



Universitat Autònoma de Barcelona

Departament de Bioquímica i Biologia Molecular

**Procarboxipeptidases humanes:
estudis funcionals i estructurals de formes
pancreàtiques i de noves formes reguladores**

**Sònia Segura Martín
Setembre 2002**



Universitat Autònoma de Barcelona

Departament de Bioquímica i Biologia Molecular

**Procarboxipeptidases humanes:
estudis funcionals i estructurals de formes
pancreàtiques i de noves formes reguladores**

Tesi presentada per Sònia Segura Martín, llicenciada en Bioquímica, per obtenir el grau de Doctora en Bioquímica.

Treball realitzat en el Departament de Bioquímica i Biologia Molecular de la Facultat de Ciències de la Universitat Autònoma de Barcelona, sota la direcció del Dr. Josep Vendrell Roca i el Dr. F.X. Avilés Puigvert.

Sònia Segura Martín

Dr. Josep Vendrell Roca

Dr. F.X. Avilés Puigvert

Bellaterra, 26 de setembre de 2002

Dedico aquest treball als meus pares, que tant m'han donat

Agraïments

Aquest treball ha estat possible gràcies a la disponibilitat i dedicació dels meus directors de tesi, el Dr. Josep Vendrell i el Dr. Francesc X. Avilés.

També vull expressar el meu agraïment als meus companys PROs, a la gent d'enzimologia i de biocomputació de l'IBB i als companys del departament.

ÍNDEX GENERAL

ÍNDEX	iii
Índex de figures i taules	vii
ABREVIATURES	ix
I.- INTRODUCCIÓ	1
I.1.- Les proteases	3
I.1.1.- Classificació de les proteases	3
I.1.2.- Classificació de les metal·loproteases	5
I.1.3.- Proproteases i proteòlisi limitada	6
I.2.- Les metal·locarboxipeptidases	7
I.2.1.- Classificació de les metal·locarboxipeptidases	7
I.2.1.1.- Subfamília de la carboxipeptidasa A1 (Subfamília A/B)	10
I.2.1.2.- Subfamília de la carboxipeptidasa E (Subfamília E/N)	12
I.2.2.- Estructura tridimensional i mecanisme catalític de les metal·locarboxipeptidases	15
I.2.3.- Evolució de les metal·locarboxipeptidases	18
I.3.- Procarboxipeptidases pancreàtiques	19
I.3.1.- Estructura tridimensional de les procarboxipeptidases	20
I.3.2.- Activació de les procarboxipeptidases pancreàtiques	26
I.3.2.1.- Procés d'activació de la PCPA1 porcina	26
I.3.2.2.- Procés d'activació de la PCPB porcina	27
I.3.2.3.- Procés d'activació de la PCPA2 humana	29
I.3.2.4.- Estudi preliminar del procés d'activació de les procarboxipeptidases humanes	29
I.4.- TAFI	30
OBJECTIUS	35
M.- MATERIAL I MÈTODES	41
M.1.- Aparells i productes químics	44
M.1.1.- Aparells	44
M.1.2.- Productes químics	45
M.2.- Mètodes de biologia molecular	45
M.2.1.- Material biològic i medis de cultiu	45
M.2.1.1.- Soques i vectors	45
M.2.1.2.- Medis de cultiu	47
M.2.2.- "Polimerase Chain Reaction" (PCR)	48
M.2.2.1.- Introducció de dianes de restricció	49
M.2.2.2.- Mutagènesi mitjançant extensió per solapament	50

M.2.3.- Seqüenciació de DNA	52
M.2.4.- Extracció de DNA plasmídic	52
M.2.4.1.- Procediment de la lisi alcalina	52
M.2.4.2.- Extracció amb kit comercial	53
M.2.5.- Determinació de la concentració de DNA	53
M.2.5.1.- Determinació espectrofotomètrica	53
M.2.5.2.- Estimació per electroforesi en gels d'agarosa	54
M.2.6.- Digestions enzimàtiques d'àcids nucleics	54
M.2.7.- Separació del DNA mitjançant gels d'agarosa	54
M.2.7.1.- Electroforesi de DNA	54
M.2.7.2.- Purificació de bandes de DNA de gels d'agarosa	56
M.2.8.- Clonatge de fragments de DNA en plàsmids	56
M.2.9.- Incorporació de DNA forani en bacteris	57
M.2.9.1.- Preparació de cèl·lules competents d' <i>E. coli</i>	57
M.2.9.2.- Transformació de cèl·lules competents d' <i>E. coli</i>	57
M.2.10.- Transformació de <i>P. pastoris</i> pel mètode dels esferoplastes	58
M.2.11.- Expressió de proteïna recombinant	59
M.2.11.1.- Producció a nivell analític	59
M.2.11.2.- Producció preparativa de proteïna recombinant	60
M.3.- Mètodes d'anàlisi de proteïnes	60
M.3.1.- Mètodes electroforètics	60
M.3.1.1.- Electroforesi discontinua en presència d'SDS	60
M.3.1.2.- Gels de tricina	63
M.3.2.- Immunotransferència	64
M.3.3.- Seqüenciació N-terminal	66
M.3.4.- Mètodes espectroscòpics	67
M.3.4.1.- Quantificació espectrofotomètrica de proteïnes	67
M.3.4.2.- Determinació espectroscòpica d'activitats enzimàtiques	67
M.3.5.- Espectrometria de masses	69
M.3.6.- Diàlisi	69
M.3.7.- Mètodes cromatogràfics	69
M.3.7.1.- Cromatografia atmosfèrica d'interacció hidrofòbica	69
M.3.7.2.- Cromatografia FPLC de bescanvi aniònic	70
M.3.7.3.- Cromatografia líquida d'alta pressió (HPLC)	70
M.3.8.- Anàlisi d'aminoàcids mitjançant dabsilació	71
M.3.8.1.- Dabsilació d'aminoàcids	71

M.3.8.2.- Anàlisi dels aminoàcids per RP-HPLC	72
M.4.- Modelat de proteïnes per homologia	72
M.4.1.- Elaboració dels alineaments múltiples	73
M.4.2.- Obtenció i avaluació dels models	73
1.- CAPÍTOL 1: Expressió de les procarboxipeptidases A1 i B humanes en <i>P. pastoris</i> i estudi detallat dels seus processos d'activació	75
1.1.- Summary	77
1.2.- Experimental procedures	78
1.2.1.- Cloning, expression in <i>P. pastoris</i> and purification of human PCPA1, PCPB and the D255K mutant	78
1.2.1.1.- Plasmid constructs	78
1.2.1.2.- PCR mutagenesis	78
1.2.1.3.- Transformation and selection of productive clones	79
1.2.1.4.- Expression and purification of recombinant human PCPA1, PCPB and the D255K mutant	80
1.2.2.- Activation studies of recombinant human PCPA1, PCPB and the D255K mutant	81
1.2.2.1.- Activation experiments	81
1.2.2.2.- Electrophoretic and N-terminal sequence analyses	81
1.2.2.3.- Chromatographic analysis by reverse-phase HPLC and mass spectrometry	81
1.2.2.4.- Amino acid analysis	82
1.3.- Results and discussion	82
1.3.1.- Construction of the expression vectors, expression in <i>P. pastoris</i> and purification of recombinant human pCPA1, pCPB and the D255K mutant	82
1.3.2.- Trypsin activation of recombinant human PCPA1	85
1.3.3.- Comparative study of the activation processes of recombinant human PCPB and its reversed-specificity variant, D255K	89
1.3.4.- Comparison of the tryptic activation processes of the members of the human procarboxypeptidase system	96
2.- CAPÍTOL 2: Estructura Tridimensional de la procarboxipeptidasa B humana i modelat de la TAFI i de tres nous membres de la família de metal·locarboxipeptidases humanes	97
2.1.- Summary	99
2.2.- Experimental procedures	100
2.2.1.- Resolution of the three-dimensional structure of human PCPB	100
2.2.1.1.- Crystallization and data collection	100

2.2.1.2.- Structure solution and refinement	100
2.2.2.- Database searches for new members of the human metallocarboxypeptidase gene family	102
2.2.3.- Model building of tafi and the three new members of the human metallo- carboxypeptidase family	103
2.2.3.1.- Template selection and alignments	103
2.2.3.2.- Building and evaluation of models	104
2.2.4.- Miscellaneous	104
2.3.- Results and discussion	105
2.3.A.- Three-dimensional structure of human PCPB and modelling of TAFI	105
2.3.A.1.- Structure of human PCPB	105
2.3.A.1.1.- Active-site architecture	107
2.3.A.2.- Modelling of TAFI	111
2.3.A.2.1.- Evaluation and refinement of the TAFI model	112
2.3.A.2.2.- Analysis of the TAFI model and comparison with the experimental PCP structures	112
2.3.B.- Identification and modelling of three new members of the human metallocarboxypeptidase family	115
2.3.B.1.- Database searches	115
2.3.B.2.- Alignments of the novel CPs	118
2.3.B.3.- Modelling of the novel CPs and prediction of their specificities	122
RESUM I DISCUSSIÓ GENERAL	129
CONCLUSIONS	139
BIBLIOGRAFIA	145

Índex de figures i taules

Figures

I.1	4	1.3	84
I.2	9	1.4	85
I.3	15	1.5	86
I.4	16	1.6	87
I.5	17	1.7	89
I.6	18	1.8	90
I.7	21	1.9	91
I.8	22	1.10	92
I.9	23	1.11	93
I.10	23	1.12	94
I.11	24	1.13	95
I.12	25	1.14	96
I.13	27		
I.14	28	2.1	106
I.15	29	2.2	107
I.16	30	2.3	109
I.17	32	2.4	110
		2.5	113
M.1	46	2.6	119
M.2	46	2.7	120
		2.8	121
1.1	79	2.9	125
1.2	83		

Taules

I.1	11
I.2	15
2.1	101
2.2	108
2.3	116
2.4	123

ABREVIATURES

Abreviatures

A	Adenosina
ACL _P	<i>Aortic carboxypeptidase-like protein</i>
ADEPT	<i>Antibody directed enzyme prodrug therapy</i>
AD	<i>Activation domain</i>
Ae	Activitat específica
AEBP	<i>Adipocyte enhancer-binding protein</i>
Aht	Absorbància hidròlisi total
α-MF	<i>α-Mating factor</i>
ATP	5´ trifosfat d´adenosina
AOX	Alcohol oxidasa
BGA	N-benzoil-L-glicil-L-arginina
BMGY	<i>Buffered Glycerol-complex Medium</i>
BMMY	<i>Buffered Methanol-complex Medium</i>
BPTI	Inhibidor de tripsina de pàncrees de bou
BSA	Seroalbúmina bovina
C	Citidina
Cbz	Benziloxicarbonil
cDNA	Àcid desoxirribonucleic complementari
CP	Carboxipeptidasa
DABS-Cl	Clorur de 4-(dimetilamino)azobenzè-4´-sulfonil
DEAE	Dietilaminoetil
DESY	<i>Deutsches elektronensynchrotron</i>
DNA	Àcid desoxirribonucleic
dNTPs	Desoxinucleòtids trifosfat
ddNTPs	Didesoxinucleòtids trifosfat
DO	Densitat òptica
DSSP	<i>Define secondary structure of proteins</i>
DTT	Ditiotreitol
EBI	<i>European bioinformatics institute</i>
EDTA	Tetracetat d´etilè diamina
ESRF	<i>European synchrotron radiation facility</i>
EST	<i>Expressed sequence tag</i>
FAPP	N-furilacrilòil-L-fenilalanil-L-fenilalanina
FPLC	<i>Fast protein liquid chromatography</i>
G	Guanosina
GPI	Glicosilfosfatidilinositol

Abreviatures

HIS4	Histidinol deshidrogenasa
HMM	<i>Hidden Markov model</i>
HPLC	<i>High performance liquid chromatography</i>
HSA	Albúmina de sèrum humà
IgG	Immunoglobulina G
kb	Kilobases
KDa	Kilodàltons
LB	Luria Bertani
mA	Miliamper
MALDI-TOF	<i>Matrix-assisted laser desorption time-of-flight</i>
MD	<i>Minimal dextrose medium</i>
MES	Àcid 2-[N-morfolino]etanesulfònic
MOPS	Àcid 3-[N-morfolino]propanesulfònic
Mr	Massa molecular relativa
MS	Espectrometria de masses
NCBI	<i>National center for biotechnology information</i>
p/p	pes/pes
PAGE	Electroforesi en gel de poliacrilamida
pb	Parells de bases
PC	Prohormona convertasa
PCR	Reacció en cadena de la polimerasa
PCP	Procarboxipeptidasa
PDB	Protein data bank
PEG	Polietilè glicol
PHD	<i>Profile fed neural network systems from HeiDelberg</i>
PMSF	Fluorur de fenilmetilsulfonil
PPE	Proteïnasa E
PVDF	Difluorur de polivinilidè
QTC	Quimotripsinogen C
RD	<i>Regeneration dextrose medium</i>
RMN	Ressonància magnètica nuclear
RMSD	<i>Root-mean-square deviation</i>
RNA	Àcid ribonucleic
rPCP	Procarboxipeptidasa recombinant
RP-HPLC	<i>Reversed-phase high performance liquid chromatography</i>
rpm	Revolucions per minut

Abreviatures

SDS	Dodecilsulfat sòdic (laurilsulfat sòdic)
SP	SwissProt
SSAP	<i>Sequential structure alignment program</i>
STI	Inhibidor de tripsina de soja
T	Timidina
TAE	Tampó Tris-acetat-EDTA
TAFI	<i>Thrombin-activatable fibrinolysis inhibitor</i>
TAFIa	Activated TAFI
TE	Tampó Tris-EDTA
TEMED	N, N, N', N'-tetrametilè diamí
TBE	Tampó Tris-borat-EDTA
TBS	Tampó Tris-salí
TFA	Àcid trifluoroacètic
T _m	Temperatura de fusió
TLCK	N-p-tosil-L-lisil-clorometil cetona
TPCK	N-p-tosil-L-fenilalanil-clorometil cetona
Tris	Tris-hidroximetilamímetà
Tricina	N-[2-hidroxi-1,1-bis(hidroximetil)etil]glicina
UDO	Unitats de densitat òptica
v/v	volum/volum
YEPD	<i>Yeast extract peptone dextrose medium</i>
w/w	<i>weight/weight</i>

INTRODUCCIÓ

I.1.- LES PROTEASES

Les proteases són enzims que catalitzen la hidròlisi d'enllaços peptídics de proteïnes i pèptids. El conjunt de les proteases engloba un gran nombre d'enzims (~2% dels productes gènics) amb funcions molt diverses i que afecten a diferents aspectes del metabolisme, variant des de la funció més puntual i específica, com el processament de prohormones i proneuropèptids, fins a la funció més exhaustiva i inespecífica, com és la degradació de les proteïnes provinents de la ingesta. Les proteases tenen un gran interès en medicina i biotecnologia i constitueixen dianes molt atractives per al disseny de fàrmacs.

Existeixen diversos termes redundants per referir-se als enzims proteolítics: proteasa, peptidasa i proteïnasa. Tot i que la Unió Internacional de Bioquímica i Biologia Molecular recomana l'ús del terme peptidasa, el terme proteasa és el que s'utilitza normalment.

El mecanisme catalític general pel qual una proteasa hidrolitza un enllaç peptídic consisteix en la polarització de l'enllaç peptídic mitjançant l'atac sobre el carboni del grup carbonil, el qual es pot produir directament per un grup nucleòfil o bé per acció d'una base a través d'una molècula d'aigua, i paral·lelament té lloc la donació d'un protó a l'àtom de nitrogen. Cada tipus de proteasa té uns residus específics que realitzen les funcions de nucleòfils i de donadors de protons.

I.1.1.- CLASSIFICACIÓ DE LES PROTEASES

Les proteases es poden classificar en base a diferents criteris: especificitat de substrat, mecanisme catalític, relacions estructurals i evolutives, localització intracel·lular, funció en l'organisme, etc.

Una de les classificacions de les proteases més usades és la descrita per Barret i Rawlings (1993), basada en el mecanisme catalític de les proteases. Aquests autors han creat una base de dades de proteases anomenada MEROPS (www.merops.ac.uk) (Rawlings i Barret, 1999; Rawlings *et al.*, 2002), on cada grup de proteases és dividit en famílies seguint criteris d'homologia estructural. Els sis grups de proteases que s'han definit són: serina proteases, treonina proteases, cisteïna proteases, aspàrtic proteases, metal·loproteases i proteases de mecanisme catalític desconegut.

El grup de les **serina proteases** engloba un conjunt d'enzims que es caracteritzen per presentar una serina reactiva en el centre actiu i el mecanisme catalític, així com l'estructura tridimensional del centre actiu, és molt semblant en totes les serina proteases estudiades. Les

diferències de funció que presenten són degudes a canvis en l'especificitat de substrat. El grup de les serina proteases es divideix en 38 famílies, cadascuna de les quals presenta un membre representatiu. La família de la quimotripsina (S1) és molt nombrosa i conté endoproteases com tripsina, elastasa, acrosina, les calicreïnes, els granzims, factors de coagulació (VII, IX, X, XI), activador del plasminogen, trombina, plasmina, proteïna C i factors del sistema del complement. Una altra família a destacar és la de les serina carboxipeptidases (S10), amb la carboxipeptidasa Y (CPY) de llevat com a enzim representatiu. En aquest grup trobem diverses carboxipeptidases presents en fongs, gramínies i dicotiledònies.

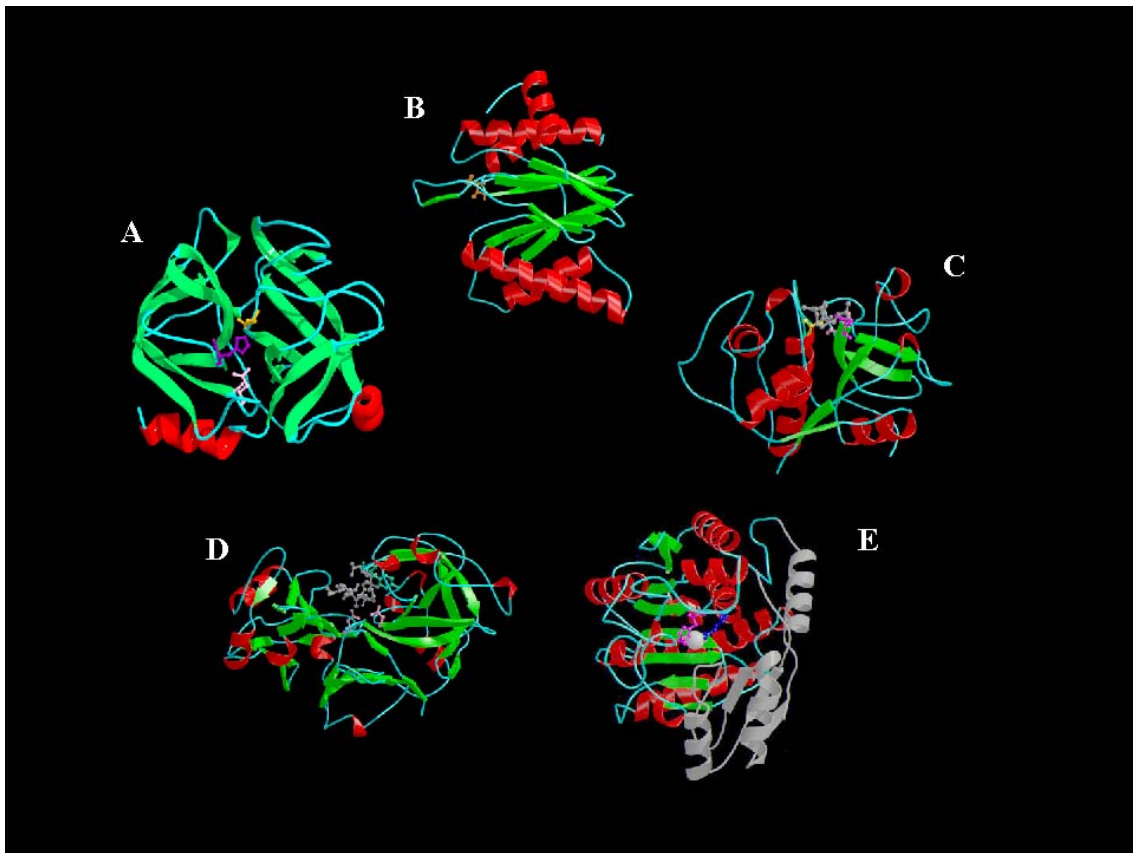


Figura I.1.- Diagrama de cintes de l'estructura d'un representant de cadascun dels cinc grups de proteases amb mecanisme catalític conegut, on es mostra en color les cadenes laterals dels diferents residus catalítics. A- Tripsina de rata (codi pdb: 1BRA), B- Subunitat catalítica del proteasoma de *Thermoplasma acidophilum* (codi pdb: 1PMA), C- Papaïna de *Carica papaya* en complex amb l'inhibidor E-64, representat en color gris (codi pdb: 1PE6), D- Pepsina A humana en complex amb l'inhibidor pepstatina, representat en color gris (codi pdb: 1PSO), E- Procarboxipeptidasa A1 porcina, on el segment pro i l'esfera que representa l'ió de zinc es mostren en color gris (codi pdb: 1PCA). Extret de MEROPS.

Les **treonina proteases** s'han descrit recentment (Seemuller *et al.*, 1995) i, de moment, és un grup de proteases poc nombrós. Només està format per quatre famílies, entre les que cal destacar la família T1, amb la subunitat catalítica del proteasoma de *Thermoplasma acidophilum* com a membre representatiu. Dins d'aquesta família trobem subunitats catalítiques tant del proteasoma procariota com de l'eucariota. Aquests enzims es caracteritzen per presentar una treonina en posició N-terminal que té la funció de nucleòfil del centre actiu.

Les **cisteïna proteases** s'anomenen també sulfhidril o tiol proteases i formen un grup de 50 famílies, la majoria de les quals corresponen a proteases de virus. El mecanisme catalític d'aquest grup s'ha caracteritzat en la família de la papaïna (C1) i s'ha observat que hi són involucrats un residu de cisteïna i un d'histidina. Altres famílies remarcables són la de les calpaïnes (C2) i la de les caspases (C14).

El grup de les **aspàrtic proteases** està constituït per 18 famílies. Aquests enzims també són coneguts com a proteases àcides o carboxil proteases. Són endoproteases que es caracteritzen per la presència d'un residu aspàrtic en el centre actiu, cosa que fa que el seu pH òptim sigui àcid. La família A1, amb la pepsina A humana com a membre representatiu és molt nombrosa i conté, a més de moltes pepsines de diferents espècies, la quimosina i la renina. La família A2 conté retropepsines del virus del SIDA de diverses espècies.

Les **metal·loproteases** són un conjunt molt variat d'enzims que es caracteritzen per presentar un ió metàl·lic en el centre actiu, que normalment és zinc. A partir de l'estudi de les estructures de diferents enzims pertanyents a aquest grup, s'ha observat que l'ió metàl·lic presenta tres lligands amb la cadena polipeptídica i un quart lligand que és una molècula d'aigua que s'activa per iniciar l'atac nucleofílic sobre l'enllaç a hidrolitzar. Es coneixen 51 famílies de metal·loproteases, algunes de les quals són endopeptidases, com les famílies de la termolisina (M4) i l'astacina (M12), i famílies d'exopeptidases, com la de la carboxipeptidasa A1 (M14) i la de l'aminopeptidasa N (M1).

En un sisè grup de proteases s'engloben aquells enzims amb mecanisme catalític desconegut. Dins d'aquest grup trobem les secretases, que són tres enzims, α , β i γ secretasa, responsables del processament de la proteïna precursora dels pèptids amiloides que es dipositen extracel·lularment en el cervell, causant la malaltia de l'Alzheimer (Dominguez *et al.*, 2001).

I.1.2.- CLASSIFICACIÓ DE LES METAL·LOPROTEASES

Rawlings i Barret (1995) van establir una classificació de les metal·loproteases basada en relacions estructurals i evolutives. Segons aquesta classificació, les diferents famílies de metal·loproteases es divideixen en quatre grups, dos dels quals contenen el motiu **HEXXH**,

que, tal i com s'ha demostrat per cristal·lografia de raigs X, forma part del lloc d'unió de l'ió metàl·lic, que generalment és zinc. Els quatre grups que s'han definit són:

- I) "**HEXXH + E**", en què els quatre lligands de l'ió de zinc són una molècula d'aigua, les dues histidines del motiu HEXXH i un residu d'àcid glutàmic. Dins d'aquest grup trobem 18 famílies, les quals corresponen en gran majoria a endopeptidases, com la família de la termolisina (M4). La família de l'aminopeptidasa N (M1), la de la carboxipeptidasa Taq de *Thermus aquaticus* (M32) i la de la glicil aminopeptidasa (M61) són famílies d'exopeptidases i la família de la oligopeptidasa timet (M3) és l'única d'entre totes les famílies de metal·loproteases que conté tant endo com exopeptidases.
- II) "**HEXXH + H/D**", que presenten l'ió de zinc coordinat amb una molècula d'aigua, dos residus d'histidina del motiu HEXXH i un tercer residu d'histidina o aspàrtic. Aquest grup està constituït per 12 famílies d'endopeptidases, entre les quals hi ha la família de l'astacina (M12) i la de la col·lagenasa I (M10).
- III) Aquest és un grup heterogeni que engloba diferents famílies de metal·loproteases que presenten un motiu d'unió a l'ió metàl·lic que és diferent de **HEXXH**. La majoria de les famílies que integren aquest grup presenten histidines com a lligand predominant de l'ió metàl·lic. Per altra banda, les famílies d'aminopeptidases que hi trobem es caracteritzen per presentar dos ions metàl·lics en el centre actiu, els quals estan coordinats majoritàriament per residus àcids, com és el cas de la família de la leucil aminopeptidasa (M17) i de la metionil aminopeptidasa (M24). En aquest últim cas, junt amb la família M52, ens trobem amb exemples de metal·loproteases que presenten en el centre actiu ions metàl·lics diferents del zinc, com són el cobalt o el manganès per a la família M24 i el níquel per a la família M52. Dins d'aquest tercer grup trobem la família M14, que conté les carboxipeptidases que són objecte d'estudi en aquest treball.
- IV) Aquest grup engloba totes les famílies (actualment un total de sis) per les quals els lligands del metall són encara desconeguts.

I.1.3.- PROPROTEASES I PROTEÒLISI LIMITADA

Moltes proteases són sintetitzades com a precursors inactius anomenats proenzims o zimògens per tal de modular la seva activitat, de manera que són activades només en el lloc i

moment adequat. Els zimògens són activats per un procés de proteòlisi limitada que implica la hidròlisi específica d'un o més enllaços per part d'una proteasa.

Els segments responsables de la inactivació de les proteases solen trobar-se en posició N-terminal. Aquests segments tenen una longitud que pot variar des d'uns pocs residus, com en el cas del tripsinogen, en què s'allibera un pèptid de 6 residus, fins als 560 residus que presenta el plasminogen, grandària major que la de la proteïna activa. El fet que existeixin segments d'activació amb una longitud considerable fa pensar que puguin tenir funcions secundàries, com passa amb el pèptid d'activació del tripsinogen, que produeix una supressió hormonal de la secreció gàstrica. D'altra banda, s'ha observat que en alguns casos els segments d'activació són essencials perquè la proteïna adopti un plegament correcte, de manera que se'ls atribueix la funció general de xaperones co-translacional (Sorensen *et al.*, 1993; Phillips i Rutter, 1996).

L'activació del quimotripsinogen, del pepsinogen i de les procarboxipeptidases constitueixen exemples de proteòlisi N-terminal, però també es dona proteòlisi activant en regions centrals de zimògens, com en el cas del factor IX de la coagulació, o en posició C-terminal, com per exemple en la proinsulina (Csorba, 1991). En alguns casos no s'allibera un pèptid, com succeeix amb la uroquinasa humana (Wun *et al.*, 1982), en què un únic tall proteolític implica la formació de dos fragments polipeptídics que es mantenen units per ponts disulfur.

Els mecanismes d'inhibició dels segments d'activació són diversos. Poden actuar com a elements d'impediment estèric per a l'accés al centre actiu, com passa en les procarboxipeptidases. En aquests casos el centre catalític està preformat als enzims i pot succeir que alguns proenzims presentin activitat intrínseca residual (Kassell i Kay, 1973; Guasch *et al.*, 1992). També s'ha observat que els segments d'activació poden provocar una distorsió en el centre actiu o en el centre de fixació de substrats, inhibint l'enzim, com en el cas del quimotripsinogen.

I.2.- LES METAL-LOCARBOXIPEPTIDASES

I.2.1.- CLASSIFICACIÓ DE LES METAL-LOCARBOXIPEPTIDASES

Les metal-locarboxipeptidases són enzims que hidrolitzen residus C-terminals d'altres proteïnes o pèptids i que presenten un ió de zinc directament involucrat en la catàlisi. Segons la classificació de les metal-loproteases que s'indica en l'apartat I.1.2, constitueixen la família M14, amb la carboxipeptidasa A1 com a enzim representatiu. Aquesta família es divideix en

dues subfamílies: la subfamília de la carboxipeptidasa A1 i la subfamília de la carboxipeptidasa E. El domini enzimàtic dels membres de la subfamília de la carboxipeptidasa A1 és d'uns 310 residus, mentre que en la subfamília de la carboxipeptidasa E té una grandària que és superior en més de 100 residus. D'altra banda, mentre tots els membres de la subfamília de la carboxipeptidasa A1 presenten segments d'activació de grandàries similars (75-76 residus per a les procarboxipeptidases de fongs i bacteris, 91-96 residus per a la resta), en la subfamília de la carboxipeptidasa E aquests propèptids són absents o bé tenen una mida petita (5-17 residus) i una funció desconeguda. La subfamília de la carboxipeptidasa E és la única que presenta membres amb diversos dominis estructurals, els quals solen estar glicosilats.

A part d'aquesta classificació de les metal·loproteases basada en relacions estructurals, s'han realitzat altres classificacions alternatives seguint diferents criteris. Així, les carboxipeptidases s'han classificat en formes A i B, segons l'especificitat que presenten pels residus C-terminals. Les formes A, en general, presenten preferència per residus C-terminals hidrofòbics, mentre que les formes B escindeixen residus bàsics. Es va observar que alguns organismes presentaven dues isoformes amb especificitat tipus A (Gardell *et al.*, 1988; Pascual *et al.*, 1989), que es van anomenar A1 i A2. La isoforma A1 correspon a la que anteriorment es coneixia com a forma A i la isoforma A2 és l'enzim que es va caracteritzar de nou, el qual mostra preferència per residus C-terminals hidrofòbics del tipus aromàtic.

També es poden classificar les metal·locarboxipeptidases d'acord amb la funció fisiològica que realitzen. Així, distingim entre les carboxipeptidases pancreàtiques o digestives, les quals estan involucrades en la degradació de les proteïnes durant la digestió, i les carboxipeptidases reguladores, que participen en processos biològics més específics, tals com el processament de pèptids bioactius. Aquestes últimes són anomenades també carboxipeptidases bàsiques, degut a l'especificitat que moltes presenten per a residus C-terminals bàsics, però no és un nom acurat, ja que no totes les carboxipeptidases reguladores presenten aquest tipus d'especificitat. Les carboxipeptidases (CP) pancreàtiques són les que es van descobrir primer i s'han estudiat en profunditat; dins d'aquest grup trobem la CPA1, la CPA2 i la CPB. També s'hi han d'incloure algunes carboxipeptidases que s'han caracteritzat en bacteris, fongs i insectes, que tenen com a funció hidrolitzar substrats per aconseguir nutrients. El grup de les carboxipeptidases reguladores, caracteritzades més recentment, està constituït per la resta de carboxipeptidases, i, en general, són més desconegudes quant a estructura i funció. Aquest grup inclou, per una part, la TAFI i la CPA3 (CPA de mastòcits), que, segons la classificació en base a l'estructura, formen part de la subfamília de la CPA1 i també són anomenats enzims "pancreatic-like". Tota la resta de carboxipeptidases reguladores pertanyen a la subfamília de la CPE i els enzims que ens trobem són la CPE, CPD, CPM, CPN, CPZ, AEBP, CPX1 i CPX2.

La lleu confusió que es pot percebre en l'anterior descripció de les classificacions de les carboxipeptidases s'ha volgut evitar darrerament mitjançant la identificació dels subgrups de metal·locarboxipeptidases a partir de la seva estructura, tant primària com terciària. Així, s'ha definit dos grans grups: el grup A/B, que engloba a totes les metal·locarboxipeptidases amb estructura semblant a les pancreàtiques i el grup E/N, que conté les que tenen una estructura en subdominis semblants a la CPE. Aquesta classificació estructural fa que convisquin, al mateix grup, enzims amb funcions "digestives" i "reguladores" (Wei *et al.*, 2002). Malgrat això, en aquesta introducció es continuarà utilitzant els termes clàssics.

	6	7	8	19	20	27	28
	0123456789012	123456789012	123456789012	901234567890123456	1234567890123456	234567890123456789012	34567890123456789012
<i>Carboxypeptidase A subfamily</i>		*	*		*		+
1	PAIWDLGIHSREWITQATGVWF	PAIWDLGIHSREWITQATGVWF	PAIWDLGIHSREWITQATGVWF	PKAFLSIHSYSQLLLYPY	PKAFLSIHSYSQLLLYPY	GIKYSFTFELRD-TGRYGFLLP	GIKYSFTFELRD-TGRYGFLLP
2	PAIWDLGIHSREWITQATGVWF	PAIWDLGIHSREWITQATGVWF	PAIWDLGIHSREWITQATGVWF	PKAFLSIHSYSQLLLYPY	PKAFLSIHSYSQLLLYPY	GIKYSFTFELRD-TGRYGFLLP	GIKYSFTFELRD-TGRYGFLLP
3	PAIWDLGIHSREWITQATGVWF	PAIWDLGIHSREWITQATGVWF	PAIWDLGIHSREWITQATGVWF	PKAFLSIHSYSQLLLYPY	PKAFLSIHSYSQLLLYPY	GIKYSFTFELRD-TGRYGFLLP	GIKYSFTFELRD-TGRYGFLLP
4	PAIWDLGIHSREWITQATGVWF	PAIWDLGIHSREWITQATGVWF	PAIWDLGIHSREWITQATGVWF	PKAFLSIHSYSQLLLYPY	PKAFLSIHSYSQLLLYPY	GIKYSFTFELRD-TGRYGFLLP	GIKYSFTFELRD-TGRYGFLLP
5	PAIWDLGIHSREWITQATGVWF	PAIWDLGIHSREWITQATGVWF	PAIWDLGIHSREWITQATGVWF	PKAFLSIHSYSQLLLYPY	PKAFLSIHSYSQLLLYPY	GIKYSFTFELRD-TGRYGFLLP	GIKYSFTFELRD-TGRYGFLLP
6	PAIWDLGIHSREWITQATGVWF	PAIWDLGIHSREWITQATGVWF	PAIWDLGIHSREWITQATGVWF	PKAFLSIHSYSQLLLYPY	PKAFLSIHSYSQLLLYPY	GIKYSFTFELRD-TGRYGFLLP	GIKYSFTFELRD-TGRYGFLLP
7	PAIWDLGIHSREWITQATGVWF	PAIWDLGIHSREWITQATGVWF	PAIWDLGIHSREWITQATGVWF	PKAFLSIHSYSQLLLYPY	PKAFLSIHSYSQLLLYPY	GIKYSFTFELRD-TGRYGFLLP	GIKYSFTFELRD-TGRYGFLLP
8	EVVMVESNIHAREWITAAATTYL	EVVMVESNIHAREWITAAATTYL	EVVMVESNIHAREWITAAATTYL	IKVYLAHHSYSQLLLYPY	IKVYLAHHSYSQLLLYPY	DIPYAVTYELPRDQWNGFLE	DIPYAVTYELPRDQWNGFLE
9	EVVLFVTCNMHAREHITVEMCLRI	EVVLFVTCNMHAREHITVEMCLRI	EVVLFVTCNMHAREHITVEMCLRI	IKVYLAHHSYSQLLLYPY	IKVYLAHHSYSQLLLYPY	KI-WSYTFEYMPKSSSPGFYPR	KI-WSYTFEYMPKSSSPGFYPR
10	EVVLYTALHAREHITVEMALYT	EVVLYTALHAREHITVEMALYT	EVVLYTALHAREHITVEMALYT	IKVYLAHHSYSQLLLYPY	IKVYLAHHSYSQLLLYPY	KI-FAFTFEMYPKSSSPGFYPR	KI-FAFTFEMYPKSSSPGFYPR
<i>Carboxypeptidase H subfamily</i>							
11	EFKYGIGNMHGNEAVGRELLIFL	EFKYGIGNMHGNEAVGRELLIFL	EFKYGIGNMHGNEAVGRELLIFL	FVLSANLHGGDLVANYPY	FVLSANLHGGDLVANYPY	SNCFEITVEL-----SCEKFFP	SNCFEITVEL-----SCEKFFP
12	EFKYGIGNMHGNEAVGRELLIFL	EFKYGIGNMHGNEAVGRELLIFL	EFKYGIGNMHGNEAVGRELLIFL	FVLSANLHGGDLVANYPY	FVLSANLHGGDLVANYPY	SNCFEITVEL-----SCDKFPN	SNCFEITVEL-----SCDKFPN
13	EFKYGIGNMHGNEAVGRELLIFL	EFKYGIGNMHGNEAVGRELLIFL	EFKYGIGNMHGNEAVGRELLIFL	FVLSANLHGGDLVANYPY	FVLSANLHGGDLVANYPY	TNCFEITVEL-----SCDKFFP	TNCFEITVEL-----SCDKFFP
14	EFKYGIGNMHGNEAVGRELLIFL	EFKYGIGNMHGNEAVGRELLIFL	EFKYGIGNMHGNEAVGRELLIFL	FVLSANLHGGDLVANYPY	FVLSANLHGGDLVANYPY	AQCFEITVELSCCKYPREERL	AQCFEITVELSCCKYPREERL
15	EFKYGIGNMHGNEAVGRELLIFL	EFKYGIGNMHGNEAVGRELLIFL	EFKYGIGNMHGNEAVGRELLIFL	FVLSANLHGGDLVANYPY	FVLSANLHGGDLVANYPY		

**Figura I.2.- Conservació de seqüència al voltant dels lligands del zinc i dels residus catalítics de la família de la CPA1 (M14). Els residus estan numerats d'acord amb la CPA1 bovina. Els asteriscs indiquen els lligands del zinc, i el símbol + indica el residu catalític Glu. Els residus idèntics a la seqüència de la CPA1 es mostren en blanc sobre negre. La clau de la numeració és: 1, CPA1 bovina; 2, CPA1 humana; 3, CPA2 de rata; 4, CPA2 humana; 5, CPB de rata; 6, CP d'*Astacus astacus*; 7, TAFI; 8, CP de *Simulium vittatum*; 9, CP de *Streptomyces capreolus*; 10, CPT de *Thermoactinomyces vulgaris*; 11, CPE humana; 12, CP de peix pulmonat; 13, CPN humana; 14, CPM humana; 15, CP de *Drosophila melanogaster*. (Rawlings i Barrett, 1995).
Nota: en la taula s'utilitza la notació H per a referir-se a la carboxipeptidasa E.**

Les carboxipeptidases tenen aplicacions biotecnològiques i biomèdiques. S'han utilitzat per a produir enzims amb especificitat inversa (Edge *et al.*, 1998) o per generar complexos enzim-anticòs amb l'objectiu d'utilitzar-los en l'anomenada ADEPT (*Antibody Directed Enzyme Prodrug Therapy*), una teràpia basada en l'ús d'enzims dirigits per anticossos capaços d'activar una prodroga en el lloc adequat (Smith *et al.*, 1997; Wolfe *et al.*, 1999; Wright i Rosowsky, 2002).

I.2.1.1.- Subfamília de la carboxipeptidasa A1 (o subfamília A/B)

A més de les carboxipeptidases pancreàtiques CPA1, CPA2 i CPB, i la TAFI, que descriurem posteriorment amb més detall, aquesta subfamília conté les carboxipeptidases que es comenten seguidament.

La **carboxipeptidasa A3**, o carboxipeptidasa A de mastòcits, es troba en els grànuls de secreció dels mastòcits, on forma complexos macromoleculars amb proteoglicans. És una proteïna que hidrolitza substrats peptídics en tàndem amb altres proteases (Reynolds *et al.*, 1989a). La funció precisa d'aquest enzim no es coneix, però ha d'estar relacionada amb els processos de resposta al·lèrgica, asmàtics o inflamatoris, en què els mastòcits estan implicats. Es sintetitza en forma de proenzim, amb un segment d'activació de 94 residus i un domini enzimàtic de 308 residus. El segment d'activació presenta dos llocs d'N-glicosilació i el tall proteolític que l'allibera del proenzim té lloc sobre un residu àcid, a diferència de les formes pancreàtiques, en què el tall activant té lloc sobre un residu bàsic. S'ha observat que el processament del zimogen a forma activa té lloc en els grànuls de secreció (Rath-Wolfson, 2001). L'enzim generat presenta activitat de tipus A, tot i que presenta un percentatge d'identitat superior respecte la forma B.

La **carboxipeptidasa A4** s'ha identificat recentment en humà en cèl·lules canceroses de pròstata. S'ha observat que l'expressió del seu gen és induïda per inhibidors de la histona desacetilasa, cosa que indica que la inducció d'aquest gen ve mediada per la via d'hiperacetilació de les histones (Huang *et al.*, 1999). Aquest enzim és sintetitzat en forma proenzimàtica, amb un segment pro que, segons la comparació amb altres procarboxipeptidases, té una longitud de 95 residus i un domini enzimàtic de 310 residus.

Recentment s'han identificat i caracteritzat carboxipeptidases de diversos insectes. En el cas del lepidòpter *Helicoverpa armigera* (Lepidoptera: *Noctuidae*), una plaga del cotó, s'ha caracteritzat una carboxipeptidasa, **CPA_H**, amb activitat tipus A (Bown *et al.*, 1998). Aquest enzim, expressat en intestí, està involucrat en la digestió de nutrients. Es sintetitza en forma proenzimàtica amb un segment pro de 91 residus i un domini actiu de 318 residus. La seva estructura (codi pdb: 1JQG) ha estat resolta recentment (Estébanez-Perpiñá *et al.*, 2001) i s'observa que, tot i tenir un nivell d'identitat baix amb els altres membres de la seva subfamília (al voltant d'un 30%), presenten estructures tridimensionals semblants. D'altra banda, s'ha identificat una altra carboxipeptidasa en els dípters hematòfags *Simulium vittatum* (Diptera: *Simuliidae*) (Ramos *et al.*, 1993), *Anopheles gambiae* (Diptera: *Culicidae*), mosquit de la malària africana, (Edwards *et al.*, 1997) i *Aedes aegypti* (Diptera: *Culicidae*), mosquit de la febre groga, (Edwards *et al.*, 2000) que presenta un nivell d'identitat entre aquestes tres

espècies superior al 60%, pel que podria correspondre a la mateixa forma enzimàtica. En contrast, el nivell d'identitat entre aquesta forma i la CPAHa és d'un 30%. L'enzim actiu té 311 residus i és sintetitzat amb un segment pro de 95 residus. Aquest enzim presenta un % d'identitat amb carboxipeptidases pancreàtiques d'un 40%, valor superior al que presenta la CPAHa.

La **MeCPA** és una carboxipeptidasa del fong entomopatogen *Metarhizium anisopliae*. Aquest enzim és secretat en els teixits infectats del seu hoste, l'insecte *Manduca sexta*, on, juntament amb altres proteases del tipus endoproteasa, s'ocupa de procurar nutrients per al fong paràsit (Joshi i St. Leger, 1999). La MeCPA és sintetitzada en forma proenzimàtica, amb un segment pro de 76 residus i un domini enzimàtic de 319 residus amb un % d'identitat amb les formes A pancreàtiques del 40%. L'enzim actiu presenta especificitat A1+A2, és a dir, és capaç d'hidrolitzar tant residus C-terminals alifàtics com aromàtics. A diferència de les carboxipeptidases bacterianes, no presenta activitat tipus B.

	CPA1	CPA2	CPB	TAFI	CPA3	CPA4	CPAHa	CPT	CPE
CPA1	100	66	47	42	41	59	28	27	14
CPA2	66	100	46	43	42	68	32	31	14
CPB	47	46	100	48	54	47	26	28	13
TAFI	42	43	48	100	48	42	26	28	17
CPA3	41	42	54	48	100	43	22	28	6
CPA4	59	68	47	42	43	100	30	27	16
CPAHa	28	32	26	26	22	30	100	24	13
CPT	27	31	28	28	28	27	24	100	19
CPE	14	14	13	17	6	16	13	19	100

Taula I.1.- Percentatges d'identitat seqüencial entre diverses carboxipeptidases de la subfamília de la CPA1 i la CPE. En tots els casos s'han considerat les seqüències humanes, excepte per a la CPAHa, que és un enzim d'*Helicoverpa armigera*, i per a la CPT, que és l'enzim de *Thermoactinomyces vulgaris*.

La **carboxipeptidasa T** és un enzim extracel·lular que s'ha trobat en bacteris com *Thermoactinomyces vulgaris* (Smulevitch *et al.*, 1991) i *Streptomyces griseus* (Narahashi, 1990). Té la característica de processar residus C-terminals bàsics i hidrofòbics amb eficiències comparables. Es sintetitza en forma proenzimàtica, amb un segment pro de 75 residus i un domini enzimàtic de 326 residus. Constitueix un homòleg remot de les carboxipeptidases de mamífers, d'aquí que presenti un nivell d'identitat baix en relació amb les altres carboxipeptidases de la seva subfamília (al voltant d'un 30%). Tot i aquesta baixa identitat

seqüencial, s'ha observat que l'estructura tridimensional de la CPT (codi pdb: 1OBR) és molt similar a la de la CPA1 (Teplyakov *et al.*, 1992).

Finalment convé comentar que, arrel de la seqüenciació dels genomes d'alguns organismes, s'han identificat gens que presenten homologia seqüencial amb la procarboxipeptidasa A1, però la potencial proteïna que se'n derivaria encara no ha estat identificada ni caracteritzada. Com a exemple, dins d'aquest grup trobem diverses hipotètiques procarboxipeptidases de *Drosophila melanogaster* (p.ex, número accés TrEMBL Q9VL86 i O61532) i una de *Saccharomyces cerevisiae* (número accés SwissProt P38836).

I.2.1.2.- Subfamília de la carboxipeptidasa E (o subfamília E/N)

La **carboxipeptidasa E**, també coneguda com a CPH, actua en el processament d'hormones peptídiques i neurotransmissors (Fricker *et al.*, 1989). La CPE es troba en teixits neuroendocrins, sobretot pituitària i cervell. Presenta un propèptid de tan sols 17 residus que no afecta a l'activitat de l'enzim, un domini enzimàtic tipus carboxipeptidasa d'uns 300 residus, un domini homòleg a la transtiretina d'uns 80 residus i una hèlix amfipàtica d'ancoratge a membrana de 25 residus. La CPE presenta especificitat pels aminoàcids C-terminals bàsics lisina, arginina i histidina i és regulada per pH (Greene *et al.*, 1992), de manera que és inactiva a pH neutre (reticle endoplasmàtic i aparell de Golgi) i té una màxima activitat a pH 5-5,5 (vesícules secretores madures). En els ratolins diabètics obesos amb la mutació *Cpe^{fat}/Cpe^{fat}* s'ha demostrat que presenten una deficiència en el processament de la proinsulina, degut a una mutació puntual Ser/Pro que es dona en la CPE, quedant la seva activitat completament suprimida (Fricker *et al.*, 1996). En aquests ratolins mutants es va observar la presència de petites quantitats de pèptids processats, cosa que indicava que una altra carboxipeptidasa contribuïa al processament de pèptids neuroendocrins. La recerca de nous enzims va conduir a la identificació de la carboxipeptidasa D.

La **carboxipeptidasa D** va ser descoberta simultàniament per tres grups. El fet que s'observessin petites quantitats de pèptids processats correctament en ratolins *Cpe^{fat}/Cpe^{fat}* implicava que havia d'existir una altra carboxipeptidasa actuant en la via secretora. Els estudis realitzats (Song i Fricker, 1995b) van mostrar que aquesta activitat carboxipeptidasa estava unida a membrana i posteriorment es va purificar aquest nou enzim, que presenta una grandària de 180 kDa. D'altra banda, la CPD també es va identificar en fetge d'ànec com una proteïna de 180 kDa que s'uneix al virus de l'hepatitis B (Kuroki *et al.*, 1994; Eng *et al.*, 1998) i com a responsable de la mutació *Silver* en *Drosophila melanogaster* (Settle *et al.*, 1995). L'any 1997 es va clonar la CPD humana (Tan *et al.*, 1997) i es va observar que conté tres

dominis semblants a la CPE seguits d'una regió transmembranal i una cua citosòlica de 58 residus. La CPD es troba en tots els teixits examinats, però en cervell i pituitària és on més abunda. El pH òptim d'activitat de la CPD és de 5,5 a 7 (Song i Fricker, 1995a). Aquest rang ampli és degut als pH òptims solapants del primer i segon domini carboxipeptidasa. El tercer domini és inactiu, ja que li manquen els residus catalítics essencials (Companys *et al.*, 2001). Recentment s'ha resolt l'estructura del domini II de la CPD d'ànec (codi pdb: 1QMU) (Gomis-Rüth *et al.*, 1999). La funció més important de la CPD sembla ser el processament de proteïnes i pèptids que transiten per la via secretora. La seva funció també inclou el processament de pèptids neuroendocrins, però la CPD no pot compensar completament la manca de CPE en ratolins *Cpe^{fat}/Cpe^{fat}*.

La **carboxipeptidasa M** és una glicoproteïna que va ser purificada i caracteritzada de microvellositats placentàries (Skidgel, 1989). És present en una gran varietat de teixits i cèl·lules, però abunda especialment en placenta i pulmons. Es troba unida a membrana gràcies a una àncora de glicosilfosfatidilinositol a la regió C-terminal, com demostra el seu alliberament de la membrana després de tractar amb una fosfolipasa específica per a fosfatidilinositol (Skidgel *et al.*, 1991). La carboxipeptidasa M és un enzim de 426 residus sintetitzat sense segment d'activació i amb especificitat front residus C-terminals bàsics. La seva localització en membrana indica que podria estar implicada en el control de l'activitat d'hormones peptídiques en la superfície cel·lular, en la modulació de receptors i en el processament o degradació de proteïna extracel·lular. Concretament, s'ha observat que la carboxipeptidasa M podria estar implicada en el metabolisme de factors de creixement degut a la seva capacitat de generar des-Arg53-EGF i a la coincidència de la distribució de l'enzim i el factor de creixement epidèrmic (McGuire i Skidgel, 1995).

La **carboxipeptidasa N** es sintetitza al fetge sense segment pro i s'allibera a la sang, on és present en elevades concentracions. Se la coneix també pels noms de lisina carboxipeptidasa (CPK), inactivador d'anafilotoxina i carboxipeptidasa B de sèrum. Presenta especificitat per residus C-terminals bàsics i en humans és una glicoproteïna que consta de quatre subunitats no associades covalentment: dues subunitats catalítiques de 438 residus cadascuna (50-55 kDa) i dues subunitats glicosilades de 83 kDa sense activitat (Erdös, 1979). Les subunitats no catalítiques estableixen les subunitats actives front la temperatura corporal i les manté en circulació evitant la seva eliminació de la sang per filtració glomerular. Aquesta proteasa està implicada en la protecció de l'organisme contra pèptids amb potent acció que van a parar al torrent sanguini, com les quinines i les anafilotoxines.

La **carboxipeptidasa Z** s'ha clonat i expressat en baculovirus (Song i Fricker, 1997). Tant la forma humana com la de rata contenen un domini N-terminal de 100 aminoàcids altament conservat que presenta un 20-30% d'identitat amb el domini "frizzled" que es troba en

proteïnes que interaccionen amb Wnt, una proteïna involucrada en la polaritat de teixits en les primeres fases d'embriogènesi. La carboxipeptidasa Z abunda en la placenta i també es troba, encara que en quantitats menors, en cervell, pulmons, timus i ronyó.

L' **AEBP**, *adipocyte enhancer-binding protein*, és un repressor transcripcional amb activitat carboxipeptidasa, l'expressió del qual és suprimida durant la diferenciació dels adipòcits (He *et al.*, 1995). S'ha observat que l'AEBP és activat a través d'un nou mecanisme, en què la unió a DNA estimula l'activitat de l'enzim (Muisse and Ro, 1999). Aquest enzim, a l'igual que la CPX1 i la CPX2 consta d'un domini similar a discoïdina i un domini carboxipeptidasa. S'ha descrit una isoforma de l'AEBP, anomenada ACLP (*aortic carboxypeptidase-like protein*), que presenta una extensió N-terminal de 380 aminoàcids. L'expressió d'aquesta isoforma augmenta durant la diferenciació de les cèl·lules del múscul llis de l'aorta (Layne *et al.*, 1998).

Les **carboxipeptidases X1 i X2** s'han identificat recentment en ratolí (Lei *et al.*, 1999; Xin *et al.*, 1998). Ambdues proteïnes contenen una regió N-terminal de 160 aminoàcids seqüencialment similar al domini de discoïdina d'altres proteïnes. Alguns residus importants per a l'activitat catalítica d'altres carboxipeptidases no es troben conservats en aquestes, característica que comparteixen amb el domini III de la CPD. Així, la carboxipeptidasa X1 no té activitat front substrats estàndard per a carboxipeptidases, mentre que per a la carboxipeptidasa X2 no s'ha detectat cap activitat. La carboxipeptidasa X1 s'ha trobat en molts teixits embrionaris, cosa que suggereix que la seva funció deu estar lligada al desenvolupament, possiblement mitjançant interaccions cel·lulars a través del seu domini de discoïdina. La carboxipeptidasa X2 té una àmplia distribució en el cervell, on probablement té una funció com a proteïna d'unió.

S'han trobat membres d'aquesta subfamília en organismes invertebrats, tals com el mol·lusc marí *Aplysia californica*, on s'ha detectat la presència d'enzims CPE-like (Juvvadi *et al.*, 1997) i CPD-like (Fan *et al.*, 1999) i l'insecte *Manduca sexta* (Stone *et al.*, 1994). En el bacteri *Porphyromonas gingivalis* s'ha caracteritzat una carboxipeptidasa que presenta homologia amb aquesta subfamília (Masuda *et al.*, 2002), però la funció que se li atribueix és la d'obtenir nutrients. D'altra banda, s'han detectat seqüències de DNA amb homologia respecte la CPE en diferents organismes, com *Caenorhabditis elegans* (número accés EMBL U80448 i Z81054) o *Pseudomonas aeruginosa* (número accés EMBL AE004730).

	CPE	CPD-II	CPM	CPN	CPA1
CPE	100	47	38	47	14
CPD-II	47	100	42	46	12
CPM	38	42	100	38	14
CPN	47	46	38	100	10
CPA1	14	12	14	10	100

Taula I.2.- Percentatge d'identitat seqüencial entre les carboxipeptidases dels membres de la subfamília de la CPE i la CPA1. En tots els casos s'han utilitzat les seqüències humanes. Per a la CPD només s'ha considerat el segon domini i per a la CPN la cadena catalítica.

I.2.2.- ESTRUCTURA TRIDIMENSIONAL I MECANISME CATALÍTIC DE LES METAL-LOCARBOXIPEPTIDASES

La CPA1 bovina és un dels enzims millor caracteritzats. S'ha determinat l'estructura cristal·logràfica de l'enzim aïllat i en complex amb inhibidors (Quioco i Lipscomb, 1971; Rees *et al.*, 1983; Christianson i Lipscomb, 1989) (codis pdb: 3CPA, 4CPA, 5CPA). Es tracta d'un enzim amb forma globular, que presenta una fulla β central de vuit cadenes paral·leles i antiparal·leles, amb un gir de 120° entre la primera i la darrera cadena, sobre la qual s'empaqueten vuit hèlixs α laterals (Figura I.3). El centre actiu es troba dins una cavitat formada per la fulla β , dues hèlixs α i una regió irregular que la cobreix parcialment.

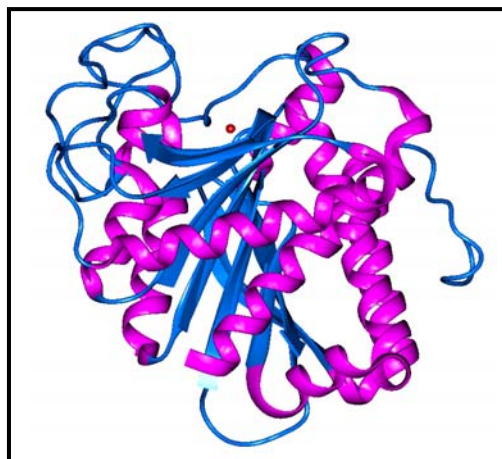


Figura I.3.- Diagrama de cintes de l'estructura de la CPA1 bovina. El punt vermell simbolitza l'ió de zinc.

Les carboxipeptidases mostren activitat peptidàsica i esteràsica i segueixen el mecanisme catalític anomenat "via promoguda per l'aigua" (Phillips *et al.*, 1990; Kim i Lipscomb, 1991; Álvarez-Santos *et al.*, 1994) (Figura I.4). Segons aquest mecanisme, el zinc del centre actiu polaritza la molècula d'aigua amb la que està coordinat, i aquesta ataca l'enllaç peptídic a hidrolitzar originant un intermediari tetraèdric. El Glu270 i l'Arg127 (numeració estàndard de la CPA1 bovina) col·laboren en aquest atac mitjançant la formació d'un pont d'hidrogen amb l'aigua i l'estabilització de l'oxoanió de l'intermediari tetraèdric, respectivament. La interacció del zinc amb l'aigua origina un hidroxil que actua com a potent nucleòfil i l'Arg127, junt amb l'ió metàl·lic, estabilitza les càrregues negatives. Aquest mecanisme funciona tant per a l'activitat peptidàsica com per a l'esteràsica, amb diferents passos limitants.

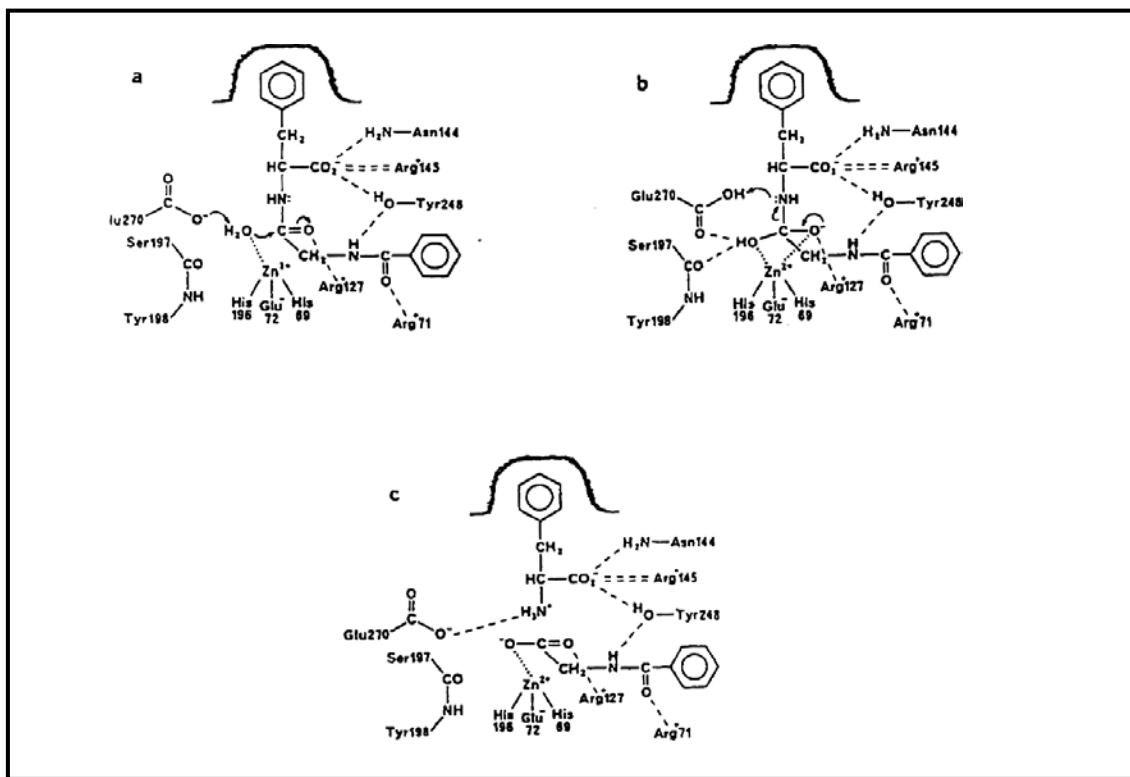


Figura I.4.- Esquema del mecanisme catalític de la carboxipeptidasa A1 bovina segons Kim i Lipscomb (1991). a) El complex precatalític format per l'oxigen carbonílic del substrat i l'Arg127 permet l'atac nucleofílic d'una molècula d'aigua amb l'ajuda de l'ió de zinc i del Glu270. b) L'intermediari tetraèdric estabilitzat prèn el protó del Glu270. c) Complex final de la reacció, després d'una segona transferència de protons pel Glu270.

Gràcies als estudis cristal·logràfics, estudis cinètics amb substrats de diferent longitud i composició i estudis de mutagènesi dirigida (Gardell *et al.*, 1985; Phillips *et al.*, 1990), s'han identificat els residus del centre actiu i els responsables de la fixació del substrat. Així, s'ha

observat que els residus d'unió al zinc són: His69, Glu72 i His196 i s'han definit una sèrie d'aminoàcids importants per a la catàlisi i unió a substrat, els quals són classificats en diversos subsetis (Figura I.5), els residus principals dels quals són els següents:

- S1': Asn144, Arg145 i Tyr248.
- S1: Arg127 i Glu270.
- S2: Arg71, Ser197, Tyr198 i Ser199.
- S3: Phe279.
- S4: Glu122, Arg124.

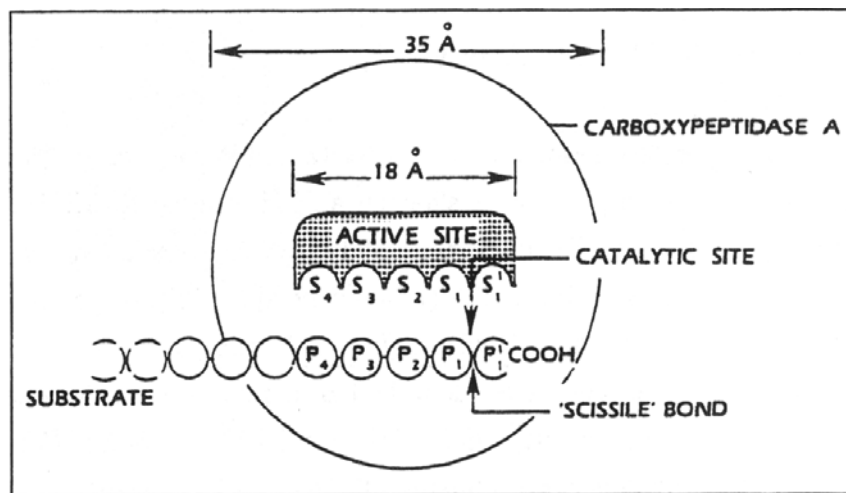


Figura I.5.- Esquema del centre actiu de la carboxipeptidasa A1. Es mostren els cinc subsetis (S_i) deduïts a partir dels estudis cinètics utilitzant substrats de diferent longitud (P_i) i estructura. Reproduït de McDonald (1985).

El grup carboxilat terminal del substrat peptídic és fixat per l'Asn144, Arg145 i Tyr248 i el grup carbonil de l'enllaç peptídic a hidrolitzar està prop de Glu270, Arg127 i el zinc. Tots aquests residus estan conservats en les carboxipeptidases digestives i molts també ho estan en les reguladores. Les diferències més importants es troben en els aminoàcids que defineixen l'especificitat de substrat o que es troben en regions superficials. Totes les carboxipeptidases presenten un motiu d'unió al zinc del tipus **HXXE + H**, on els quatre lligands del zinc són una histidina, un glutamat situat tres residus més enllà de la histidina, una altra histidina localitzada 108 – 135 residus cap al C-terminal de la primera histidina i una molècula d'aigua.

En la CPB porcina, els residus implicats en l'activitat catalítica es conserven espacialment respecte la CPA1 bovina (Coll *et al.*, 1991) (codi pdb:1NSA), així com en la CPA2 humana

(García-Sáez *et al.*, 1997) (codi pdb: 1AYE). En aquest últim cas les diferències es troben en el segment flexible que forma el centre fixador de substrats. S'ha resolt l'estructura cristal·logràfica del complex de la CPA2 humana amb l'LCI (*Leech Carboxypeptidase Inhibitor*) (Reverter *et al.*, 2000) (codi pdb: 1DTD). També s'han determinat les estructures tridimensionals de la CPB bovina (codi pdb: 1CPB) i la CPT de *Thermoactinomyces vulgaris* (codi pdb: 1OBR).

I.2.3.- EVOLUCIÓ DE LES METAL-LOCARBOXIPEPTIDASES

La família M14 conté membres procariotes i eucariotes i s'ha pogut elaborar un arbre filogenètic gràcies a la disponibilitat de suficients seqüències (Rawlings i Barret, 1995). A partir d'aquest, ha estat possible determinar que la divergència entre els membres procariotes i eucariotes de la família es va produir fa aproximadament 1344 milions d'anys. Com que la divergència dels regnes es va produir fa uns 3000 milions d'anys, es dedueix que va existir una transferència horitzontal de gens de carboxipeptidases procariotes cap a cèl·lules eucariotes, mitjançant la incorporació d'organismes endosimbiòtics, que donaren lloc als orgànuls en els eucariotes primitius.

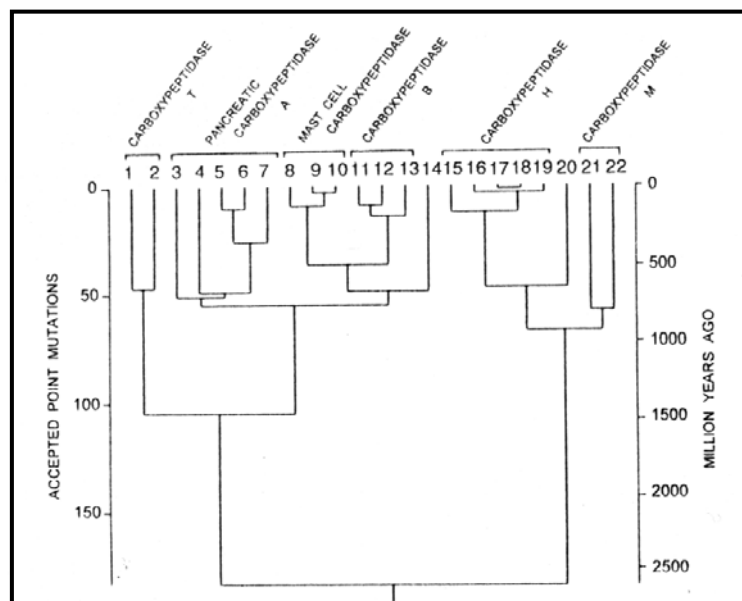


Figura I.6.- Arbre filogenètic de la família de la CPA1. 1 i 2, CPT de *T.vulgaris* i d' *S.griseus*; 3, CP d' *Astacus astacus*; 4, CP de *Simulium vittatum*; 5-7, CPA1 humana, bovina i de rata; 8-10, CPA3 humana, de rata i de ratolí; 11-13, CPB bovina, humana i de rata; 14, TAFI humana; 15-19, CPE de peix pulmonat, humana, de rata, de ratolí i bovina; 20, CPN humana; 21, CPE de *Drosophila melanogaster*; 22, CPM humana. Extret de Rawlings i Barret (1995).

Com ja s'ha dit abans, el nivell d'identitat seqüencial entre la subfamília de la CPA1 i la de la CPE és baix, però s'observa una conservació dels residus importants per a la catàlisi, com: His69, His169 i Glu72, que coordinen el zinc; Arg145, necessària per a l'ancoratge del substrat i Glu270, Arg127, Arg71 i Tyr248 als subestis del centre actiu (numeració de la carboxipeptidasa A1 bovina). Les diferències de grandària entre les carboxipeptidases són degudes als dominis d'unió a membrana o a altres dominis addicionals o repetits. S'ha proposat com a arbre evolutiu de les metal·locarboxipeptidases el que es mostra en la figura I.6.

I.3.- PROCARBOXIPEPTIDASES PANCREÀTIQUES

Fins ara, s'han descrit tres tipus de carboxipeptidases pancreàtiques: A1, A2 i B. La CPA1 hidrolitza preferentment residus C-terminals alifàtics de cadena llarga (Anson, 1935), mentre que la CPA2 té preferència per residus aromàtics voluminosos com el triptòfan (Gardell *et al.*, 1988; Pascual *et al.*, 1989; Catasús *et al.*, 1995). La CPB presenta afinitat per residus bàsics. En el procés de digestió dels vertebrats, la CPA1 continua la degradació dels productes de la hidròlisi quimotríptica i la CPB dels de la hidròlisi trípica (Puigserver *et al.*, 1986). Les carboxipeptidases pancreàtiques es troben en forma soluble i no glicosilada i es sintetitzen com a precursors inactius amb un segment d'activació d'uns 94-96 residus. Són estables en la forma de zimògens mentre estan emmagatzemats en els grànuls pancreàtics i fins el moment de la secreció. Posteriorment, mitjançant proteòlisi limitada, la tripsina genera enzims actius de 305-309 residus (Avilés *et al.*, 1993) que contribueixen a la degradació de les proteïnes alimentàries en l'intestí. Gràcies a aquesta clara compartimentalització de l'activitat, aquests proenzims poden servir de marcadors de diverses patologies. Així, per exemple, la procarboxipeptidasa B (PCPB) trobada en sèrum constitueix un marcador de pancreatitis aguda (Yamamoto *et al.*, 1992), i també de deficiència renal crònica (Chen *et al.*, 1996).

En mamífers, la PCPA1 pot trobar-se en forma monomèrica, formant complexos binaris amb la proproteïnasa E (PPE) o amb el quimotripsinogen C (QTC), o bé formant un complex ternari amb ambdues serina proteases. Les formes B, en canvi, apareixen tan sols com a monòmers. El complex ternari (PCPA1/PPE/QTC) s'ha descrit únicament en el sistema boví (Pascual *et al.*, 1990; Ventura *et al.*, 1997) i s'ha resolt la seva estructura tridimensional per difracció de raigs X (Gomis-Rüth *et al.*, 1997) (codi pdb: 1PYT). Existeixen diverses hipòtesis sobre quin és el paper biològic de les estructures quaternàries de les PCPA1: control i potenciació de l'activació, modulació de la activitat, protecció dels enzims front el pH àcid del suc pancreàtic i coordinació de l'aparició de les diferents activitats proteolítiques.

El sistema de les PCPs pancreàtiques humanes ha estat caracteritzat detalladament pel nostre grup (Pascual *et al.*, 1989). A partir d'extractes aquosos de polvoritzat pancreàtic humà s'han aïllat tres PCPAs (A1, A2 i A3) i dues PCPBs (B1 i B2). La PCPA1 i la PCPA2 s'obtenen en forma monomèrica, mentre que l'espècie PCPA3 és un complex binari format per la PCPA1 i la propteïnasa E. La PCPB1 i la PCPB2 corresponen a dues formes al·loestomèriques del mateix enzim. Més tard es van aïllar els cDNAs de les diferents procarboxipeptidases humanes i es va determinar les seves seqüències nucleotídiques i les aminoacídiques derivades (Catasús *et al.*, 1992; Catasús *et al.*, 1995; Aloy *et al.*, 1998) (Figures I.7 i I.8).

I.3.1.- ESTRUCTURA TRIDIMENSIONAL DE LES PROCARBOXIPEPTIDASES

S'ha determinat l'estructura cristal·logràfica de la PCPB porcina (Coll *et al.*, 1991) (Figura I.9) (codi pdb: 1NSA). També s'ha resolt l'estructura del domini globular del segment d'activació de la PCPB porcina en dissolució mitjançant RMN (Vendrell *et al.*, 1990a; 1991) (codi pdb: 1PBA). De la mateixa manera, per a la PCPA1 porcina s'ha resolt l'estructura cristal·logràfica del proenzim sencer (Guasch *et al.*, 1992) (Figura I.10) (codi pdb: 1PCA) i s'ha fet estudis estructurals del seu domini d'activació per RMN (Avilés *et al.*, 1985; Vendrell *et al.*, 1990b). L'estructura cristal·lina de la PCPA2 humana s'ha determinat pel proenzim sol (Figura I.11) (codi pdb: 1AYE) i en complex amb l'inhibidor benzilsuccínic (García-Sáez *et al.*, 1997). També es coneix l'estructura tridimensional de la PCPA1 bovina en el complex ternari amb la propteïnasa E i el quimotripsinogen C (Gomis-Rüth *et al.*, 1997) (codi pdb: 1PYT).

Tot i que els segments d'activació de les diferents formes proenzimàtiques presenten un nivell d'identitat baix, la seva disposició espacial és molt similar. Tots presenten una regió globular amb una topologia antiparal·lela α /antiparal·lela β en forma de sandvitx obert. La fulla β té una posició central mirant a la carboxipeptidasa per un costat i amb dues hèlixs en el costat oposat. En el cas de la PCPB existeix una hèlix 3_{10} que se situa al costat d'una cadena β en direcció a l'enzim. El segment que connecta la regió globular amb l'enzim comprèn els residus 82A-99A de la PCPA1 (18 residus) i està estructurat en una hèlix de quatre girs més un llaç de 4 residus; en la PCPA2, aquesta hèlix és una volta més llarga (residus 82A-101A). En la PCPB, comprèn els residus 82A-95A amb una hèlix de dues voltes més un llaç de 6 residus desestructurat. La numeració dels segments pro que s'utilitza en aquest treball correspon a la de la PCPB porcina (Coll *et al.*, 1991), la primera estructura tridimensional que contenia informació sobre el segment d'activació. En aquesta numeració els residus pertanyents a aquest segment són identificats per una "A" que segueix al número que identifica la posició.

Introducció

```

-110                                     -100                                     -94
ATG CCG GGG TTG CTG GTG TTG AGT GTC CTG TTG GGG GCT GTC TTT GGC AAG GAG GAC TTT
Met arg gly leu leu val leu ser val leu leu gly ala val phe gly lys glu asp phe
-90
GTG GGG CAT CAG GTG CTC CGA ATC TCT GTA GCC GAT GAG GCC CAG GTA CAG AAG GTG AAG
val gly his gln val leu arg ile ser val ala asp glu ala gln val gln lys val lys
-70
GAG CTG GAG GAC CTG GAG CAC CTG CAG CTG GAC TTC TGG CGG GGG CCT GCC CAC CCT GGC
glu leu glu asp leu glu his leu gln leu asp phe trp arg gly pro ala his pro gly
-50
TCC CCC ATC GAC GTC CGA GTG CCC TTC CCC AGC ATC CAG GCG GTC AAG ATC TTT CTG GAG
ser pro ile asp val arg val pro phe pro ser ile gln ala val lys ile phe leu glu
-30
TCC CAC GGC ATC AGC TAT GAG ACC ATG ATC GAG GAC GTG CAG TCG CTG CTG GAC GAG GAG
ser his gly ile ser tyr glu thr met ile glu asp val gln ser leu leu asp glu glu
-10
CAG GAG CAG ATG TTC GCC TTC CGG TCC CGG GCG CGC TCC ACC GAC ACT TTT AAC TAC GCC
gln glu gln met phe ala phe arg ser arg ala arg ser thr asp thr phe asn tyr ala
30
ACC TAC CAC ACC CTG GAG GAG ATC TAT GAC TTC CTG GAC CTG CTG GTG GCG GAG AAC CCG
thr tyr his thr leu glu glu ile tyr asp phe leu asp leu leu val ala glu asn pro
50
CAC CTT GTC AGC AAG ATC CAG ATT GGC AAC ACC TAT GAA GGG CGT CCC ATT TAT GTG CTG
his leu val ser lys ile gln ile gly asn thr tyr glu gly arg pro ile tyr val leu
70
AAG TTC AGC ACG GGG GGC AGT AAG CGT CCA GCC ATC TGG ATC GAC ACG GGC ATC CAT TCC
lys phe ser thr gly gly ser lys arg pro ala ile trp ile asp thr gly ile his ser
90
CGG GAG TGG GTC ACC CAG GCC AGT GGG GTC TGG TTT GCA AAG AAG ATC ACT CAA GAC TAT
arg glu trp val thr gln ala ser gly val trp phe ala lys lys ile thr gln asp tyr
110
GGG CAG GAT GCA GCT TTC ACC GCC ATT CTC GAC ACC TTG GAC ATC TTC CTG GAG ATC GTC
gly gln asp ala ala phe thr ala ile leu asp thr leu asp ile phe leu glu ile val
120
ACC AAC CCT GAT GGC TTT GCC TTC ACG CAC AGC ACG AAT CGC ATG TGG CGC AAG ACT CGG
thr asn pro asp gly phe ala phe thr his ser thr asn arg met trp arg lys thr arg
140
TCC CAC ACA GCA GGC TCC CTC TGT ATT GGC GTG GAC CCC AAC AGG AAC TGG GAC GCT GGC
ser his thr ala gly ser leu cys ile gly val asp pro asn arg asn trp asp ala gly
160
TTT GGG TTG TCC GGA GCC AGC AGT AAC CCC TGC TCG GAG ACT TAC CAC GGC AAG TTT GCC
phe gly leu ser gly ala ser ser asn pro cys ser glu thr tyr his gly lys phe ala
180
AAT TCC GAA GTG GAG GTC AAG TCC ATT GTA GAC TTT GTG AAG GAC CAT GGG AAC ATC AAG
asn ser glu val glu val lys ser ile val asp phe val lys asp his gly asn ile lys
200
GCC TTC ATC TCC ATC CAC AGC TAC TCC CAG CTC CTC ATG TAT CCC TAT GGC TAC AAA ACA
ala phe ile ser ile his ser tyr ser gln leu leu met tyr pro tyr gly tyr lys thr
220
GAA CCA GTC CCT GAC CAG GAT GAG CTG GAT CAG CTT TCC AAG GCT GCT GTG ACA GCC CTG
glu pro val pro asp gln asp glu leu asp gln leu ser lys ala ala val thr ala leu
240
GCC TCT CTC TAC GGG ACC AAG TTC AAC TAT GGC AGC ATC ATC AAG GCA ATT TAT CAA GCC
ala ser leu tyr gly thr lys phe asn tyr gly ser ile ile lys ala ile tyr gln ala
260
AGT GGA AGC ACT ATT GAC TGG ACC TAC AGC CAG GGC ATC AAG TAC TCC TTC ACC TTC GAG
ser gly ser thr ile asp trp thr tyr ser gln gly ile lys tyr ser phe thr phe glu
280
CTC CGG GAC ACT GGG CGC TAT GGC TTC CTG CTG CCA GCC TCC CAG ATC ATC CCC ACA GCC
leu arg asp thr gly arg tyr gly phe leu leu pro ala ser gln ile ile pro thr ala
300
AAG GAG ACG TGG CTG GCG CTT CTG ACC ATC ATG GAG CAC ACC CTG AAT CAC CCC TAC TGA
lys glu thr trp leu ala leu leu thr ile met glu his thr leu asn his pro tyr OPA
GCTGACCCCTTGACACCCCTCTTGTCTCTCTCTGCCCCCTCCAGGCAACCAATAAAGTTTGAGTGT...

```

Figura I.7.- Seqüència nucleotídica i aminoacídica de la PCPA1 humana. Els 16 primers aminoàcids corresponen al pèptid senyal (blau), els 94 següents corresponen al segment d'activació (verd) i els 309 restants a l'enzim actiu (vermell).

```

-110                                     -100                                     -95
ATG TTG GCA CTC TTG GTT CTG GTG ACT GTG GCC CTG GCA TCT GCT CAT CAT GGT GGT GAG
Met leu ala leu leu val leu val thr val ala leu ala ser ala his his gly gly glu
-90
CAC TTT GAA GGC GAG AAG GTG TTC CGT GTT AAC GTT GAA GAT GAA AAT CAC ATT AAC ATA
his phe glu gly glu lys val phe arg val asn val glu asp glu asn his ile asn ile
-70
ATC CGC GAG TTG GCC AGC ACG ACC CAG ATT GAC TTC TGG AAG CCA GAT TCT GTC ACA CAA

```

Introducció

```

ile arg glu leu ala ser thr thr gln ile asp phe trp lys pro asp ser val thr gln
-50
ATC AAA CCT CAC AGT ACA GTT GAC TTC CGT GTT AAA GCA GAA GAT ACT GTC ACT GTG GAG
ile lys pro his ser thr val asp phe arg val lys ala glu asp thr val thr val glu
-30
AAT GTT CTA AAG CAG AAT GAA CTA CAA TAC AAG GTA CTG ATA AGC AAC CTG AGA AAT GTG
asn val leu lys gln asn glu leu gln tyr lys val leu ile ser asn leu arg asn val
-10
GTG GAG GCT CAG TTT GAT AGC CGG GTT CGT GCA ACA GGA CAC AGT TAT GAG AAG TAC AAC
val glu ala gln phe asp ser arg val arg ala thr gly his ser tyr glu lys tyr asn
20
AAG TGG GAA ACG ATA GAG GCT TGG ACT CAA CAA GTC GCC ACT GAG AAT CCA GCC CTC ATC
lys trp glu thr ile glu ala trp thr gln gln val ala thr glu asn pro ala leu ile
40
TCT CGC AGT GTT ATC GGA ACC ACA TTT GAG GGA CGC GCT ATT TAC CTC CTG AAG GTT GGC
ser arg ser val ile gly thr thr phe glu gly arg ala ile tyr leu leu lys val gly
60
AAA GCT GGA CAA AAT AAG CCT GCC ATT TTC ATG GAC TGT GGT TTC CAT GCC AGA GAG TGG
lys ala gly gln asn lys pro ala ile phe met asp cys gly phe his ala arg glu trp
80
ATT TCT CCT GCA TTC TGC CAG TGG TTT GTA AGA GAG GCT GTT CGT ACC TAT GGA CGT GAG
ile ser pro ala phe cys gln trp phe val arg glu ala val arg thr tyr gly arg glu
100
ATC CAA GTG ACA GAG CTT CTC AAC AAG TTA GAC TTT TAT GTC CTG CCT GTG CTC AAT ATT
ile gln val thr glu leu leu asn lys leu asp phe tyr val leu pro val leu asn ile
120
GAT GGC TAC ATC TAC ACC TGG ACC AAG AGC CGA TTT TGG AGA AAG ACT CGC TCC ACC CAT
asp gly tyr ile tyr thr trp thr lys ser arg phe trp arg lys thr arg ser thr his
140
ACT GGA TCT AGC TGC ATT GGC ACA GAC CCC AAC AGA AAT TTT GAT GCT GGT TGG TGT GAA
thr gly ser ser cys ile gly thr asp pro asn arg asn phe asp ala gly trp cys glu
160
ATT GGA GCC TCT CGA AAC CCC TGT GAT GAA ACT TAC TGT GGA CCT GCC GCA GAG TCT GAA
ile gly ala ser arg asn pro cys asp glu thr tyr cys gly pro ala ala glu ser glu
180
AAG GAG ACC AAG GCC CTG GCT GAT TTC ATC CGC AAC AAA CTC TCT TCC ATC AAG GCA TAT
lys glu thr lys ala leu ala asp phe ile arg asn lys leu ser ser ile lys ala tyr
200
CTG ACA ATC CAC TCG TAC TCC CAA ATG ATG ATC TAC CCT TAC TCA TAT GCT TAC AAA CTC
leu thr ile his ser tyr ser gln met met ile tyr pro tyr ser tyr ala tyr lys leu
220
GGT GAG AAC AAT GCT GAG TTG AAT GCC CTG GCT AAA GCT ACT GTG AAA GAA CTT GCC TCA
gly glu asn asn ala glu leu asn ala leu ala lys ala thr val lys glu leu ala ser
240
CTG CAC GGC ACC AAG TAC ACA TAT GGC CCG GGA GCT ACA ACA ATC TAT CCT GCT GCT GGG
leu his gly thr lys tyr thr tyr gly pro gly ala thr thr ile tyr pro ala ala gly
260
GGC TCT GAC GAC TGG GCT TAT GAC CAA GGA ATC AGA TAT TCC TTC ACC TTT GAA CTT CGA
gly ser asp asp trp ala tyr asp gln gly ile arg tyr ser phe thr phe glu leu arg
280
GAT ACA GGC AGA TAT GGC TTT CTC CTT CCA GAA TCC CAG ATC CGG GCT ACC TGC GAG GAG
asp thr gly arg tyr gly phe leu leu pro glu ser gln ile arg ala thr cys glu glu
300
ACC TTC CTG GCA ATC AAG TAT GTT GCC AGC TAC GTC CTG GAA CAC CTG TAC TAG TTG AGA
thr phe leu ala ile lys tyr val ala ser tyr val leu glu his leu tyr AMB
307
AAG CTG ATG GCC TTG TTT CAA AAT TCT CAT TTT TCA TTT CTT TTC TTT

```

Figura I.8.- Seqüència nucleotídica i aminoacídica de la PCPB1 humana. Els 15 primers residus corresponen al pèptid senyal (blau), els 95 següents al segment d'activació (vermell) i els 307 últims residus a l'enzim actiu (verd).

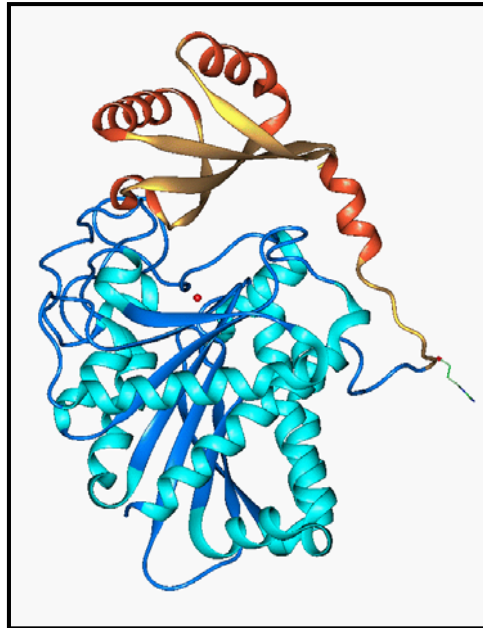


Figura I.9.- Diagrama de cintes de l'estructura de la PCPB porcina. En groc i taronja es representa el segment d'activació i en blau el domini enzimàtic. El punt vermell simbolitza l'ió de zinc. Es mostra la cadena lateral de l'arginina 95A, on té lloc el tall activant.

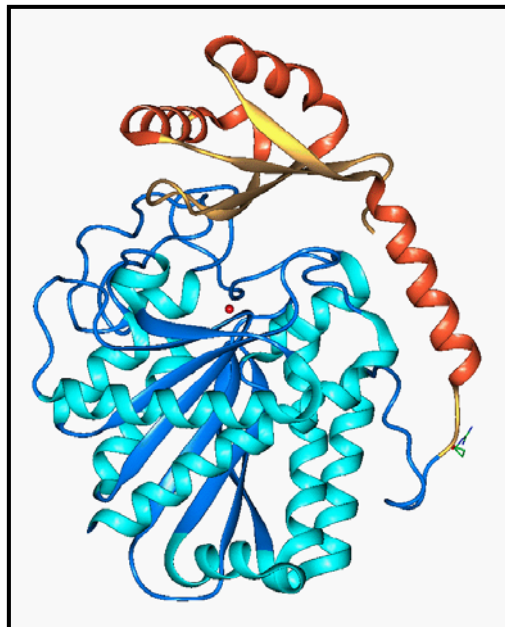


Figura I.10.- Diagrama de cintes de l'estructura de la PCPA1 porcina. El punt vermell simbolitza l'ió de zinc. S'ha representat la cadena lateral de l'arginina 99A, on té lloc l'activació del proenzim.

La regió globular del segment d'activació de la PCPB interacciona amb el subseti S1' (responsable de l'ancoratge del substrat) a través d'un pont salí entre l'Arg145 de l'enzim i l'Asp41A del segment. Aquesta interacció contribueix a l'absència d'activitat intrínseca de la PCPB. En el cas de les formes A no existeix un residu equivalent a Asp41A en el segment *pro* i la conformació d'aquesta regió és diferent degut a que no presenta l'hèlix 3₁₀, factors que contribueixen a què aquestes formes presentin activitat intrínseca. El segment d'activació de la PCPB inhibeix completament la CPB per a la majoria de substrats peptídics. La PCPA1 presenta un 10% de la V_{max} de l'enzim per a substrats peptídics curts i la seva activitat disminueix amb la grandària del substrat fins un 1,5% per a hexapèptids (Serra, 1992). En el cas de la PCPA2, el canal situat entre l'enzim i el segment d'activació per on difonen els substrats cap al centre actiu és més estret que en la PCPA1, de manera que presenta un 10% d'activitat per al substrat Cbz-Gly-Tyr, però aquest valor baixa al 3% si es canvia la tirosina en posició C-terminal per un residu més voluminós com el triptòfan (Reverter *et al.*, 1997). En la figura I.12 es presenta un alineament de les seqüències dels segments d'activació dels tres proenzims.

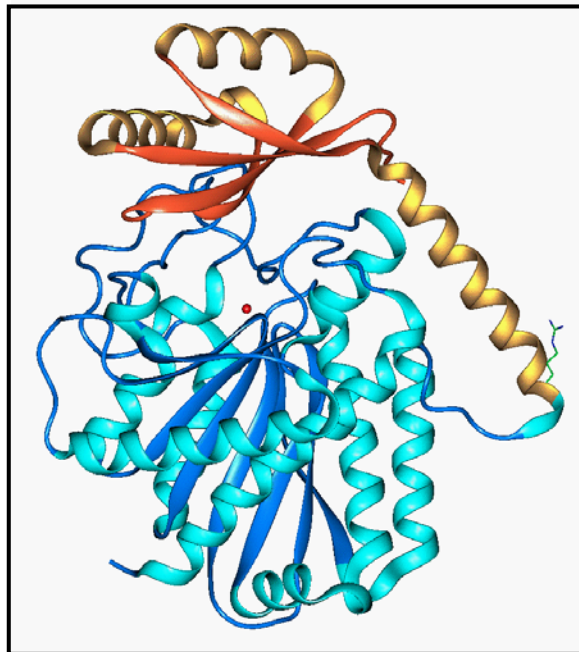


Figura I.11.- Diagrama de cintes de l'estructura de la PCPA2 humana. El punt vermell representa l'ió de zinc. Es mostra la cadena lateral de l'arginina 101A, on té lloc el tall activant.

Existeixen una sèrie de contactes específics entre segment d'activació i enzim que bloquegen els subetis implicats en la unió de substrats peptídics. Els responsables d'aquests contactes són, en la cadena $\beta 2$, Asp36A, Phe37A i Trp38A; en $\beta 3$, Asp53A en la PCPA1 i Arg55A en la PCPB i, per últim, en el llaç que connecta ambdues cadenes, Arg47A en la PCPA1, Asp41A, Gln45A i Lys47A en la PCPB. De tots aquests residus, la tríada Asp36A-Phe37A-Trp38A està molt conservada i es creu que és essencial per a la interacció. L'Arg127 de l'enzim s'uneix a Asp41A del segment de la PCPB per un pont salí mediat per aigua. L'estructura de les molècules d'aigua al voltant del Glu270 i del zinc difereixen de les trobades en la CPA1 bovina. Respecte la Tyr248, aquesta canvia de la posició *up* a la *down* quan un anàleg de substrat s'uneix al centre actiu. Encara que aquest residu no està implicat en la catàlisi, intervé en la fixació del carbonil P1' i del carbonil P2 i podria influir en l'accessibilitat de les molècules d'aigua al centre actiu, així com formar una xarxa aromàtica amb altres residus de l'enzim i del substrat (Christianson i Lipscomb, 1989; Philips *et al.*, 1990).

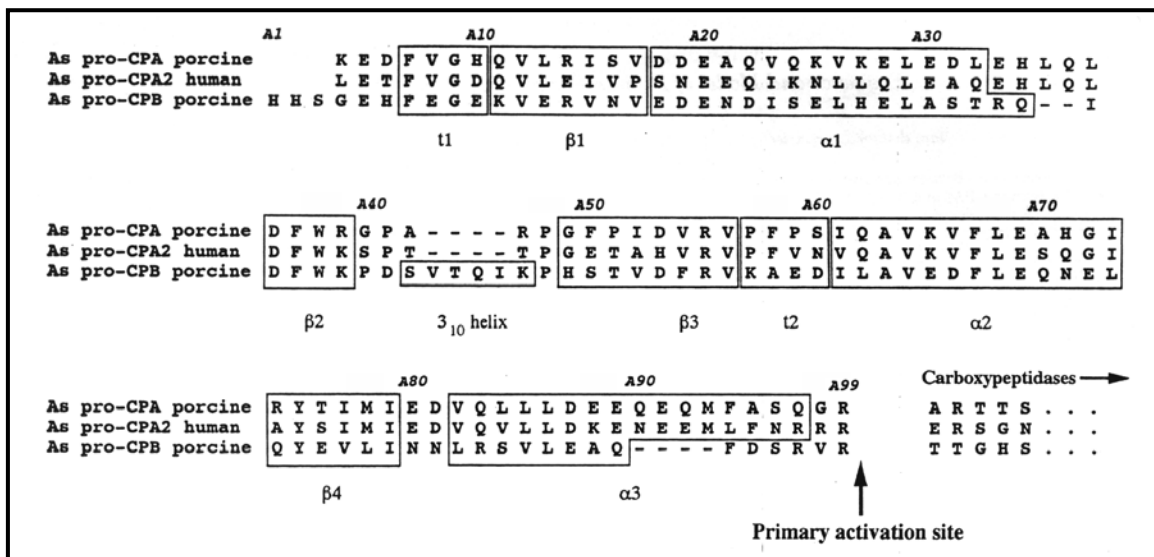


Figura I.12.- Comparació de l'estructura primària i secundària dels segments d'activació de les PCPs pancreàtiques A1 i B porcines i A2 humana. Els quadres indiquen l'estructura secundària, que s'especifica sota en cada cas. Observi's que en l'alineament s'ha tingut en compte, a més de l'homologia seqüencial, consideracions de caire estructural. Modificat de Catasús *et al.* (1995).

I.3.2.- ACTIVACIÓ DE LES PROCARBOXIPEPTIDASES PANCREÀTIQUES

Les procarboxipeptidases pancreàtiques s'activen per acció de la tripsina, que allibera el segment *pro* o d'activació. Fins ara s'han estudiat detalladament els processos d'activació de la PCPA1 porcina (Vendrell *et al.*, 1990c), de la PCPB porcina (Burgos *et al.*, 1991; Villegas *et al.*, 1995a; Ventura *et al.*, 1999; Companys, 2002), del sistema de PCPs de rata (Opezco *et al.*, 1994) i de la PCPA2 humana (Reverter *et al.*, 1998). D'altra banda, pel sistema de PCPs humanes s'ha realitzat un estudi preliminar dels seus processos d'activació (Pascual *et al.*, 1989). Aquest conjunt d'estudis mostren que les diferents PCPs segueixen processos d'activació característics. Així, l'activació progressa més ràpidament en les formes B i A2 que en la PCPA1, degut a l'efecte inhibitor que té el segment d'activació sobre la CPA1 un cop s'ha produït el tall tríptic. També s'ha observat que el segment d'activació de la PCPA1 es degrada més lentament que el de la PCPB, efecte més marcat quan la primera es troba formant estructures quaternàries (Vendrell *et al.*, 1982; Ventura *et al.*, 1997). El segment d'activació de la PCPA2 és alliberat ràpidament, però l'elevada especificitat de l'enzim generat fa que aquest no intervingui en el seu procés de degradació (Opezco *et al.*, 1994; Reverter *et al.*, 1998).

I.3.2.1.- Procés d'activació de la PCPA1 porcina

La generació d'activitat carboxipeptidasa durant el procés d'activació de la PCPA1 segueix una corba bifàsica que no es correlaciona amb la formació de l'enzim actiu (Vendrell *et al.*, 1990c), de manera que si bé electroforèticament s'observa que al cap d'un minut tot el proenzim s'ha transformat en enzim actiu, en aquest moment només es detecta un 30% de l'activitat potencial total.

El seguiment per HPLC de la generació dels diferents fragments del segment d'activació permet relacionar-los amb l'aparició d'activitat. En la figura I.13 es presenta un esquema general de l'activació de la PCPA1 porcina. En primer lloc, la tripsina fa un tall en l'enllaç Arg99A-Ala1. El segment generat, SaA1 (de 94 residus), no s'allibera immediatament, sinó que roman interaccionant amb l'enzim inhibint-lo (K_i de 2 nM) (San Segundo *et al.*, 1982; Vilanova *et al.*, 1985). A continuació, la CPA1 escindeix l'arginina C-terminal d'SaA1 i es genera el segment SaA2, que mostra una capacitat inhibidora sensiblement disminuïda, encara que fins que no desapareixen ambdós segments no s'observa la inflexió en la corba d'activitat. Simultàniament a partir dels dos segments, per acció tríptica sobre els enllaços Gln92A-Met93A, Arg74A-Tyr75A i Ala70A-His71A, es generen els fragments 3, 4 i 5 respectivament, els quals han perdut pràcticament la capacitat inhibidora i l'aparició del fragment 4 es pot associar al

100% de l'activitat. SaA4 i SaA2 són els fragments més importants quantitativament. Durant el procés d'activació de la PCPA1 porcina s'observa que es dona un tall tríptic sobre la CPA1 generada en l'enllaç Arg2-Thr3, de manera que s'obtenen dues formes de l'enzim actiu, anomenades CPA1 α i β .

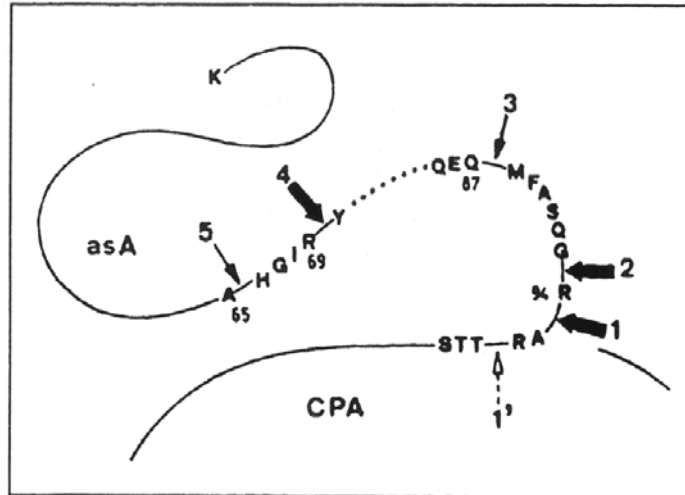


Figura I.13.- Esquema dels principals punts de tall durant l'activació tríptica de la PCPA1 porcina. Les fletxes en negreta indiquen aquells que donen lloc als fragments detectats majoritàriament. 1' indica l'enllaç peptídic hidrolitzat en la conversió de carboxipeptidasa α a β . Noti's que en l'esquema s'ha emprat una numeració correlativa dels residus del segment d'activació, de manera que l'arginina en posició 94 correspon al residu Arg99A segons la notació convencional del segment d'activació. Reproduït de Vendrell *et al.*, (1990c).

I.3.2.2.- Procés d'activació de la PCPB porcina

El procés d'activació de la PCPB porcina s'ha descrit detalladament (Burgos *et al.*, 1991; Villegas *et al.*, 1995a; Ventura *et al.*, 1999; Companys, 2002). El seu procés d'activació segueix una corba monofàsica i és molt més ràpid que el de la PCPA1 porcina, pel que s'han d'utilitzar condicions suaus d'activació: relació PCPB/tripsina de 400/1 (p/p) i 0 °C en comptes d'una relació 40/1 i els 25 o 37 °C necessaris per a la PCPA1 porcina. L'enzim actiu que es genera durant el procés d'activació està implicat en el processament proteolític de la PCPB, cosa que s'ha demostrat realitzant l'activació en presència d'inhibidors de carboxipeptidases, de manera que s'han pogut detectar els fragments generats que contenen una arginina C-terminal, residu que altrament és escindit per la CPB generada després de l'acció tríptica (Villegas *et al.*, 1995a).

L'activació tríptica de la PCPB porcina té lloc mitjançant proteòlisi de l'enllaç Arg95A-Thr1, com a conseqüència del qual s'allibera la CPB (307 residus) i el fragment F1 α o SaB1 α (1-95 residus). Aquest fragment no és detectable durant el procés d'activació, ja que l'arginina C-terminal que conté és ràpidament escindida per la CPB per donar lloc a F1 β (1-94). A continuació la tripsina talla F1 β en l'enllaç Arg83A-Ser84A i es genera F2 α (1-83) i F4 β (84-94). F2 α és indetectable perquè la seva arginina C-terminal és proteolitzada ràpidament per la CPB, però si es detecta F2 β (1-82), ja que conté una leucina C-terminal que la CPB degrada molt lentament per originar F3 (1-81). F3 és un domini globular estable a la proteòlisi anomenat domini d'activació. En la figura I.14 es mostra un esquema general de l'activació de la PCPB porcina.

Quan es realitza l'activació en presència d'inhibidors, la tripsina pot tallar l'enllaç Arg93A-Val94A amb baixa eficiència, de manera que es generen dues formes de carboxipeptidasa activa (Villegas *et al.*, 1995a). Sembla ser que l'inhibidor és capaç d'induir un canvi en la posició del segment d'activació en el proenzim i exposar una nova diana tríptica, a més, s'observa una acceleració del procés d'activació. En el cas de la PCPB porcina, a diferència del que s'observa en la PCPA1 porcina, cap dels fragments generats durant l'activació sembla ser capaç de mantenir-se unit a l'enzim (Ventura *et al.*, 1999) i, per tant, no poden inhibir-la després del primer tall tríptic.

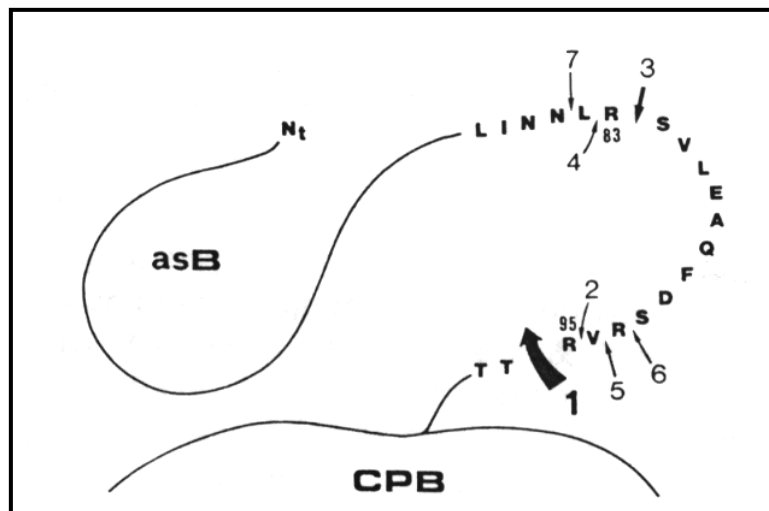


Figura I.14.- Esquema dels principals punts de tall que es donen durant l'activació tríptica de la PCPB porcina. 1 i 3 representen els talls tríptics que donen lloc als fragments F1 i F2. Extret de Burgos *et al.*, (1991).

I.3.2.3.- Procés d'activació de la PCPA2 humana

L'activació de la PCPA2 humana és un procés ràpid que segueix una corba monofàsica (Reverter *et al.*, 1998). Com en el cas de la PCPB porcina, cap dels fragments generats després del primer tall tríptic pot inhibir l'enzim.

Durant el procés d'activació de la PCPA2 humana, el primer tall tríptic té lloc en l'enllaç Arg101A-Ser1, com a conseqüència del qual s'allibera un fragment de 96 residus anomenat fragment α . Després del tall que allibera el segment d'activació de la PCPA2 la tripsina actua sobre l'Arg99A, alliberant-se el dipèptid Glu-Arg en posició C-terminal del segment primari d'activació i generant-se el fragment β . Per últim, el tall tríptic de l'enllaç Arg97A-Arg98A dona lloc al fragment γ , que és resistent a la proteòlisi. L'alta especificitat de la CPA2 no li permet escindir les arginines C-terminal. La figura I.15 mostra un esquema d'aquest procés d'activació.

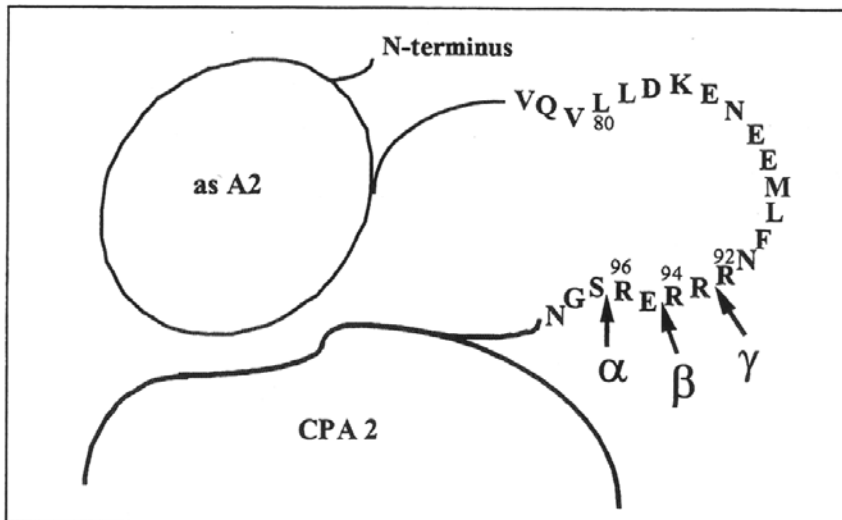


Figura I.15.- Esquema de l'activació tríptica de la PCPA2 humana. Es mostra la seqüència de la regió de connexió entre el domini d'activació i l'enzim actiu. Les fletxes indiquen els punts d'acció tríptica. Noti's que en l'esquema s'ha emprat una numeració correlativa dels residus del segment d'activació, de manera que l'arginina en posició 96 correspon al residu Arg101A segons la notació convencional del segment d'activació. Extret de Reverter *et al.*, (1998).

I.3.2.4.- Estudi preliminar del procés d'activació de les PCPs humanes

S'ha realitzat un estudi preliminar de l'activació tríptica del sistema de les procarboxipeptidases humanes (Pascual *et al.*, 1989). Degut a què es va treballar amb

proenzims purificats de pàncrees humà i es disposava de petites quantitats de proteïna, aquests estudis no es van realitzar amb la concentració de proenzim que s'usa habitualment, que és d'1mg/mL, sinó que es va emprar una concentració de 45 µg/mL. Les activacions es van fer a 25 °C i per a les formes A es va usar una relació proenzim/tripsina de 40/1 (p/p) i per a les formes B una relació 4/1. El fet que s'utilitzés una quantitat relativa de tripsina tan elevada per a les formes B és degut a què amb relacions més baixes els proenzims no poden arribar a expressar el seu màxim d'activitat. En les condicions esmentades, es va observar que tant les formes B, com l'A2 segueixen una activació monotònica i ràpida, assolint-se en tots els casos el 50% de l'activitat màxima al cap d'uns 10 minuts. La PCPA1 i la PCPA3 presenten una corba d'activació lleugerament sigmoïdal i el 50% d'activitat màxima no s'assoleix fins després d'uns 60 minuts. La figura I.16 mostra una representació de les corbes d'activació comentades.

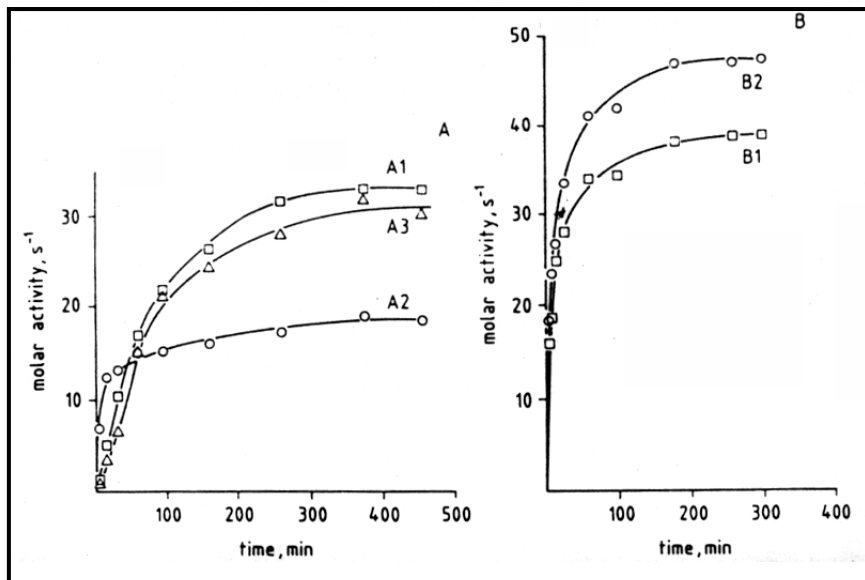


Figura I.16.- Corbes d'activació de les PCPs humanes amb tripsina. A- Activació de PCPAs amb tripsina (relació 40/1). B- Activació PCPBs amb tripsina (relació 4/1). Reproduït de Pascual *et al.*, (1989).

I.4.- TAFI

TAFI (**T**hrombin-**A**ctivable **F**ibrinolysis **I**nhibitor) és una procarboxipeptidasa sintetitzada en el fetge que circula per la sang unida a plasminogen i que, en activar-se, genera un enzim inestable amb activitat front a residus bàsics. Ha estat descrita per diferents grups de recerca, d'aquí que sigui coneguda per diversos noms: CPR (Campbell *et al.*, 1989), CPU (Hendriks *et*

al., 1990), CPB de plasma (Eaton *et al.*, 1991) i TAFI (Bajzar *et al.*, 1995). La TAFI presenta més d'un 40% d'identitat seqüencial amb les procarboxipeptidases pancreàtiques, i les grandàries del seu segment *pro* (92 residus) i domini enzimàtic (309 residus) són semblants a les de les formes pancreàtiques, però, a diferència d'aquestes, la TAFI presenta glicosilacions en el seu segment *pro*. S'ha descrit la presència de TAFI en diversos mamífers i en pollastre.

Tot i que la TAFI és activable per plasmina (Wang *et al.*, 1994) i per trombina, s'ha observat que la velocitat d'activació s'incrementa 1250 vegades quan la trombina es troba formant un complex amb la trombomodulina (Bajzar *et al.*, 1996), cosa que indica que l'activador fisiològic de la TAFI deu ser el complex trombina/trombomodulina. Aquest procés d'activació ve modulats pel domini *EGF-like* 3 de la trombomodulina i és inhibit competitiuament per la proteïna C (Kokame *et al.*, 1998). La TAFI constitueix una connexió molecular entre la cascada de coagulació i la fibrinolítica. La trombina produïda per la cascada de coagulació activa la TAFI, donant lloc a un enzim actiu (TAFIa) que inhibeix la fibrinòlisi. Aquesta funció la realitza mitjançant l'escissió dels residus de lisina i arginina carboxi-terminals exposats en la fibrina parcialment degradada, cosa que fa que la capacitat d'unió entre la fibrina i el plasminogen disminueixi, de manera que no pot donar-se la degradació de fibrina mediada per l'activador tissular de plasminogen (t-PA) (Wang *et al.*, 1998). En la figura I.17 es representa un esquema simplificat de les cascades de coagulació i fibrinòlisi.

La TAFIa és intrínsecament inestable a 37°C, amb una vida mitja de 8-9 minuts (Boffa *et al.*, 1998). Aquesta inestabilitat és fortament depenent de la temperatura, de manera que a temperatures més baixes augmenta l'estabilitat de l'enzim. La inactivació de TAFIa és deguda a una desestabilització conformacional, com a conseqüència de la qual es dona una degradació proteolítica de l'enzim (Boffa *et al.*, 2000, Marx *et al.*, 2000). S'han identificat diferents isoformes de l'enzim degudes a variacions en la posició 147 (Ala 147 o Thr 147) i en la posició 325 (Ile 325 o Thr 325). La variació A147T no comporta canvis en l'activitat antifibrinolítica de l'enzim (Zhao *et al.*, 1998), però s'ha observat que la variació T325I dona lloc a enzims que difereixen en quant a estabilitat i activitat antifibrinolítica (Schneider *et al.*, 2002).

Donada la importància de la TAFI en l'hemostàsia, la concentració de TAFI en plasma s'ha convertit en un paràmetre de gran interès mèdic i s'han posat a punt diversos mètodes per tal de poder-la determinar (Schatteman *et al.*, 1999, Schatteman *et al.*, 2001). El nivell de TAFI en plasma presenta una gran variació interindividual que no és explicable en base a l'estil de vida (Tregouet *et al.*, 2001). S'han identificat una sèrie de polimorfismes en el promotor i en el 3'UTR del gen de la TAFI que semblen lligats a la determinació de la concentració de TAFI en plasma (Franco *et al.*, 2001, Henry *et al.*, 2001). Degut a l'acció antifibrinolítica de la TAFI, s'ha associat un nivell alt de TAFI en plasma amb risc de trombosi venosa (Franco *et al.*, 2001, van Tilburg *et al.*, 2000) i de malaltia arterial coronària (Silveira *et al.*, 2000). Per altra banda,

s'ha observat una baixa concentració de TAFI en malalts amb insuficiència hepàtica crònica (van Thiel *et al.*, 2001) o amb cirrosi (Lisman *et al.*, 2001) i també s'ha relacionat la TAFI amb l'hemofília (Mosnier *et al.*, 2001) i amb la pressió sanguínia (Koschinsky *et al.*, 2001).

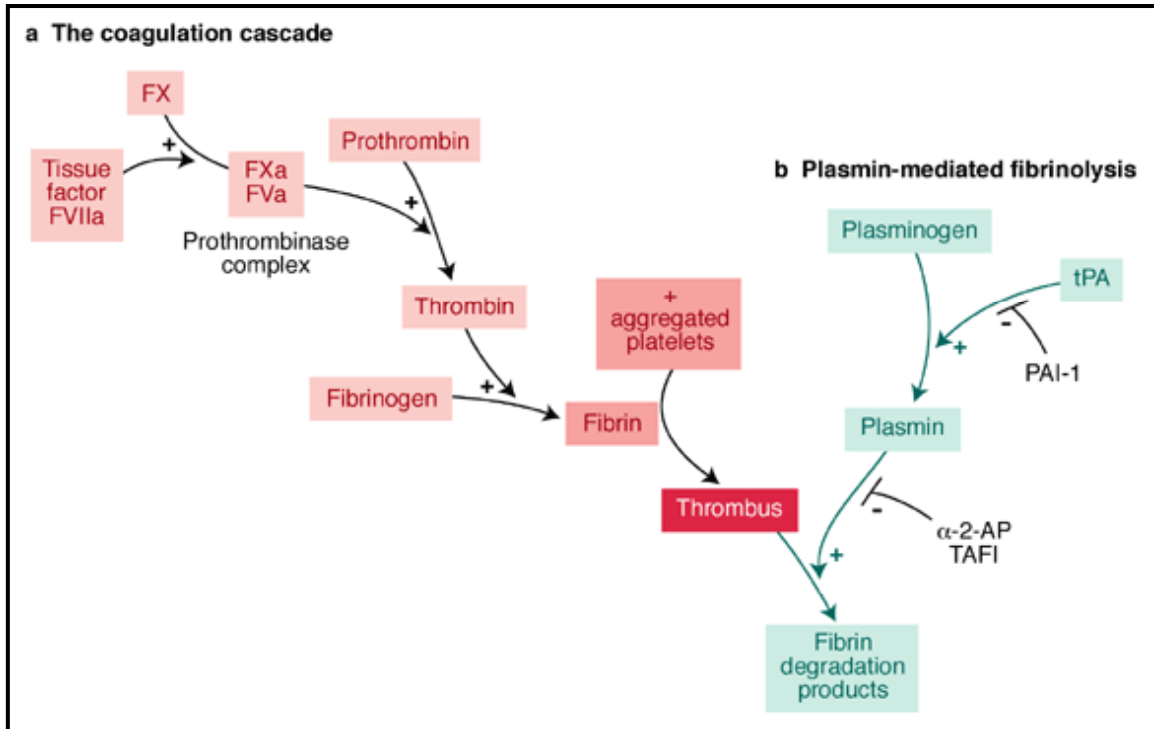


Figura I.17.- Esquema simplificat de les cascades de coagulació i fibrinòlisi. (a) La cascada de coagulació s'inicia *in vivo* pel factor tissular i el factor VIIa (FVIIa) i dona lloc a la conversió de protrombina a trombina a través del complex protrombinasa (FXa i FVa). A continuació la trombina transforma el fibrinogen a fibrina que, juntament amb l'agregació de plaquetes, forma el coàgul. (b) La fibrinòlisi actua desfent el coàgul mitjançant l'acció de la plasmina. El responsable de la conversió del plasminogen a plasmina és l'activador tissular de plasminogen (tPA). L'inhibidor 1 de l'activador del plasminogen (PAI-1) inhibeix el tPA ràpidament. L' α -2-antiplasmina (α -2-AP) inactiva la plasmina formant amb ella un complex inhibidor. La TAFI talla els residus de lisina C-terminals de la fibrina, evitant la co-activació del plasminogen per part de la fibrina. Extret de Bodary *et al.*, 2002.

A part de la seva funció en la regulació de la coagulació sanguínia, la TAFI també està involucrada en el processament de pèptids bioactius generats durant la inflamació (Campbell *et al.*, 1989, Campbell *et al.*, 2001, Campbell *et al.*, 2002). Entre els seus substrats s'han descrit quinines i anafilotoxines, algunes de les quals són processades amb més eficiència per la TAFI

que per la CPN, el clàssic inactivador d'anafilotoxina. En relació amb aquesta funció de la TAFI, s'ha descrit que en pacients amb artritis reumatoide la concentració de TAFI en sèrum està reduïda o absent, cosa que es podria relacionar amb la inflamació local característica d'aquesta malaltia (Campbell *et al.*, 1990).

La determinació de l'estructura tridimensional de la TAFI ens permetria entendre el seu comportament i, donada la seva importància biomèdica, constituïria la base per al disseny racional de fàrmacs. De moment, no ha estat possible cristal·litzar la TAFI, probablement degut a la presència de glicosilacions en el segment *pro* o a la inestabilitat intrínseca de l'enzim.

OBJECTIUS

Els objectius i el treball d'aquesta tesi s'inscriuen dins l'àmbit descrit a la Introducció i que constitueix una de les línies de recerca principals del nostre grup: la relació estructura/funció en metal·locarboxipeptidases. En la recerca dins d'aquesta línia les tècniques experimentals de biologia molecular i química de proteïnes es complementen amb estudis teòrics per computador.

El treball realitzat s'ha dividit en una primera i una segona part, centrades fonamentalment en estudis funcionals i estructurals, respectivament, però en ambdues parts s'ha seguit l'objectiu d'aprofundir en el coneixement de la relació estructura/funció de les metal·locarboxipeptidases humanes i els seus zimògens.

En el primer capítol s'ha treballat amb aspectes funcionals de les procarboxipeptidases pancreàtiques humanes PCPA1 i PCPB; concretament, s'ha volgut aprofundir en l'estudi dels seus processos d'activació, ja que fins ara només s'havia realitzat un estudi detallat per a la PCPA2, mentre que per a la PCPA1 i la PCPB tan sols es disposava de dades d'un estudi preliminar realitzat amb petites quantitats de proteïna purificada de pàncrees humà. La impossibilitat d'obtenir elevades quantitats d'aquests zimògens a partir de pàncrees humà, degut tant a la labilitat d'aquestes proteïnes com a la dificultat d'aconseguir aquest tipus de material, ens va fer plantejar la necessitat de posar a punt sistemes d'expressió heteròloga per a la PCPA1 i la PCPB per tal d'emprendre estudis funcionals detallats i, si el rendiment de l'expressió ho permetia, estudis estructurals.

Cal assenyalar que les formes actives d'aquests zimògens tenen una aplicació biomèdica i biotecnològica derivada del seu ús en la teràpia contra el càncer anomenada *Antibody-directed enzyme prodrug therapy* o ADEPT, en què la carboxipeptidasa és dirigida cap al tumor mitjançant la unió a un anticòs específic contra cèl·lules tumorals, de manera que es produeix una activació de la prodroga antitumoral de forma localitzada en el lloc on es desitja la seva actuació. L'interès en l'ús d'aquests enzims és degut al fet que gràcies al seu origen humà s'espera que presentin una baixa immunogenicitat en l'ús clínic. Així, és de gran interès desenvolupar sistemes d'expressió recombinant eficients per a aquests enzims amb vistes al seu interès biotecnològic, així com per realitzar estudis de relació estructura/funció que ens permetin aprofundir en el seu coneixement i poder millorar d'aquesta manera la seva aplicabilitat biotecnològica.

Els objectius concrets d'aquesta primera part de la tesi han estat:

- Desenvolupament d'un sistema d'expressió recombinat eficient en *P. pastoris* per aconseguir elevades quantitats de PCPA1 i PCPB humanes en forma soluble.

- Caracterització detallada dels processos d'activació trípica de la PCPA1 i PCPB humanes i comparació d'aquests processos amb d'altres ja descrits amb la finalitat de treure conclusions generals sobre l'activació de les procarboxipeptidases.
- Realització de mutagènesi dirigida sobre la PCPB humana per tal d'obtenir un mutant amb especificitat reversa, expressió del proenzim mutant i estudi dels efectes que té aquesta mutació sobre el procés d'activació de la PCPB humana.

La segona part d'aquesta tesi s'ha centrat en l'estudi d'aspectes estructurals de procarboxipeptidases humanes pancreàtiques i reguladores pertanyents a la subfamília A/B de les metal·locarboxipeptidases de la família M14. Els estudis realitzats s'han dut a terme fent ús de tècniques experimentals, com és la determinació d'estructures cristal·logràfiques mitjançant difracció de raigs X, i teòriques, com és el modelat de proteïnes per homologia. Els objectius concrets d'aquesta part del treball han estat:

- Cristal·lització i resolució per difracció de raigs X de les estructures tridimensionals de la PCPA1 i PCPB humanes, per tal que el coneixement detallat de les seves estructures permeti procedir al seu redisseny amb vistes a la seva aplicabilitat en biomedicina i biotecnologia, així com per tal d'esbrinar quins són els trets estructurals determinants dels seus processos d'activació característics.
- Obtenció d'un model de la TAFI humana, un proenzim d'elevada rellevància biomèdica, mitjançant la tècnica de modelat per homologia, gràcies a la disponibilitat de l'estructura tridimensional de la PCPB humana, amb qui la TAFI comparteix el percentatge d'identitat més elevat. Aquest model, en absència d'una estructura tridimensional obtinguda experimentalment, hauria de permetre el redisseny de la TAFI per enginyeria de proteïnes, així com l'establiment d'hipòtesis que expliquin les seves característiques funcionals, tals com la seva inestabilitat conformacional intrínseca.
- Cerca en la base de dades del genoma humà de seqüències nucleotídiques de nous membres de la família de les metal·locarboxipeptidases humanes i obtenció, mitjançant modelat per homologia, de models dels nous enzims trobats que ens permetin predir si es tracten de carboxipeptidases actives i quina podria ser llur especificitat.

En referència a aquest capítol, cal dir que gràcies a l'èxit aconseguit en l'expressió heteròloga eficient de la PCPA1 i PCPB, es van emprendre estudis de cristal·lització amb

ambdós zimògens humans, però tan sols el darrer proenzim va produir cristalls que difractessin adequadament. Així, tan sols s'ha pogut resoldre l'estructura tridimensional de la PCPB humana. Aquesta part del treball es va realitzar en col·laboració amb el grup del Dr. Miquel Coll de la unitat de cristal·lografia del Centre d'Investigació i Desenvolupament del CSIC a Barcelona i l'estructura va ser resolta pel Dr. Pedro José Barbosa Pereira. El coneixement de l'estructura tridimensional d'aquests enzims humans és imprescindible per tal de procedir al seu redisseny amb l'objectiu de millorar la seva aplicabilitat biomèdica i biotecnològica dins la teràpia de l'ADEPT. D'altra banda, a arrel de la disponibilitat de la seqüència quasi completa del genoma humà, es va decidir cercar gens de nous membres de la família de les metal·locarboxipeptidases humanes i amb les noves seqüències trobades es va procedir a l'obtenció de models estructurals per a les proteïnes corresponents. Aquesta part del treball es va realitzar en col·laboració amb el Dr. Lloyd D. Fricker de l'Albert Einstein College of Medicine (Bronx, NY), amb l'objectiu d'assignar estructura a totes les seqüències de metal·locarboxipeptidases humanes existents, en el que seria un treball de genòmica estructural aplicat al cas concret d'una família enzimàtica.

Les proteïnes que s'han modelat en aquest treball pertanyen a la subfamília A/B de la família M14 de metal·locarboxipeptidases i es tracten d'enzims anomenats *pancreatic-like*, ja que comparteixen un percentatge d'identitat superior amb les carboxipeptidases pancreàtiques. Els estudis realitzats han volgut aprofundir en el coneixement estructural d'aquestes proteïnes, per alguna de les quals, malgrat la seva importància fisiològica derivada de la seva involucració en processos reguladors, es disposa de poques o cap dada estructural. Així, a partir dels models obtinguts, s'espera poder extreure hipòtesis que ens permetin explicar determinades observacions funcionals, com succeeix en el cas de la TAFI, o, en absència de dades funcionals, hipòtesis que ens portin cap a l'elucidació de les seves possibles funcions.