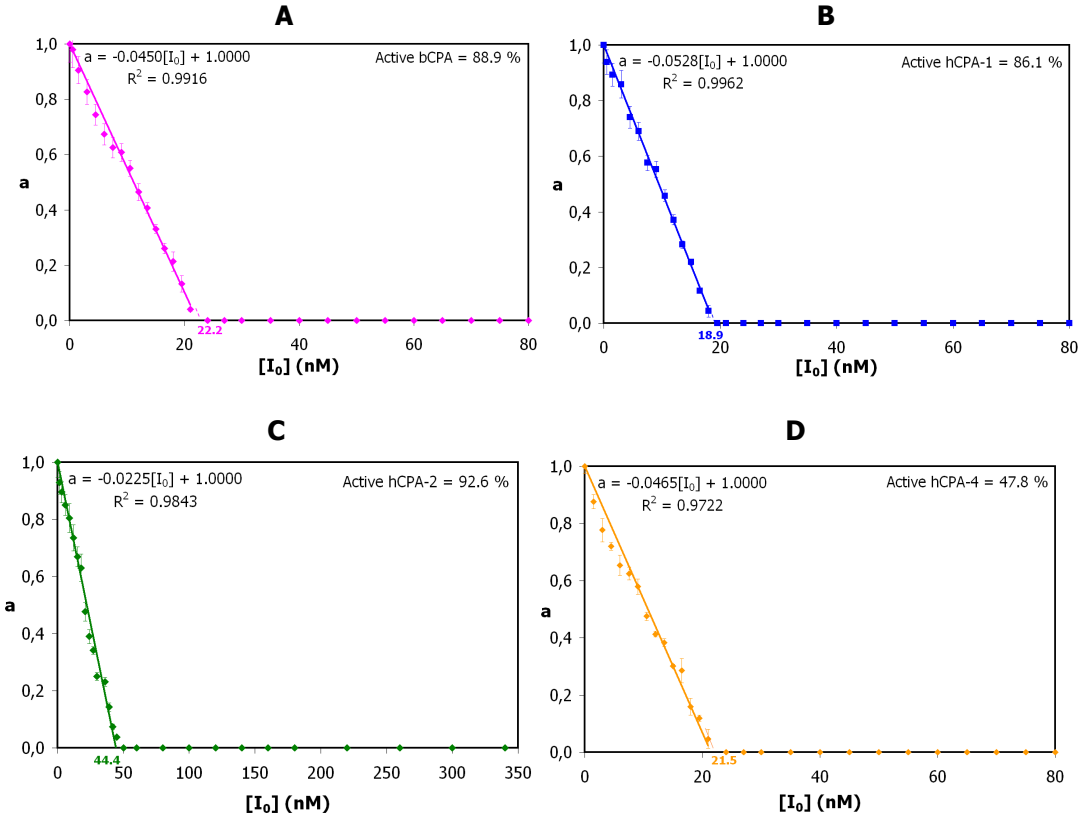


MALDI MS/MS spectrum of parent m/z=1680 from the enzymatic digestion with trypsin of NvCI. Sequence derived from MS/MS spectrum by *de novo* sequencing: FHVPDDRPCI NPGR

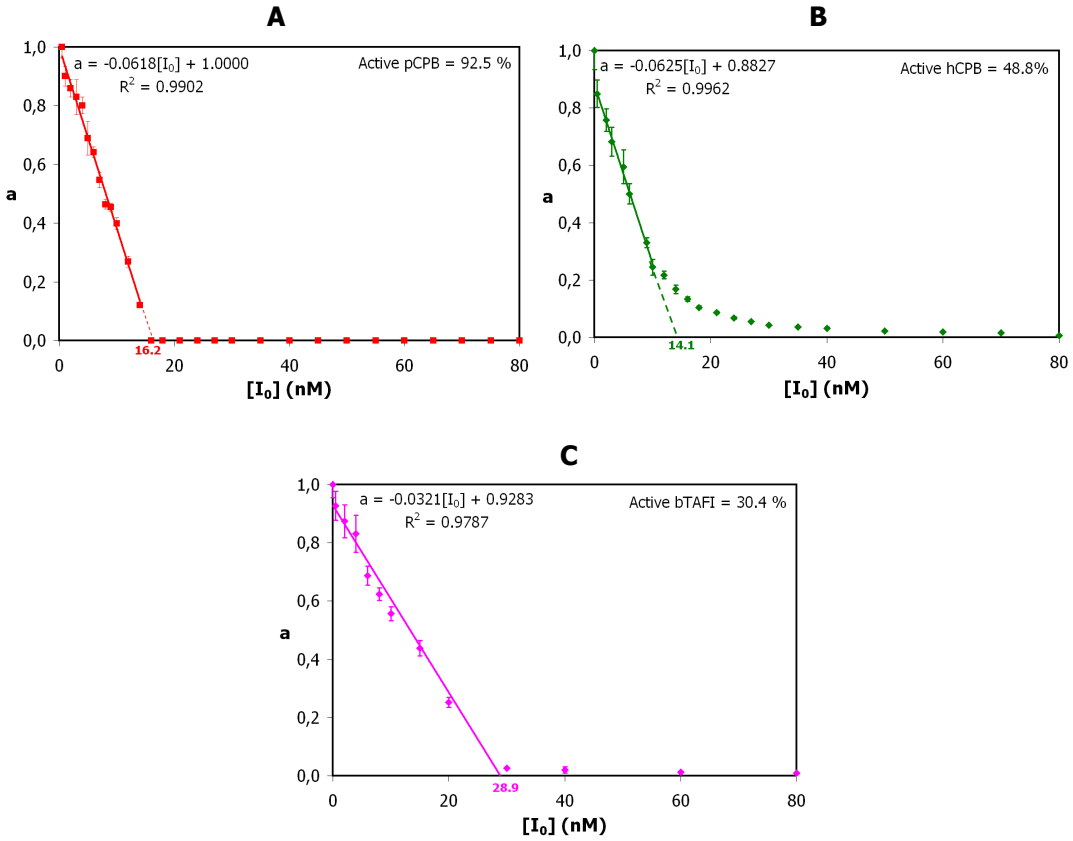


MALDI MS/MS spectrum of parent m/z=2430 from the enzymatic digestion with trypsin of NvCI. Sequence derived from MS/MS spectrum by *de novo* sequencing: AADNDFGYECQHVWTFEGQR

ANNEX IV.

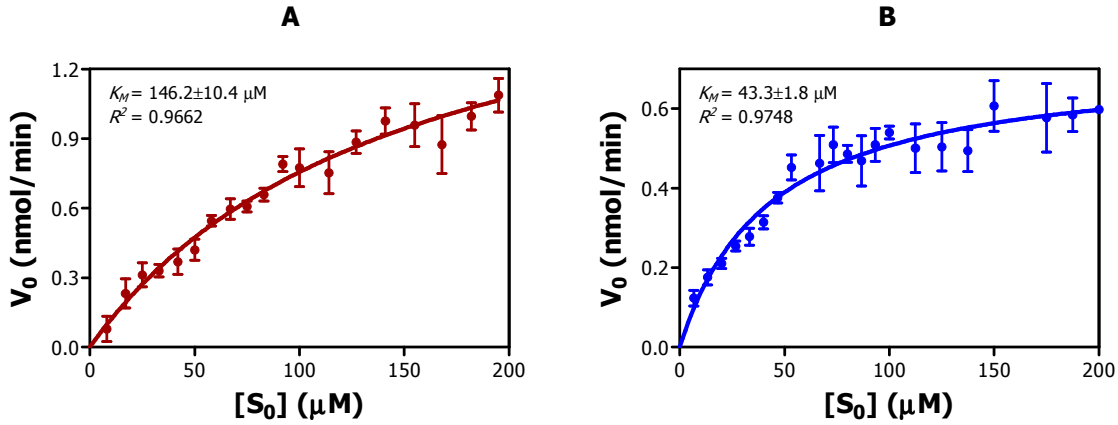


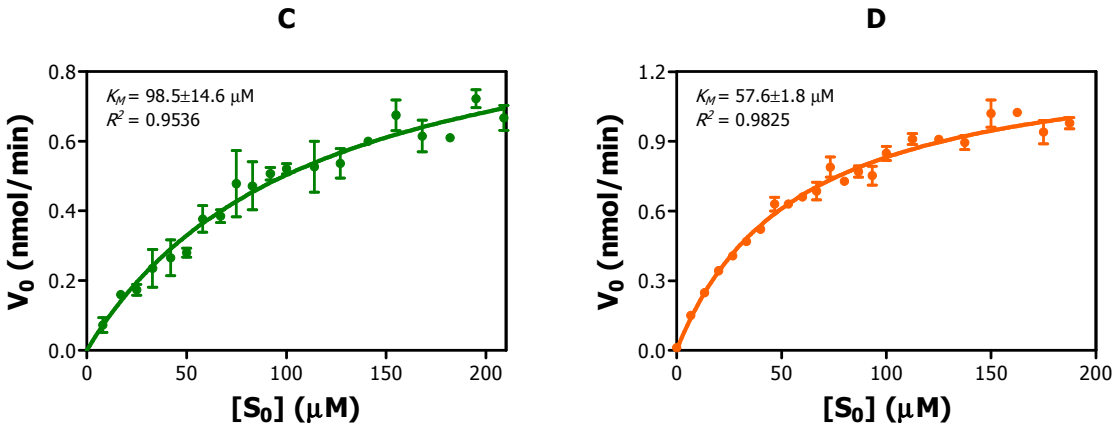
CPA-like enzyme titration curves with Tick Carboxypeptidase Inhibitor (TCI).
A. [bCPA]=25.0 nM, $[E_0]/K_i$=22.7. **B.** [hCPA1]=22.0 nM, $[E_0]/K_i$=18.3. **C.** [hCPA2]=48.0 nM, $[E_0]/K_i$=13.3. **D.** [hCPA4]=45.0 nM, $[E_0]/K_i$=56.2. Substrate: AAPP. Preincubation time: 15 min, T=37ºC. Other specifications are as described in Materials and Methods. $a=v_i/v_0$: fraction of enzymatic activity in the presence (v_i) and absence of inhibitor (v_0) in terms of initial velocities. Data are means (n=3) ± S.D.



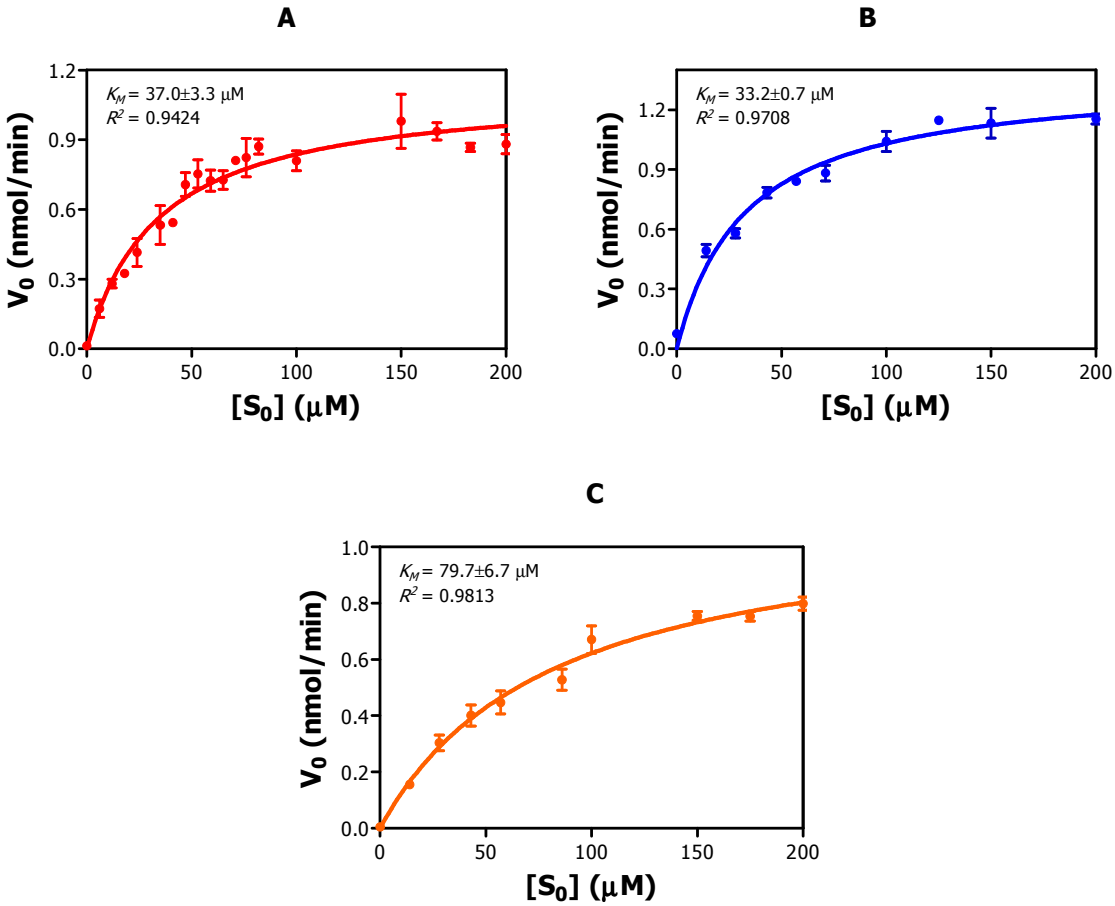
CPB-like enzyme titration curves with Tick Carboxypeptidase Inhibitor (TCI).
A. [pCPB]=17.5 nM, $[E_0]/K_i$=10.7. **B.** [hCPB]=29.0 nM, $[E_0]/K_i$=22.2. **C.** [bTAFI]=95.0 nM, $[E_0]/K_i$=73.1. Substrate: AAPA. Preincubation time: 15 min, T=37ºC. Other specifications are as described in Materials and Methods. $a=v_i/v_0$: fraction of enzymatic activity in the presence (v_i) and absence of inhibitor (v_0) in terms of initial velocities. Data are means (n=3) ± S.D.

ANNEX V.



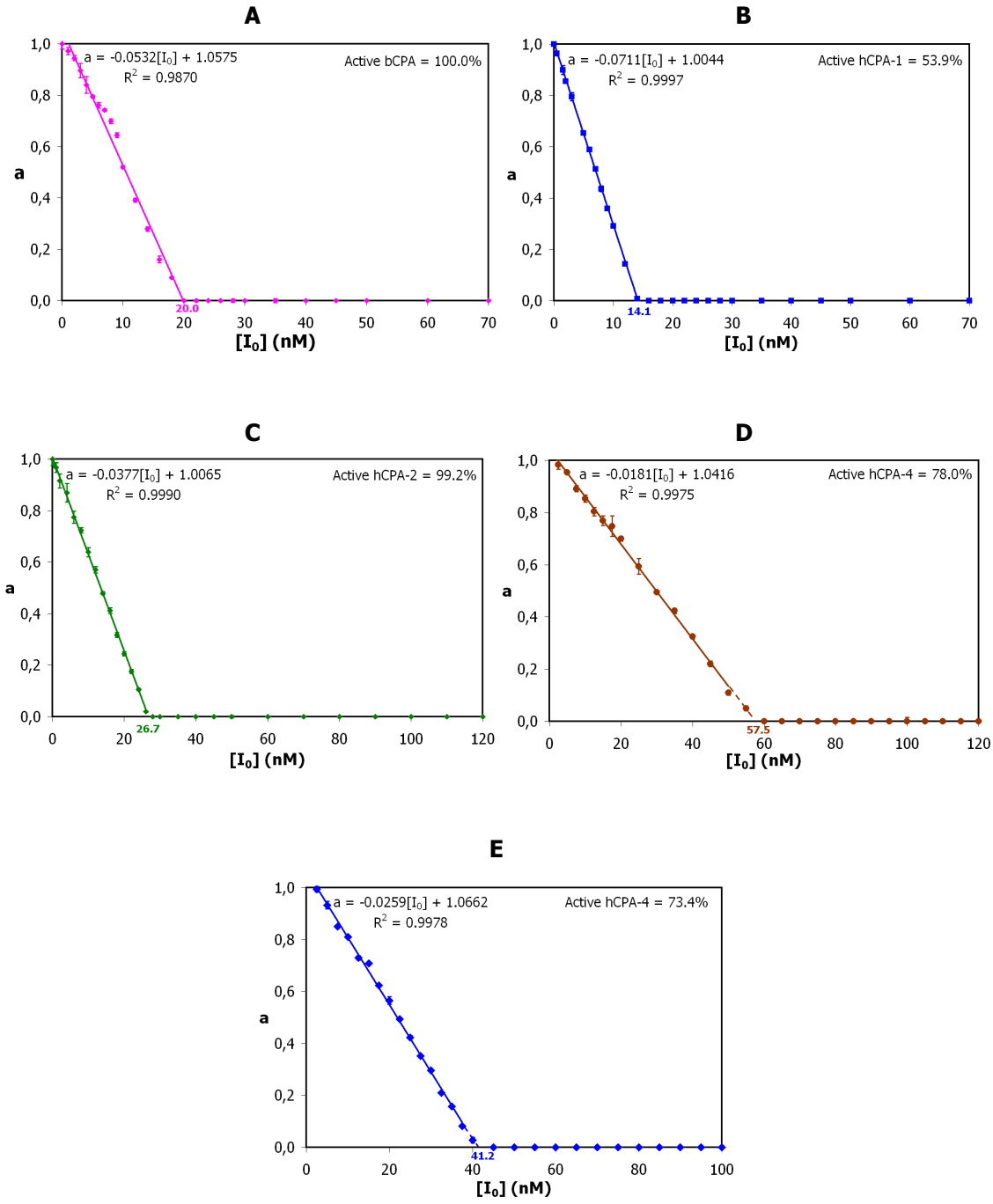


Michaelis-Menten constant (K_M) for AAPP hydrolysis catalyzed by various metalloproteases. **A.** bCPA. **B.** hCPA1. **C.** hCPA2. **D.** hCPA4. Proteolytic activity was followed by its ability to hydrolyse AAPP synthetic substrate in a final volume of 250 µl, buffer: 20 mM Tris-HCl pH 7.5, 500 mM NaCl, 1% v/v DMSO, 0.05% w/v BRIJ-35 and T=37ºC. The best-fit value of K_M was performed by adjusting the experimental values to the Michaelis-Menten equation (Eq. 22) using the program GraphPad Prism 5 (GraphPad Software, Inc.) at p<0.05. Data are means (n=3) ± S.D.

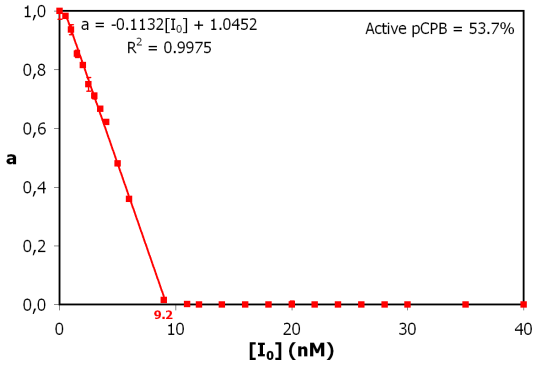


Michaelis-Menten constant (K_M) for AAPP hydrolysis catalyzed by various metalloproteases. **A.** pCPB. **B.** hCPB. **C.** bTAFI. Proteolytic activity was followed by its ability to hydrolyse AAPA synthetic substrate in a final volume of 250 µl, buffer: 20 mM Tris-HCl pH 7.5, 100 mM NaCl, 1% v/v DMSO, 0.05% w/v BRIJ-35 and T=37ºC. The best-fit value of K_M was performed by adjusting the experimental values to the Michaelis-Menten equation (Eq. 22) using the program GraphPad Prism 5 (GraphPad Software, Inc.) at p<0.05. Data are means (n=3) ± S.D.

ANNEX VI.



CPA-like enzyme titration curves with Tick Carboxypeptidase Inhibitor (TCI).
A. [bCPA]=20.0 nM, $[E_0]/K_i$=18.2. **B.** [hCPA1]=26.0 nM, $[E_0]/K_i$=21.7. **C.** [hCPA2]=27.0 nM, $[E_0]/K_i$=7.5. **D.** [hCPA4]=73.5 nM, $[E_0]/K_i$=91.9 nM, pH=7.5. **E.** [hCPA4]=56.0 nM, $[E_0]/K_i$=70.2, pH=8.5. Substrate: AAPP. Preincubation time: 15 min, T=37ºC. Other specifications are as described in Materials and Methods. $a=v_i/v_0$: fraction of enzymatic activity in the presence (v_i) and absence of inhibitor (v_0) in terms of initial velocities. Data are means (n=3) ± S.D.



Porcine CPB titration curve with Tick Carboxypeptidase Inhibitor (TCI).
[pCPB]=17.0 nM, $[E_0]/K_i$=10.7. Substrate: AAPA. Preincubation time: 15 min, T=37ºC. Other specifications are as described in Materials and Methods. $a=v_i/v_0$: fraction of enzymatic activity in the presence (v_i) and absence of inhibitor (v_0) in terms of initial velocities. Data are means (n=3) ± S.D.