Characterization of two metallocarboxypeptidases of M14D subfamily from bacterial and human origin

Anabel Otero Bilbao

Characterization of two metallocarboxypeptidases of M14D subfamily from bacterial and human origin

Doctoral thesis presented by Anabel Otero Bilbao
to obtain the degree of PhD in Biochemistry, Molecular Biology and Biomedicine
from the Universitat Autònoma de Barcelona

Institut de Biotecnologia i Biomedicina

Thesis supervised by
Prof. Francesc Xavier Avilés Puigvert and Dra. Julia Lorenzo Rivera

Anabel Otero Bilbao

Prof. Francesc Xavier Avilés Puigvert   Dra. Julia Lorenzo Rivera

Barcelona, June 2014

Abstract

Metallocarboxypeptidases (MCP) of M14 family catalyze the hydrolysis of C-terminal amino acid residues in protein and peptide substrates. Recently, our group described the M14D subfamily or cytosolic carboxypeptidases (CCPs) after genomic searching of the aminoacidic sequences similar to CCP1 or Nna1 (nervous system nuclear protein induced by axotomy). CCPs were identified in around 100 Gram-negative bacteria that usually contain one gene copy. The domains architecture of bacterial CCPs is simpler, consisting in the amino-terminal (N-domain) and the carboxypeptidase domain (CP domain) both conserved in the subfamily. In humans there are six paralogous genes (CCP1-CCP6) and, besides the N- and CP-domains, they contain additional extensions at the N- and/or the C-terminus. At the protein level, they are quite complex, labile and difficult to express and handle experimentally. It has been proposed a primary function for eukaryotic CCPs in processing the C-terminal posttranslational modifications of tubulins, including the glutamates side chains and the encoded glutamate. Additionally, human CCP1 shortens acidic tails, implicated in molecular interactions, from other proteins therefore it has been proposed a general function in the regulation of protein-protein and protein-DNA interactions. Recently, evidence from phylogenetic analysis and cellular localization suggests that eukaryotic CCPs functions could be primarily related to cilia and basal bodies. Nevertheless, many of the molecular mechanisms underlying the proposed functions remain unclear. On the other hand, prokaryotic CCPs are poorly characterized from the biochemical and functional point of view.

The present work aimed to provide more information on key structural and functional features of M14D subfamily using the CCP from *Pseudomonas aeruginosa* (PaCCP) and the human CCP6 (hCCP6) as models. The crystal 3D structure of recombinant PaCCP was solved showing a novel β-sandwich N-terminal domain followed by a well defined α/β-hydrolase CP domain. Although most residues seem to be well positioned in the active site, no activity against a wide array of substrates was detected. This suggests that PaCCP is in an inactive form or requires a very specific substrate(s) and/or an alosteric molecule to display the CP activity. Co-crystallization experiments with inhibitors typical of A and B specificities, traditionally used for M14A/B MCPs, indicate that the enzyme might require structural

rearrangements to be active. The structure of PaCCP provides a first detailed structural insight into mammalian CCPs evidencing a conserved global fold of the N- and CP-domains and the position of key residues. The functional processes involving PaCCP in *P. aeruginosa* were explored by characterizing a mutant strain deficient in the protein (Δ*paccp*). PaCCP production was dependent of cellular density suggesting a relation with Quorum Sensing (QS). Δ*paccp* strain showed increased extracellular proteolytic activity and swarming motility, decreased biofilm formation and a complete loss of twitching motility (TM). We have also demonstrated that intracellular low levels of pilin structural protein, PilA, prevented TM. It is proposed that PaCCP could be involved in the regulation of the expression of *pilA* and possibly other genes in *P. aeruginosa*. Decreased virulence of Δ*paccp* assayed in the *C. elegans* "slow killing" model pointed a role of PaCCP in *P. aeruginosa* pathogenesis. hCCP6 was produced in mammalian cells and showed the same deglutamylating activity on tubulins as a mouse ortholog. The protein acted specifically on α-tubulin what could represent a regulatory mechanism of CCPs activity. hCCP6 was able to hydrolyze the C-terminus glutamates stretch of TRAFD1, a substrate of CCP1, although it should be determined whether it is a natural substrate. The enzymatic activity of hCCP6 was inhibited by chelating agents but not by proteinaceus CP inhibitors. Interactomic assays have been performed and potential interacting partners of PaCCP and hCCP6 were identified. Some of them are proposed for further validation, which will contribute to a better understanding the functional relationships of these enzymes.

Abbreviations

Arm: armadillo

ASPA: aspartatoacylase

BBS: Bardet-Biedl Syndrome

BcCCP : Cytosolic carboxypeptidase of *Burkholderia cenocepacia* J2315

bCPA: bovine carboxypeptidase A

BmCCP: Cytosolic carboxypeptidase of *Burkholderia mallei* ATCC 23344

BZS: benzylsuccinic acid

c-di-GMP: cyclic diguanosine-5´-monophosphate

CCP: Cytosolic carboxypeptidase

CBB: cilia and basal bodies

COFRADIC: Combined FRActional Diagonal Chromatography

CP: carboxypeptidase

CPA: Carboxypeptidase A

CPB: Carboxypeptidase B

CPO: Carboxypeptidase O

CV: crystal violet

E. coli: *Escherichia coli*

EDTA: Ethylenediaminetetraacetic acid

EPS: extracellular polymeric substance

FliC: flagellin

GEMSA: guanidinoethylmercaptosuccinic acid

HA: hemagglutinin

hCCP6: human cytosolic carboxypeptidase 6

HMGB3: High mobility group protein B3

HRP: Horseradish peroxidase

IPTG: Isopropyl β-D-1-thiogalactopyranoside

Ki: Equilibrium dissociation constant

LB: Luria-Bertani

LCI: Leech carboxypeptidase inhibitor

MALDI-TOF: Matrix-assisted laser desorption ionization time-of-flight mass

MAPs: Microtubules Associated Proteins

MCP: Metallocarboxypeptidase

MLCK1: Myosin Light Chain Kinase 1

N-terminal: Amino-terminal

nano-ESI-LC-MS/MS: nano Electro Spray Liquid chromatography-tandem mass spectrometry

Nna1: Nervous system nuclear protein induced by axotomy

NvCl: Nerita versicolor carboxypeptidase inhibitor

PaCCP: Cytosolic carboxypeptidase of *Pseudomonas aeruginosa* MPAO1

PBS-T 0.01%: PBS containing 0.01% Tween 20 (v/v)

PBS-T: PBS containing 0.05% Tween 20 (v/v)

PBS: Phosphate Buffered Saline

Pcd: Purkinje cell degeneration

PCI: Potato carboxypeptidase inhibitor

PCP: Pro-carboxypeptidase

PCR: polimerase chain reaction

PDB: Protein Data Bank

PilA: pilin

Poly glutamate: polyE

PTM: posttranslational modifications

PVDF: Polyvinylidene difluoride

QS: Quorum Sensing

RARRES1: Retinoic Acid Receptor RESponder 1

RMS: Root Mean Square

RT: room temperature

SdCCP: Cytosolic carboxypeptidase of *Shewanella denitrificans* OS-217

SDS-PAGE: Sodium dodecyl sulfate- polyacrylamide gel electrophoresis

SSA: sensory, structural, and assembly

T4P: Type IVa pili

TCI: Tick carboxypeptidase inhibitor

TRAFD1: TRAF-type zinc finger domain-containing protein

TVCPT: *Thermoactinomyces vulgaris* carboxypeptidase T

wt: wild type

1. Introduction

1.1. Proteases

Proteases or peptidases are enzymes that hydrolyze peptide bonds between amino acids. These enzymes are critical for the complete degradation of proteins and peptides, so that the component amino acids can be re-used. Additionally, they are involved in post-translational processing of proteins, so that activation is delayed until the protein has reached the correct cellular or extracellular compartment, or the necessary developmental, physiological or environmental conditions prevail. Proteolytic cascades exist to amplify the original biological signal, such as in blood coagulation and apoptosis. The latter process highlights a function of proteases in the destruction of proteins and in switching off biological signals, including the degradation of peptide hormones and neurotransmitters.

A protease is targeted to a specific substrate by recognition of one or more binding sites. A substrate may bind directly at the protease active site, where the protease recognizes a sequence or structural motif in the substrate, or to a site known as an exosite, some distance from the enzyme active site (Rawlings et al., 2014).

1.2 Classification of proteases

One of the useful methods of grouping proteases currently in use is according to the protein residues that are essential for catalysis. Proteolytic enzymes can be described as serine, cysteine, threonine, aspartic, glutamic, asparagine or metallo catalytic type (and just a few remain of unknown catalytic type). Metallo-proteases exert catalysis through one or two bound divalent metal ions, mostly zinc but also sometimes cobalt, nickel, or manganese.

Proteolytic enzymes can also be grouped by the kinds of reaction they catalyze. On this basis, they can be classified into groups such as endopeptidases, omega-peptidases, exopeptidases, aminopeptidases, carboxypeptidases, dipeptidyl-peptidases, tripeptidyl-peptidases, peptidyl-dipeptidases and dipeptidases. These groupings are important in the Enzyme Commission List of the Nomenclature Committee of International Union of Biochemistry and Molecular Biology (IUBMB, 1992).

The newest method of classification of proteases is by molecular structure and homology because it depends on the availability of data for amino acid sequences and three-dimensional structures. This method is specifically implemented in the MERO database (Rawlings et al., 2014) .

The MEROPS database classifies proteases along evolutionary principles, so that homologous sequences are included in the same family, and homologous tertiary structures are included in the same clan, according to the methods of Rawlings and Barrett (1993). There are nearly 250 families of proteases in MEROPS Release 9.9 (Rawlings et al., 2014) . Each family is identified by a letter, representing the catalytic type of the proteolytic enzymes it contains, Aspartic (A), Cysteine (C), Glutamic (G), Metallo (M), Asparagine (N), Mixed (P), Serine (S), Threonine (T), Unknown (U), together with a unique number. Some families are divided into subfamilies because there is evidence of a very ancient divergence within the family (Rawlings et al., 2014).

Although the families are the largest groupings of proteases that can be proven rigorously to be homologous, there are persuasive lines of evidence that many of the families do share common ancestry with others. The term "clan" is used to describe sets of families in which all of the proteins have diverged from a single ancestral protein, but they have diverged so far that their relationship can no longer be proved by comparison of the primary structures. The best kind of evidence to support the formation of a clan is similarity in three-dimensional structures when the data are available, but the arrangement of catalytic residues in the polypeptide chains and limited similarities in amino acid sequence around the catalytic amino acids are also taken into account. The name of each clan is formed from the letter for the catalytic type of the proteases it contains followed by an arbitrary second capital letter (Rawlings et al., 2014).

1.3 Substrate specificity of proteases and related nomenclature

Based on the hydrolytic activity of the endopeptidase papain on alanine peptides of varying length, Schechter and Berger (2012) hypothesized that the active-site of the enzyme had seven "subsites" interacting with the side chains of the residues flanking the scissile bond. They subdivided the cleft into an N-terminal or upstream region with respect to the scissile bond, the "non-primed" side, and a C-terminal or downstream region, the "primed"

side (Figure 1.1). Consequently, the substrate side chains on the non-primed side away from the scissile bond are now termed P1, P2, P3, etc. and their cognate enzyme subsites S1, S2, S3, etc. On the primed side, substrate side chains P1', P2', P3', etc. are accommodated in subsites S1', S2', S3', etc (Figure 1.1).
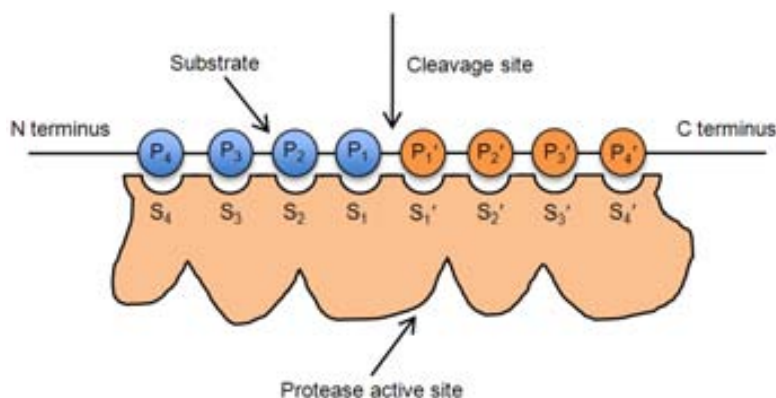
Figure 1.1 Nomenclature of the substrate specificity of proteases according to Schechter and Berger (2012).

These findings arise from the fact that proteases must provide a scaffold that is large enough to bind a substrate efficiently and then to exert the cleavage reaction. This also entails that the reactivity of a scissile peptide bond depends not only on the flanking residues in S1 and S1' but also on the side chains of residues further upstream and downstream (Schechter and Berger, 2012). In carboxypeptidases (CP), the C-terminal residue to be excised had to be placed in S1' and that residues upstream in the substrate occupied subsites S1, S2, etc., thus giving rise to a shorter active-site cleft than in endopeptidases.

1.4 Metallocarboxypeptidases of the M14 family

Metallocarboxypeptidases (MCPs) of Clan MC, family M14 (MEROS database classification) are zinc-containing exopeptidases that catalyze the removal of C-terminal amino acids from proteins and peptides. The two best-studied subfamilies are M14A and M14B according to the MEROPS database (Arolas et al., 2007; Vendrell et al., 2000). Much less information is available for the M14C subfamily that comprises the bacterial orthologs of d-glutamyl-(l)-meso-diaminopimelate peptidase I (Hourdou et al., 1993) and they will not be described in this work. More recently, a group of cytosolic carboxypeptidases (CCPs) have been described and classified into a new M14D subfamily (Kalinina et al., 2007; Rodriguez de la Vega et al., 2007).

MCPs can also be classified, regarding to their substrate specificity into A-like enzymes, with a preference for hydrophobic C-terminal residues, and B-like enzymes, which cleave only C-terminal lysine or arginine residues (Aviles et al., 1993; Vendrell et al., 2000). O-type enzymes recognize specifically aspartate or glutamate residues (Lyons and Fricker, 2011) and some MCPs have broad substrate specificity (Estebanez-Perpina et al., 2001; Grishin et al., 2008).

An alternative classification of MCPs is also possible based on the physiological role of these enzymes. Some of them participate only in alimentary digestion and are sometimes called pancreatic MCPs, while others are involved in more selective reactions (Reznik and Fricker, 2001; Vendrell et al., 2000; Wei et al., 2003) and have been named regulatory MCPs. These latter proteins perform a variety of important biological functions mainly in non-digestive tissues and fluids, acting in blood coagulation/fibrinolysis, inflammation and local anaphylaxis, pro-hormone and neuropeptide processing, cellular response, and others.

1.5 M14A metallocarboxypeptidases

There are nine members of the M14A subfamily of MCPs in mammals: CPA1, CPA2, CPA3, CPA4, CPA5, CPA6, CPB, CPU, and CPO. All appear to be soluble proteins that are normally synthesized as inactive precursors (pro-carboxypeptidases, PCPs) containing a preceding signal peptide of 15-22 residues. The PCPs of the M14A subfamily contain a 90-95-residue N- terminal pro-region that folds in a globular independent unit called the "activation segment", linked to the enzyme through a connecting part (Coll et al., 1991; Guasch et al., 1992). This zymogenic activation is important to restrict activity to where it is needed. In addition to its role in enzymatic inhibition, the N-terminal pro-domain is also thought to function in protein folding. All types of substrate specificity can be seen in this subfamily; CPA1 to 6 are A-type enzymes, CPB and CPU are B-type enzymes and CPO is an O-type MCP (Vendrell et al., 2000).

In mammals, PCPA1, PCPA2, and PCPB function in the degradation of dietary proteins after secretion and trypsin-promoted limited proteolysis in the duodenum (Vendrell et al., 2000). CPA3 is found in the secretory granules of mast cells mainly in its active form and has a role in regulating innate immunity responses (Pejler et al., 2009). CPU is produced and secreted by the liver as the inactive precursor, commonly referred to as thrombin-

activatable fibrinolysis inhibitor (TAFI). Its active form (TAFIa) controls the rate of fibrinolysis by cleaving C-terminal Lys residues that serve as plasminogen binding sites (Foley et al., 2013). CPA4 is expressed in some carcinomas and probably functions in neuropeptides processing (Tanco et al., 2010). CPA5 cleaves aliphatic C-terminal residues and is present in discrete regions of rodent pituitary. It is probably involved in the non-basic processing of pro-opiomelanocortin-derived peptides (Wei et al., 2003). CPA6 has preference for large hydrophobic C-terminal amino acids and is highly expressed in the adult mouse olfactory bulb and broadly expressed in embryonic brain and other tissues. Unlike most members of the M14A subfamily, CPA6 is secreted in its active form after processing in the secretory pathway by furin-like endoproteases. A disruption in the human CPA6 gene is linked to Duane syndrome and is supposed to play a role in the regulation of neuropeptides in the extracellular environment (Lyons et al., 2008). CPO is a membrane-anchored intestinal enzyme that cleaves acidic residues from dietary proteins. CPO is functional without the presence of a prodomain and does not appear to be regulated by pro-peptide cleavage (Lyons and Fricker, 2011) .

1.6 M14B metallocarboxypeptidases

M14B subfamily of MCPs is formed by five enzymatically active CPs (CPN, CPE, CPM, CPD, and CPZ) and other non-active members including CP-like protein X1 (CPX1), CP-like protein X2 (CPX2) and a protein designated aortic CP-like protein (ACLP). CPX1, CPX2 and ACLP lack critical active site and substrate-binding residues that are considered necessary for CP activity. These proteins may function as binding proteins or may display catalytic activity toward other substrates (Arolas et al., 2007). The active members of the M14B subfamily exhibit a strict specificity for removing C-terminal basic amino acids arginine and lysine from peptides/proteins and are therefore referred to as basic metallo-CPs (Vendrell et al., 2000).

M14B MCPs does not contain an activation domain; instead they have an 80-90 amino acid domain located immediately after the C-terminus of the CP domain. This domain has structural homology to transthyretin and other proteins and has been proposed to function as a folding domain. The M14B family members do not require proteolytic cleavages to

generate the active form and rely on subcellular localization to prevent inappropriate cleavages that would otherwise damage the cell (Arolas et al., 2007).

CPE is found primarily in endocrine and neuroendocrine cells and is involved in the biosynthesis of peptide hormones and neuropeptides by removing C-terminal basic residues. CPE also functions as a pro-hormone sorting receptor for the regulated secretory pathway. In humans, deregulated CPE has been associated with numerous diseases, such as diabetes, Alzheimer's disease and cancers (Arolas et al., 2007; Pla et al., 2013). CPN circulates in plasma in a high molecular weight complex that includes a binding protein. CPN plays an important role in protecting the body from excessive buildup of potentially deleterious peptides that normally act as local autocrine or paracrine hormones. Among CPN substrates are kinins, anaphylatoxins, creatine kinase, plasminogen receptors hemoglobin and SDF-1α (Skidgel and Erdos, 2007).

Unique among the MCPs is the attachment of CPM to the cell surface via a glycosyl-phosphatidyl-inositol glycan anchor although also circulates in body fluids. It is supposed that CPM regulates or processes peptides generated by inflammatory processes (Deiteren et al., 2007). CPZ contains a cysteine-rich domain with 20–30% homology to the one present in the frizzled family of Wnt receptors. CPZ is secreted and located in the extracellular matrix and binds to Wnt-4 enhancing its activity by proteolytically removing the arginine at the C-terminus. It has been proposed that CPZ-mediated Wnt activation might be a general mechanism of thyroid hormone action during development (Wang et al., 2009).

CPD was localized to the trans-Golgi network, the secretory and reuptake pathways, and transiently on the cell surface. Unlike all other members of the CP family, CPD contains multiple CP domains followed by a transmembrane domain and cytosolic tail (Reznik and Fricker, 2001). In all vertebrates examined, CPD contains three CP domains; the first two of which are enzymatically active while the third domain is not. CPD is thought to play a role in the processing of proteins that are initially processed by furin or related endopeptidases present in the trans Golgi network, such as growth factors and receptors (Sidyelyeva et al., 2010).

1.7 M14D cytosolic metallocarboxypeptidases

CCPs were identified after a detailed genomic scan of nervous system nuclear protein induced by axotomy (Nna1)-like peptidases (Rodriguez de la Vega et al., 2007). The Nna1 gene transcript was upregulated in spinal cord of mice subjected to sciatic nerve transection or crush injury (Harris et al., 2000) and was identified as the gene mutated in the classical recessive mouse mutant Purkinje cell degeneration (*pcd*) (Fernandez-Gonzalez et al., 2002). *Pcd* mice exhibit a complex phenotype although ataxia associated with the loss of cerebellar Purkinje cells during early adulthood is its most prominent features (Wang et al., 2006).

Prokaryotic CCPs are restricted to around 100 gram-negative bacteria most of them proteobacteria. Eukaryotic CCPs are widely distributed in holozoans and protists, although they are only found in the basal lineages of plants and fungi (Rodriguez de la Vega et al., 2007). Most bacteria contain a single copy of the *CCP* gene while the number of paralogs varies between different eukaryotic organisms. For instance, it is possible to find one copy of the gene in fungi to six copies in most mammals (Rodriguez de la Vega et al., 2007).

CCPs range from 400 to 2000 residues in length and comprise a common MCP domain of around 300 residues. M14D peptidases conserve the Zinc ligands and the catalytic residues of MCPs although the motifs where they occur differ from the ones defined for M14A, M14B, or M14C subfamilies. In addition to zinc ligands, other important MCPs residues involved in catalysis or in substrate binding are conserved in CCPs (Rodriguez de la Vega et al., 2007). An amino-terminal (N-terminal) conserved domain was also identified that is around 150 residues in length and contains three highly conserved motifs (Figure 1.2). Sequence homology and fold recognition analysis showed that the conserved N-domain of CCPs is different to the M14A pro-domain. It was suggested that it could act as a folding domain and might also act as a regulatory or a binding domain (Rodriguez de la Vega et al., 2007). Some eukaryotic CCPs present large extensions at the N- and/or the C-terminus and orthologous of CCP5 have an insertion in the CP domain (Figure 1.2). Additional conserved domains are found in sequence extensions including armadillo (Arm) like domain that has been linked to protein–protein interactions (Madhurantakam et al., 2012).
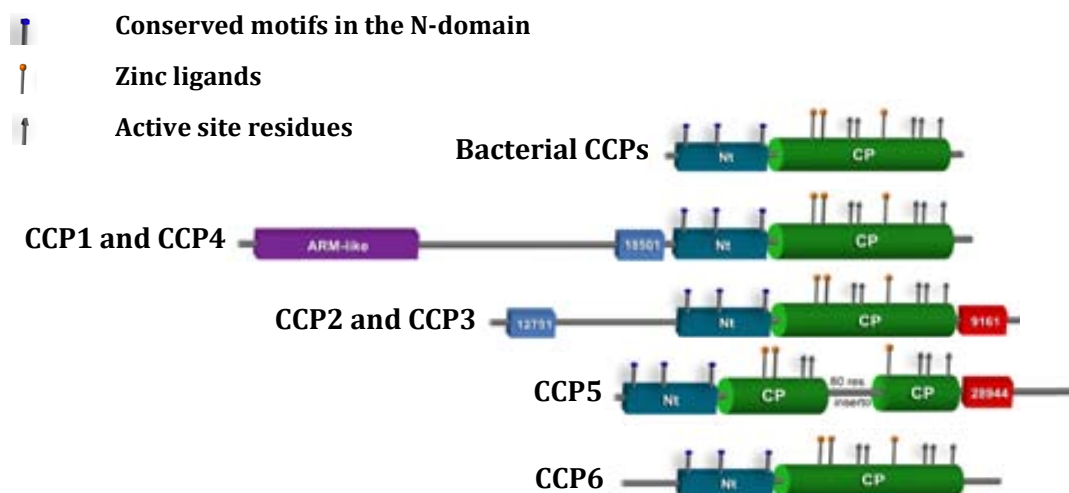
Figure 1.2. Domain architecture of bacterial and mammalian CCPs. Conserved N-terminal domain (Nt) and carboxypeptidase domain (CP) are represented as blue boxes and green tubes, respectively. Other conserved motifs among subgroups are represented as colored boxes, with numbers indicating the PfamB family. Conserved motifs at the N-terminal domain and active site residues of the peptidase unit are indicated. This figure is adapted from Rodríguez de la Vega et al., (2007).

The majority of CCPs lack signal peptide, except some bacterial CCPs, and it has been shown that eukaryotic CCPs display intracellular localization. The specific enzymatic activity of some members of the eukaryotic CCPs subfamily was described by Rogowski et al., (2010). CCP1, CCP4 and CCP6 catalyze the shortening of polyglutamate chains attached to the C-terminal of tubulins as a reversible posttranslational modification (PTM) (Figure 1.3). CCP5 specifically removes the branching point glutamates. Additionally, CCP1, CCP4 and CCP6 also remove gene-encoded glutamates from the C-terminal of previously detirosynated tubulins to generate Δ2 tubulin (Figure 1.3) and from the C-terminal of other proteins like Myosin light chain kinase 1 (MLCK1) and telokin (Rogowski et al., 2010). New substrates of CCP1, including ribosomal proteins, translation factors and high mobility group proteins were identified by positional proteomics approach. It was shown that CCP1 processes glutamates as well as C-terminal aspartates (Tanco et al., 2014 submitted). Sahab et al., (2011) proposed CCP2 as the candidate for the long sought-after tubulin tyrosine carboxypeptidase. Nevertheless, results from our laboratory suggest that CCP2 and CCP3 also act as deglutamylating enzymes (Tort et al., 2014, submitted).
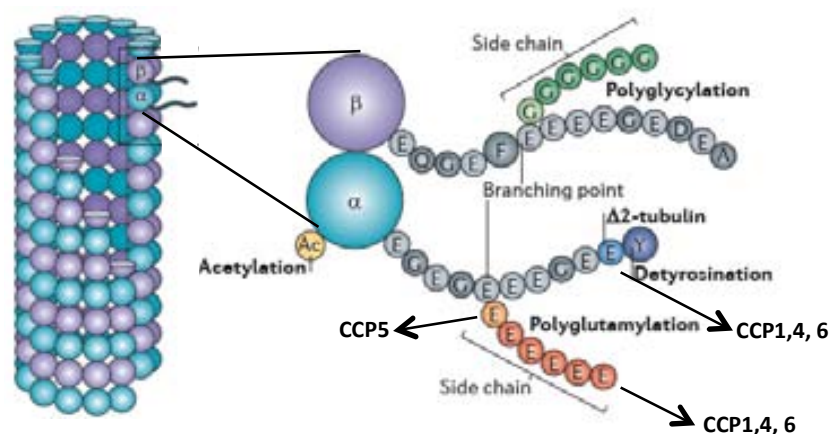
Figure 1.3. Tubulin posttranslational modifications and CCPs. Schematic representation of the α-tubulin–β-tubulin dimer and its associated modifications. The carboxy-terminal tails of both tubulins are represented as amino acid sequences (mouse α1A-tubulin from residue 441 to the end and β2B-tubulin from 432 to the end). Both α-tubulin and β-tubulin can be modified by polyglutamylation and polyglycylation on different Glu residues within those tails. Together with detyrosination at the C-terminus and the follow-up removal of the penultimate Glu residue (which generates Δ2-tubulin), these modifications are specific to the C-terminal tails of tubulin. Residues hydrolyzed by CCP1, 4, 5 and 6 are indicated with arrows. This figure is adapted from Janke and Bulinski (2011).

Recently, it has been hypothesized that the primary function of CCPs could be related to cilia and basal bodies (CBBs), based on phylogenetic analysis and intracellular localization(Rodriguez de la Vega Otazo et al., 2012). It was shown that the taxonomic distribution of eukaryotic CCP genes correlates with the presence of CBBs. The potential functions of CCP orthologs were mapped according to their occurrence either in organisms containing only primary cilia or only motile flagella to correspondingly assign them a sensory, structural, and assembly (SSA) or motile function. It was inferred that CCP1 and CCP4 orthologs could be associated with the assembly and elongation of the microtubule network, which takes place at the CBB level in the building of primary or motile ciliary axonemes and at the axonal level during reinervation and axogenesis. The main function of CCP2 and CCP3 orthologs may be associated with motile cilia. CCP5 orthologs have motile functions and CCP6 orthologs may play a role in the *de novo* assembly of CBB and have an SSA function not restricted to primary cilia (Rodriguez de la Vega Otazo et al., 2012).

The intracellular localization of human CCPs was study in HeLa and ciliated NIH-3T3 cells (Rodriguez de la Vega Otazo et al., 2012). CCP1 in HeLa cells was localized in the nucleus and cytoplasm, showing a granular distribution in interphase and dividing cells. Human CCP2 was found in centrioles and basal bodies of ciliated and nonciliated HeLa cells

across all stages of the cell cycle. CCP5 was localized in the nucleus of nondividing HeLa cells, in the mitotic spindle during the different steps of cell division and in the intercellular bridges formed during cytokinesis. Finally, CCP6 was found in the Golgi apparatus and centrioles of interphase cells. In mitotic cells, the CCP6 signal was stronger in centrioles and colocalized with γ-tubulin (centrosomes and spindle poles bodies). Moreover, CCP6 was detected in the basal bodies of ciliated NIH-3T3 cells (Rodriguez de la Vega Otazo et al., 2012).

The identification of the CCPs, together with other enzymes that modify the C-terminal of tubulins has enabled functional studies that strengthen the "tubulin code" hypothesis. The numerous tubulin gene products and PTMs have the potential of fine-tuning of processes like the attachment strength, processivity or speed of motor proteins, or the coordination of MAPs and motors with one another (Ghosh-Roy et al., 2012; Sirajuddin et al., 2014; Wehenkel and Janke, 2014). Aberrations in the normal pattern of PTMs could directly or indirectly lead to disease like ciliopathies and neurodegeneration (Janke and Bulinski, 2011).

Unlike mammalian CCPs, the information available for bacterial CCPs is very scarce. This thesis is focus in the CCP from *Pseudomonas aeruginosa* and the human CCP6.

1.8 Structure of M14A and B metallocarboxypeptidases

1.8.1. Catalytic domain

There are several three-dimensional structures available for A/B-MCPs, either in their active or zymogenic state, unbound or as inhibitor or product complexes (Vendrell et al., 2000). The study of these structures reveals that the catalytic domains are closely related with a compact globular shape, which resembles the volume obtained when a cone is extracted from a sphere (Figure 1.4). The domain is called funnelin for the funnel-like opening at the top, inside which lies the active-site cleft at the base of the opening (Pallares et al., 2005).
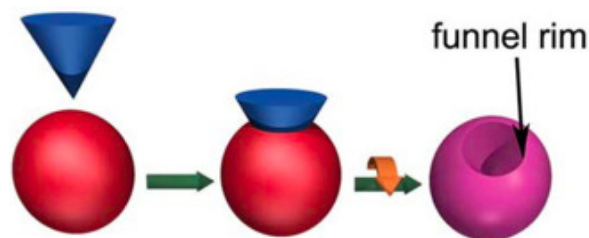
Figure 1.4. Carboxypeptidase funnel-like structure. Scheme illustrating the generation of an inverted spherical cone (magenta) through the intersection of a regular cone (blue) and a sphere (red). The resulting border is like a funnel rim. This figure is adapted from Gomis-Ruth, (2008).

Funnelin catalytic domains show a consensus α8/β8-topology conforming to an α/β-hydrolase fold (Figures 1.5A,B) (Ollis et al., 1992). They contain a central doubly-wound eight-stranded β-sheet (β1–β8) of mixed parallel/antiparallel topology with strand connectivity +1, +2, −1x, −2x, −2, +1x, −2. The core of the sheet consists of four parallel coplanar central strands (β3–β5 and β8) and the catalytic site is located at the C- terminal end of these strands. These four strands are flanked by two parallel strands on the top (β6–β7) and a β-ribbon (β1–β2) at the bottom (Figures 1.5A,B), which confer the overall twist on the sheet. This leads to a concave front side of the sheet, which accommodates helices α5 and α7, as well as the active-site cleft. All funnelins studied except *Lophonetta specularioides* (crested duck) CPD2, *Homo sapiens* CPM, and *Thermoactinomyces vulgaris* CPT (TVCPT) contain an additional short helix inserted after the upper strand β7 (Figure 1.5B). At the back of the molecule, the convex side of the sheet embraces helices α1–α4, α6 and α8 and the surface N- and C-termini of the molecule.

The funnel-like access to the active site is shaped by a series of irregular segments of varying length, namely Lβ8α8, Lβ5β6, Lβ7α7, Lβ3α2, and, in particular, the loop Lα4α5 in the lower left of the molecule (Figure 1.5B). These segments are required for interactions of the catalytic moiety with the pro-domain in the zymogen and with cognate protein inhibitors (Pallares et al., 2005). Lα4α5 delimits the active-site cleft and its conformation is stabilized by a disulfide bond (Cys138-Cys161 in bovine CPA1 (bCPA1)).

The catalytic zinc ion resides at the bottom of the active-site cleft and is coordinated by His69 Nδ1 and by Glu72 atoms Oε2 and Oε1 in a bidentate manner. These two residues are provided by Lβ7α5 in a consensus sequence characteristic of funnelins, HXXE (Hooper, 1994). A third protein zinc ligand is His196 Nδ1 at the end of strand β5 (Figure 1.5A), which

together with a metal-bound solvent molecule gives rise to an overall distorted tetrahedral coordination sphere.
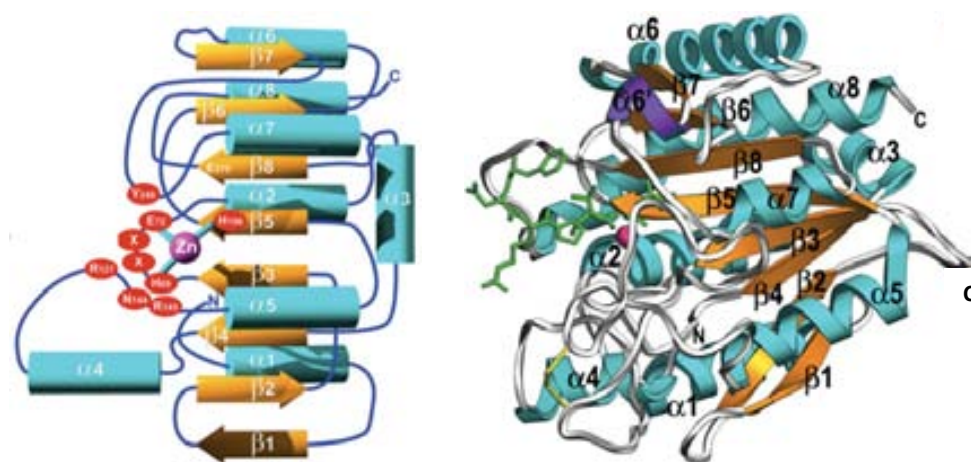
Figure 1.5. Structural features of the funnelin catalytic domain of MCPs. A. Topology scheme illustrating the consensus secondary-structure elements (α-helices in cyan and β-strands in orange) found in all funnelins analyzed. The highlighted amino acids are contained in the characteristic set of conserved residues, HXXE+R+NR+H+Y+E, and are numbered according to the bovine CPA1 numbering. The metal ion is shown as a magenta sphere. B. Richardson diagram of HCPA4. The consensus secondary-structure elements and the metal site are colored and labeled as in A. An additional element to the consensus α8/β8-structure is helix α6' (purple ribbon). A bound and subsequently cleaved hexapeptide (green stick model) is also displayed. The N- and C-termini are labeled for reference. This figure is adapted from Gomis-Ruth, (2008).

The only major differences of CP domains of M14B MCPs in comparison with those of M14A MCPs are within the length of the loops between the 16-sheets and 16-helices. These differences cause the M14B subfamily members to have a more restrictive substrate-binding pocket. The larger loops within the area surrounding the substrate-binding pocket explain the inability of proteinaceous inhibitors of M14A MCPs to bind to M14B MCPs (Arolas et al., 2007).

1.8.2. Pro-domains of M14A subfamily MCPs

The conformation of the structurally-determined pro-regions of M14A subfamily is very similar despite a low degree of sequential homology observed in some cases. The topology of these pro-regions consists of an open sandwich with a four-stranded antiparallel16-sheet on one side and two anti-parallel 16helices on top with an extended 16helix at the C-terminus connecting the globular domain to the enzyme moiety (Figure 1.6).
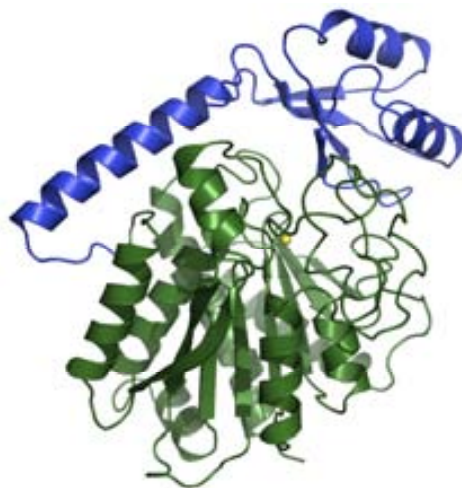
Figure 1.6. Structure of pro-CPA4. Ribbon plot of human pro-CPA4 (PDB access code 2BOA), with the prodomain in blue and the mature enzyme moiety in green. The zinc ion is shown as a yellow sphere. Graphic prepared with PYMOL.

The interaction with the enzyme takes place at the side of the sheet. There are no disulfide bonds in the activation segment of PCPs. Limited proteolysis at the end of this segment activates the zymogen through release of the pro-domain, which is degraded *in vivo*. *In vitro*, the isolated pro-domain acts as an autologous inhibitor in trans (Vendrell et al., 2000).

Overall, the pro-domain/mature enzyme complexes of funnelins involve large interfaces (Garcia-Castellanos et al., 2005) but few direct contacts. In A-type zymogens, pro-domain residues Asp36A and Trp38A interact with Arg71 and Phe279/Tyr198, respectively, of the mature enzyme moiety. The human and porcine B-type zymogens have an additional salt bridge made up by Asp41 of the prodomain and Arg145 in the CP domain. The lack of this interaction accounts for the capacity of A-type zymogens to cleave small fluorigenic substrates, while B-type zymogens are inactive (Arolas et al., 2007).

1.8.3. Transthyretin-like domain of M14B MCPs

Transthyretin-like domain has a rod-like shape and is located downstream of the catalytic domain in M14B MCPs (Figure 1.7). Its N- and C-termini are on opposite sides of the rod, which folds into an all-β seven-stranded flat β-sandwich. The β-sandwich consists of two layers of three mixed and four antiparallel strands, respectively, which are held together by a hydrophobic core. Interactions between the transthyretin-like and the catalytic domains are established by hydrophobic residues and a salt bridge formed by an

aspartate, which is conserved among M14A funnelins, with arginine residues within the transthyretin-like domains.
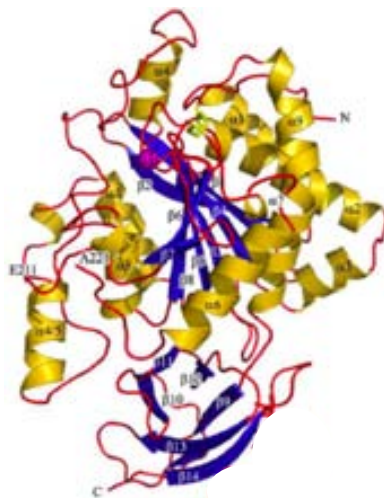
Figure 1.7. Structure of the human CPN. Ribbon representation of the catalytic domain on top and the transthyretin-like domain at the bottom. α-helices and β-strands are represented as golden helices and blue arrows, respectively. The zinc ion is shown as a purple sphere. This figure is adapted from Keil et al., (2007).

1.8.4. Active-site of MCPs

The length, shape, and nature of the active-site cleft have been investigated by a series of biochemical and structural studies with transition-state analogs, reaction products and inhibitors (Gomis-Ruth, 2008). The cleft is rather shallow, which explains the capacity of funnelins to cleave a large variety of well-folded proteins, as only few contacts between the enzyme and the C-terminal stretch of a substrate are required to fix the latter to the active-site (Figures 1.5A,B). The peptide chain, in a nearly extended conformation at P1, P1, P2 and P3, is bent at P4 and folds back to creep along the molecular surface (Figure 1.5B). The scissile residue side chain is accommodated in the S1' site, the specificity pocket, which is delimited by the protein side chains in positions 194, 203, 207, 243, 247, 250, 253–255 and 268 (bCPA1 numbering).

The residues of M14A MCPs, important for substrate binding and catalysis, can be grouped in five subsites (Aviles et al., 1993; Vendrell et al., 2000): S1' (including Asn144, Arg145, Tyr248 and Xaa255); S1 (Arg127, Leu/Ile247 and Glu270); S2 (Arg71, Asp142, Ser197, Tyr198 and Ser199); S3 (Phe279); and S4 (Glu122, Arg124 and Lys128). The terminal carboxylate group of the peptide substrate is fixed by Asn144, Arg145 and Tyr248, while the

carbonyl group of the scissile peptide bond becomes positioned near Glu270, Arg127 and the zinc ion.

Residues in the different subsites are highly conserved among M14A MCPs, and differences are only restricted to residues of position 255 that define the substrate specificity. In most cases, an Asp255 accounts for B-type specificity in M14A MCPs (CPB and TAFIa) (Sanglas et al., 2010) while in A-type enzymes, a hydrophobic residue is found in this position (Ile in CPA1, CPA2 and CPA4; Leu in CPA3; Val in CPA5 and Met in CPA6) providing specificity for aliphatic and aromatic residues (Gomis-Ruth, 2008; Lyons et al., 2008; Tanco et al., 2010). The selectivity of CPO for acidic amino acids is unique among the mammalian M14A MCPs and contains an Arg at the position 255 (Lyons and Fricker, 2011). TVCPT, with mixed A+B specificity, has a Thr at this position (Grishin et al., 2008) and the CPA from *Helicoverpa armigera*, able to remove hydrophobic, positively and negatively charged C-terminal residues, has a Ser (Estebanez-Perpina et al., 2001).

The critical active site residues are conserved in the M14B MCPs, however, some of the important substrate-binding residues are different to M14A MCPs. M14B active members contain either a Gln (CPE, CPD-I, CPD-II, CPN, CPM) or a Ser (CPZ) in position 255. The specificity of M14B MCPs is provided by an Asp in the position equivalent to 207 of BCPA1 (a glycine in M14A MCPs). Position 207 plays a minor role in the S1' binding pocket.

In mammalian CCPs, the residue in the position equivalent to Ile 255, also appears to be determinant of substrate specificity. The alignment of CCP1 and members of the M14A and B subfamilies with aspartatoacylase (ASPA), a related MCP, showed that CCP1 has an Arg in the position equivalent to Ile255 that was essential for CCP1 activity *in vivo* (Wu et al., 2012). It has been shown that mammalian CCPs display specificity for C-terminal glutamates (Rogowski et al., 2010) (Tort et al., unpublished results).

1.9 Catalytic mechanisms of MCPs

The studies of the mechanisms of MCPs have been based on complexes with substrates, products and transition-state analogs of bCPA1 and related enzymes. Two hypotheses have been proposed to explain the catalytic mechanism of bCPA1, the acyl pathway hypothesis and the promoted-water pathway also known as the general base-general acid pathway (Lipscomb and Strater, 1996; Mock and Zhang, 1991). Arguments are

currently in favor of the promoted-water pathway where a penta-coordinated zinc polarizes a water molecule to attack the scissile peptide bond of the substrate, leading to a tetrahedral intermediate/transition state (Lipscomb and Strater, 1996; Wu et al., 2010).

In the unbound enzyme, the catalytic zinc ion coordinates to His69, Glu72, His196, and a solvent molecule, which in turn hydrogen-bond to the side chain of Glu270. The presence of the polarizing glutamate carboxylate of Glu270 may provide the degree of solvent nucleophilicity required for the enzymatic reaction. Protease mutants, in which this glutamate was replaced by non-acidic residues, were catalytically inactive or showed only residual activity (Gomis-Ruth, 2008).

Before catalysis, the substrate is accommodated within the active-site cleft with P´1 side chain of the scissile residue interacting with the specificity pocket. Substrate binding leads to the rearrangement of the side chain of Tyr248, which moves up to 12 Å from a surface "up" position in the unbound enzyme to a "down" position to participate in substrate fixation (Garcia-Saez et al., 1997). The function during catalysis of Tyr248 has been a matter of much debate although X-ray crystallography and computational analysis of protein motion, indicated that the phenolic hydroxyl is hydrogen bonded to the substrate during the entire process (Fernandez et al., 2010).

Once the substrate is bound, the scissile carbonyl oxygen enters the metal coordination sphere, which becomes penta-coordinate. The carboxylate of Glu270, acts as a general base and abstracts a proton from the water and the resulting hydroxide performs a nucleophilic attack on the scissile amide carbonyl carbon. A negatively-charged tetrahedral reaction intermediate is form, stabilized by the zinc ion, and the side chain of Arg127. Glu270 subsequently acts as a general acid catalyst resulting in the reaction intermediate collapsing to two products containing a carboxylate and an α-ammonium group, respectively. Eventually, the amino acid in S´1 leaves and the enzyme is ready for subsequent catalysis.

1.10 Inhibitors of MCPs

Kinetic and spectroscopic experiments have shown that MCPs are all inhibited by chelating agents such as 1,10-phenanthroline (Felber et al., 1962) or EDTA, by small carboxylic acid anions, D-amino acids or certain dipeptides that bind to the S1' and S1 subsites (Vendrell, 1999). The most popular organic inhibitors are benzyl succinic acid for A-

like enzymes, and guanidinoethylmercaptosuccinic acid (GEMSA) for B-like enzymes, with equilibrium dissociation constants (Ki) in the order of mM to nM (Lee et al., 1999). GEMSA has been crystallized occupying the specificity pocket of the second domain of duck CPD (Figure 1.8) (Aloy et al., 2001).
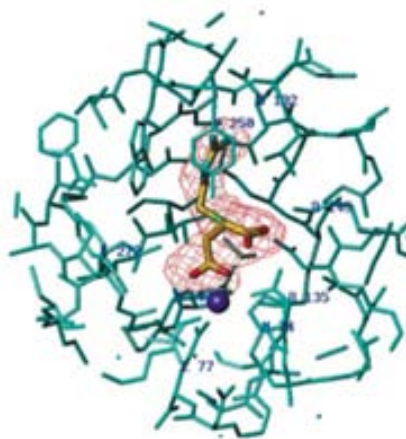
Figure 1.8. Complex between the second domain of duck CPD with inhibitor GEMSA. The inhibitor carboxylate groups coordinate the catalytic zinc ion (blue sphere) in a bidentate manner and Arg145/Arg135 from the protein, respectively. The 2-guanidinoethyl moiety of the inhibitor is placed in the specificity pocket. Some residues are labeled. This figure is adapted from Aloy et al., (2001).

Few proteinaceous inhibitors, that specifically inhibit M14A MCPs, have been identified so far. Two endogenous inhibitors, latexin and RARRES1, have been obtained from rat and human tissues. Exogenous inhibitors, PCI, ACI, LCI, TCI, HLTCI (Gong et al., 2007), SmCI (Alonso-del-Rivero et al., 2012) and NvCI, have been found in *Solanacea* plants, parasites and marine invertebrates. No proteinaceous M14B MCPs inhibitors have been found to date. It has been proposed that RARRES1 inhibits CCP2.

1.10.1. Endogenous MCP inhibitors

The endogenous inhibitor latexin is larger (223 residues in humans) than exogenous inhibitors and does not display sequence similarity with known MCP inhibitors. Latexin is expressed in the developing brain and apparently functions in inflammation (Aagaard et al., 2005). The recombinant form of latexin appears as a hardly reversible, non-competitive and potent inhibitor of M14A MCPs (Ki 1-4 nM). The structure of latexin (Pallares et al., 2005) comprises two topologically equivalent domains linked by a connecting segment encompassing an α-helix. These domains are reminiscent of cysteine protease inhibitors (cystatins) and consist of an α-helix surrounded by a curved β-sheet (Figure 1.9A). The

native target(s) of latexin could include CPA6, which is expressed in brain and other tissues and/or cytosolic MCPs like CCPs, due to the presumed cytosolic localization of the protein (Arolas et al., 2007).
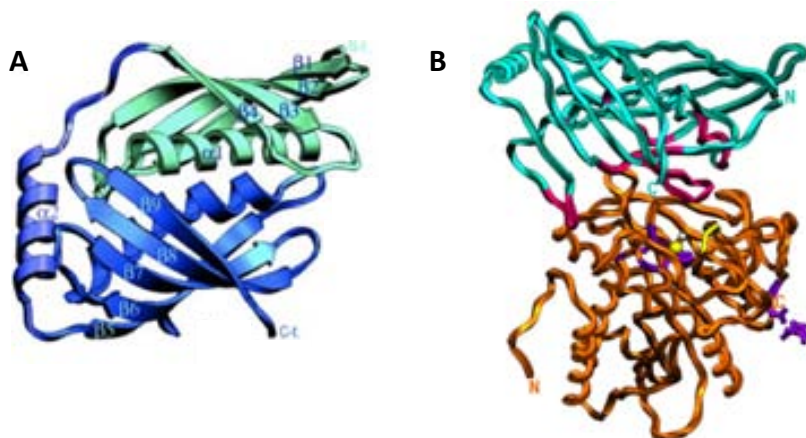
Figure 1.9. Structure of latexin and of its complex with human CPA4. A. Ribbon plot of human latexin, the N-terminal subdomain (cyan) and the C-terminal subdomain (blue) are shown. B. Ribbon plot of the inhibitor-enzyme complex showing hCPA4 in orange and latexin in cyan. Inhibitor segments contacting the protease are in magenta. The metal ion is shown as a yellow sphere. This figure is adapted from Pallares et al., (2005).

Latexin binds to M14A MCPs in a way that mimics the autologous inhibition of the activation segment of PCPs. The structure of latexin in complex with human CPA4 (Figure 1.9B) showed that the inhibitor interacts with the top of the enzyme through the interface of its two domains, clamping a loop encompassing residues Asp273-Pro282 of CPA4 moiety (Pallares et al., 2005). The contact area of the complex is larger than the typical protease-inhibitor interface. However, only a few interactions are formed with residues of the active site (with S1' and S2 subsites), which explains the flexibility of latexin in inhibiting all vertebrate M14A MCPs tested.

Retinoic acid receptor responder 1 (RARRES1) is a 294 amino acids protein that shares sequence similarity with latexin (~30% identity over 220 residues). RARRES1 is expressed in skin fibroblasts and epithelia, and is encoded by a retinoic acid receptor-responsive gene. It has been implicated in both the therapeutic effects of retinoic acid, in psoriasis (Nagpal et al., 1996) and in tumor suppression (Jing et al., 2002). RARRES1 is a type III membrane protein found in a complex with several proteins involved in the regulation of microtubule function. Among the interacting partners was found the CCP2, a member of the M14D subfamily (Sahab et al., 2011).

1.10.2. Exogenous MCP inhibitors

The potato carboxypeptidase inhibitor (PCI), highly homologous to the inhibitor from tomatoes, is a small protein (39 residues) that competitively inhibits M14A MCPs with a Ki in the nanomolar range (Hass et al., 1975). PCI is synthesized with a 29-residue N-terminal signal peptide, a 27-residue N-terminal pro-region, the 39-residue mature region, and a 7-residue C-terminal extension (Villanueva et al., 1998). The structure of PCI was solved in complex with bCPA (Rees and Lipscomb, 1982), displaying a 27-residue globular core that lacks regular secondary structures (Figure 1.10A) except for a short 3_{10} helix and a very small β-sheet made by two short antiparallel strands. PCI is stabilized by three disulfide bridges and displays a peculiar disulfide-stabilized loop scaffold known as cystine knot or T-knot.
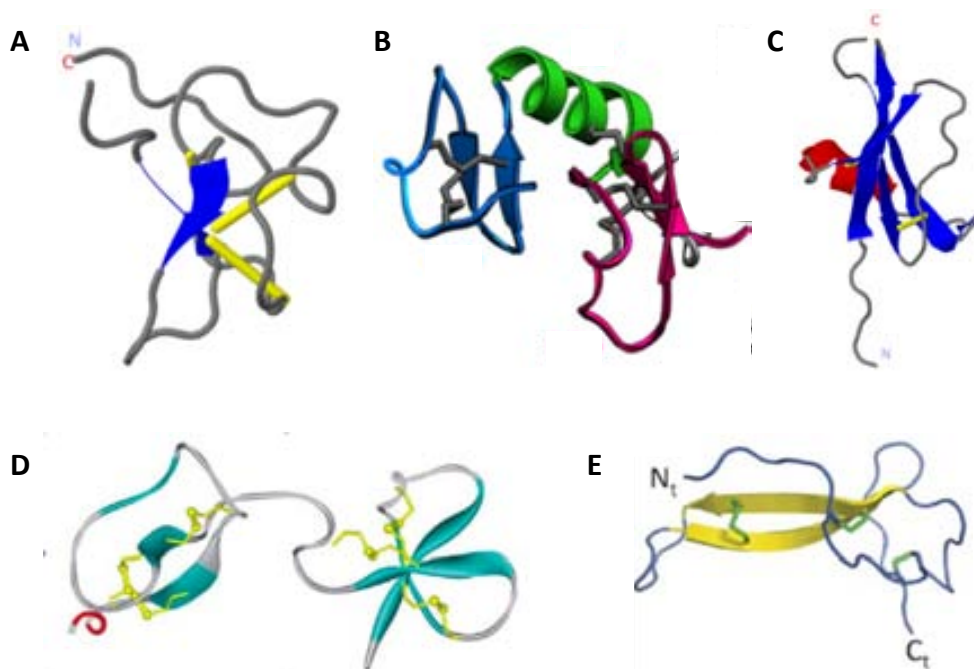
Figure 1.10. Structures of exogenous metallocarboxypeptidases inhibitors. A. PCI B. ACI C. LCI D. TCI and E. NvCI

The MCP inhibitor isolated from the parasite *Ascaris suum* (ACI) is a 67-residue amino acid sequence with no significant homology with any known protein. Heterologous overexpression and purification of ACI rendered a functional molecule with namolar Ki against M14A MCPs (Sanglas et al., 2009). The crystal structure of the complex of ACI with human CPA1 was determined and the inhibitor has a fold consisting of two tandem homologous domains, each containing a β-ribbon and two disulfide bonds (Figure 1.10B).

These domains are connected by an α-helical segment and a fifth disulfide bond (Sanglas et al., 2009).

The leech carboxypeptidase inhibitor (LCI), found in the medical leech *Hirudo medicinalis*, is a protein of 66 residues that behaves as a tight binding, competitive inhibitor of M14A MCPs with a nanomolar Ki. The LCI cDNA codes for an 81-residue sequence preceded by a 15-residue signal peptide. The 3D structure of its recombinant form was determined in solution by NMR and in crystal in complex with human CPA2 (Reverter et al., 2000). LCI folds in a compact domain consisting of a five-stranded anti-parallel β-sheet and a short α-helix that interacts with the end and the beginning of the first and second β-strands (Figure 1.10C) This inhibitor is highly stabilized by the presence of four disulfide bridges formed between cysteine residues located within secondary structure elements.

Inhibitor isolated from tick *Rhipicephalus bursa* (TCI) also competitively inhibits M14A MCP with namolar affinity (Arolas et al., 2005b). Its full-length cDNA codes for a precursor protein that contains a signal peptide of 22 residues followed by the mature TCI, a 75-residue cysteine-rich protein (six disulfide bridges). The sequence of the C-terminus resembles those of the other exogenous MCP inhibitors. The determined crystal structures of TCI bound to bCPA and to human CPB (Arolas et al., 2005b) have shown that it consists of two domains that are structurally similar despite the low degree of sequence homology (Figure 1.10D). The domains, each consisting of a short α-helix followed by a small twisted antiparallel β-sheet, show a high level of structural homology to proteins of the β-defensin-fold family.

The inhibitor isolated from the marine snail *Nerita versicolor* (NvCI) is the first proteinaceous inhibitor of MCPs isolated and characterized from a marine organism organism (Covaleda et al., 2012). NvCI is a 53-residue protein that tightly binds to the M14A MCPs with Ki in the picomolar range. Such Ki values are 3 orders of magnitude lower than for other known exogenous protein inhibitors. The crystal structure of NvCI in complex with human CPA4 have been resolved at 1.7 Å resolution (Covaleda et al., 2012). The structure of NvCI displays a new extended globular protein fold, which is basically formed by a central two-stranded antiparallel β-sheet connected by three major loops and two short tails that extend at the N- and C- termini (Figure 1.10E). The β-strands and the three loops are cross-connected and stabilized by three disulfide bridges.

The inhibition of exogenous CP inhibitors relies on the interaction of the C-terminal tail with the active site groove of the CP in a mechanism that mimics substrate binding. Although there are neither sequences nor three-dimensional structure similarities between the main bodies of these inhibitors, they all have effects on the key residues for the activity of the enzyme.

The C-terminal tail of PCI and LCI is inserted into the active-site cleft and forms the primary contact region (Figure 1.11A). The inhibitor affects most of the CP residues that are essential for substrate binding and catalysis (Rees and Lipscomb, 1982; Reverter et al., 2000). The interaction of PCI and LCI with M14A MCPs are also strongly stabilized by the presence of additional secondary contacts (Arolas et al., 2004).
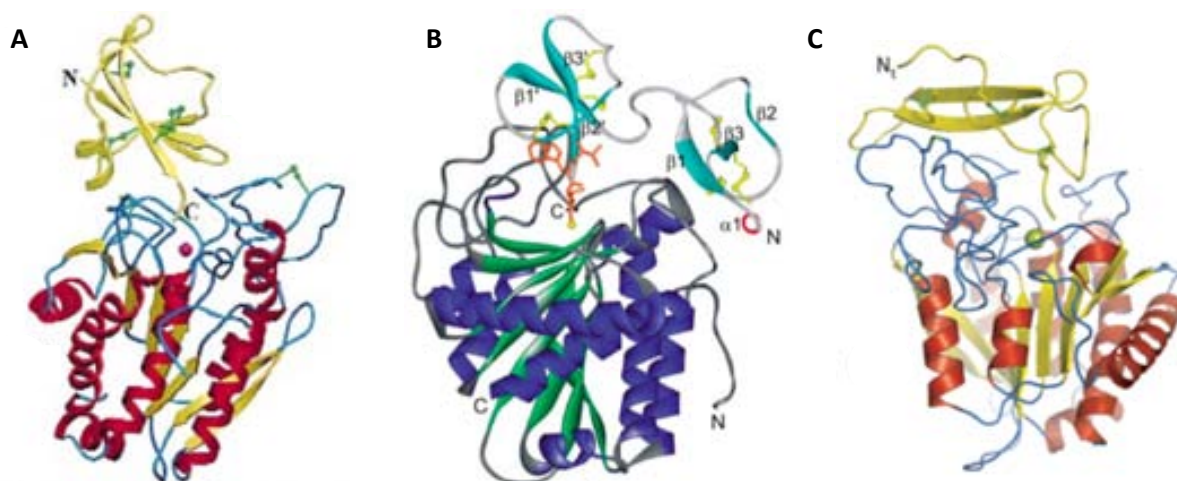
Figure 1.11. Crystal structures of complexes between exogenous inhibitors and M14A metallocarboxypeptidases. A. Cartoon representation of the structure of LCI in complex with human CPA2 (Reverter et al., 2000). For CPA2 α-helices and β-strands are colored in red and orange, respectively. LCI is colored in yellow and disulfide bridges are in green. B. Cartoon representation of the structure of TCI in complex with bovine CPA (Arolas et al., 2005a). The α-helix and β-strands in TCI are colored in red and light blue, respectively, and the disulfide bridges in yellow. The α-helices and β-strands of CPA are shown in dark blue and dark green, respectively. C. Cartoon representation of the structure of NvCI in complex with human CPA4 (Covaleda et al., 2012). The α-helices and β-strands of CPA4 are highlighted in red and yellow, respectively. The NvCI is colored in yellow and disulfide bridges are in green.

TCI anchors to the surface of M14A MCPs in a double-headed manner not seen for the other inhibitors (Figure 1.11B). Apart from several additional secondary contacts similar to those of PCI and LCI, the N-terminal domain of TCI establishes interactions with the enzyme surface away from the active site groove (Arolas et al., 2005b).

The lower Ki values observed for NvCI regarding other MCP inhibitors can be attributed to both the primary and secondary interaction regions (Figure 1.11C), which create an extended interface with the enzyme that minimizes the product release of the catalytic reaction (Covaleda et al., 2012) .

1.11 Pseudomonas aeruginosa

Pseudomonas aeruginosa is member of the gamma proteobacteria class of bacteria. It is a gram-negative, aerobic rod belonging to the bacterial family Pseudomonadaceae. This free-living bacterium has the ability to adapt to and thrive in many ecological niches, from water and soil to plant and animal tissues. Members of the genus are one of the few groups of bacteria that are true pathogens of plants. The bacterium is capable of utilizing a wide range of organic compounds as food sources, thus giving it an exceptional ability to colonize ecological niches where nutrients are limited (Ramos, 2004).

P. aeruginosa has become increasingly recognized as an emerging opportunistic pathogen of clinical relevance. Epidemiological studies track its occurrence as a nosocomial pathogen and indicate that antibiotic resistance is increasing in clinical isolates (de Bentzmann and Plesiat, 2011; Poole, 2011). It is the cause of urinary tract infections, respiratory system infections, dermatitis, soft tissue infections, bacteremia, bone and joint infections, gastrointestinal infections and a variety of systemic infections, particularly in patients with severe burns, cancer, patients suffering from cystic fibrosis and AIDS patients who are immunosuppressed (de Bentzmann and Plesiat, 2011).

The virulence machinery of *P. aeruginosa* comprises both cell-associated determinants (such as lipopolysaccharides, pili, flagella) and numerous secreted factors (such as elastases, proteases, exotoxins, pyocyanin, extracellular polysaccharides). One of the mechanisms by which *P. aeruginosa* senses external signals is using sensor proteins that, through phosphorelay, activate specific transcriptional regulators. These sensor–regulator protein pairs, called two-component systems, also feed into major regulatory pathways and play a critical role in allowing cells to modulate gene expression in response to environmental conditions (Gooderham and Hancock, 2009). Many of the secreted virulence factors and phenotypes, such as extracellular proteases, iron chelators, efflux pump expression, biofilm formation, motility and the response to host immune signals, are under the control of a cell

density recognition mechanism called quorum sensing (QS) that aids in the coordinated expression of genes. In QS, small molecules called autoinducers are produced and secreted and when reach a concentration threshold, they diffuse back into the cell to elicit a coordinated response promoting group survival (Cornforth et al., 2014). In *P. aeruginosa* this is achieved using two types of autoinducers, N-acyl-homoserine lactones and 2-alkyl-4 quinolones (Jimenez et al., 2012).

In the following sections, some critical virulence factors of *P. aeruginosa* will be described.

1.12 Type IV pili structure

Type IV pili provides bacteria the ability to adhere to chemically diverse surfaces, promotes cell aggregation involved in microcolony formation and virulence and mediates flagellum-independent motility. From an evolutionary standpoint, the type IV pili machinery is ancient, present in gram-positive*,* gram-negative, and archaeal species (Pelicic, 2008). In *P. aeruginosa* is possible to find three type IV pili systems: type IVa pili, type IVb pili and the tight adherence pili but only type IVa pili is associated with motility (hereafter T4P).

The T4P and its assembly apparatus can be envisioned as four interdependent subcomplexes that together span the entire cell envelope, with most components located in the inner membrane (Figure 1.12). The outer membrane subcomplex, consists of the secretin, a highly stable dodecamer of PilQ subunits, and its lipoprotein pilotin, PilF, composed entirely of protein-protein interaction motifs. PilF is responsible for correct localization and oligomerization of PilQ in the outer membrane. The secretin provides an outer membrane channel for the pilus fiber to exit the cell (Koo et al., 2013).

The motor and alignment subcomplexes are located in the inner membrane. In *P. aeruginosa,* the exact composition of the T4P motor subcomplex has not been defined but, on the basis of available studies and evidence from related systems, potentially comprises PilB, PilC, PilT, and PilU (Chiang et al., 2005), as well as regulatory proteins (Figure 1.12). PilB is a hexameric ATPase and is predicted to convert chemical energy from ATP hydrolysis to mechanical energy required for pilus assembly. PilT and PilU are PilB-like ATPases required for pilin depolymerization. PilD, a membrane-bound aspartyl protease that processes the type III signal sequence at the pilins' N-terminus, and N-methylates the new terminus

(Strom et al., 1993), might be transiently or permanently associated with the motor subcomplex.
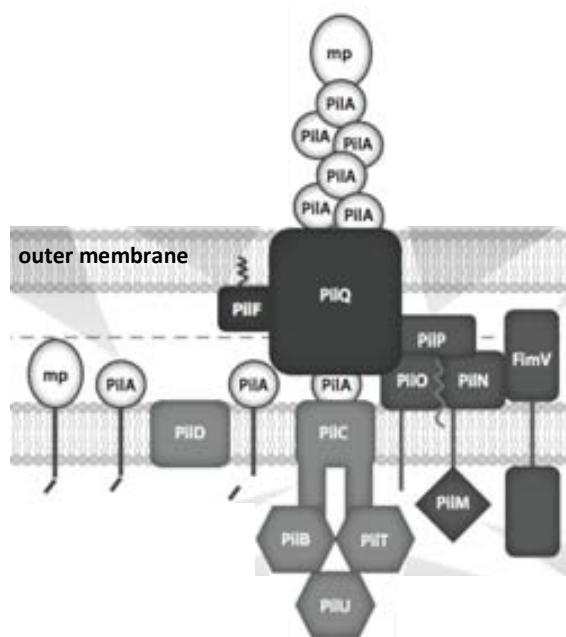
Figure 1.12. Type IVa pili (T4P) assembly subcomplex components. The outer membrane subcomplex (PilF and PilQ) is shown in dark gray; the inner membrane motor subcomplex (PilB, PilC, PilD, PilT, and PilU) in light gray; the alignment subcomplex (PilM, PilN, PilO, PilP and FimV) in medium gray; and the pilus subcomplex [the minor pilins (mp) FimU, PilV, PilW, PilX, and PilE and the major pilin, PilA] in white. This figure is adapted from Burrows (2012).

The alignment subcomplex is proposed to physically connect the inner membrane motor and outer membrane subcomplexes to ensure that the growing, highly flexible pilus is positioned correctly for egress as it begins to extend through the periplasm from its assembly site. The components of the alignment subcomplex include the broadly conserved PilMNOP proteins (Figure 1.12) that are encoded with the PilQ secretin monomer (Pelicic, 2008), and potentially the peptidoglycan-binding protein FimV (Wehbi et al., 2011).

PilM is a cytoplasmic, actin-like protein bound to PilN via the latter's cytoplasmic N-terminus (Karuppiah and Derrick, 2011). Both PilN and PilO proteins are predicted to have short cytoplasmic N-terminus, a single transmembrane domain, and a periplasmic C-terminus. PilP is an inner membrane lipoprotein that forms a stable 1:1:1 heterotrimer with PilN and PilO (Tammam et al., 2011). The last member of the proposed alignment subcomplex is FimV, a protein containing a periplasmic domain with a peptidoglycan-binding LysM motif connected via a single transmembrane segment to a highly acidic cytoplasmic domain. The peptidoglycan-binding function of FimV is required for correct

formation of the PilQ secretin (Wehbi et al., 2011). The function of the cytoplasmic domain of FimV is currently unknown, but it may participate in regulating pilus biogenesis (Fulcher et al., 2010).

The final subcomplex is the pilus itself. It is the most dynamic of the four, as it is repeatedly assembled and disassembled. It is composed of the major pilin subunit (PilA) (Paranchych et al., 1979), as well as the minor (based on abundance) pilins FimU, PilVWX, and PilE. Although the majority of the pilus fiber consists of PilA, the minor components are also present in surface-exposed pili (Burrows, 2012). Both major and minor pilins share the highly conserved N-terminal domain but their periplasmic C-termini are divergent. The hydrophobic N-terminus of pilins act as transmembrane domains prior to assembly and as protein-protein interaction domains in the assembled fiber (Craig et al., 2006).

The extended N-terminal α-helix of pilins proteins (Figure 1.13A) has two sudomains. The first (spanning amino acids ~1 to 28) forms the central core of the assembled pilus fiber (Craig et al., 2004)and acts as a transmembrane segment to retain individual pilin subunits in the cytoplasmic membrane prior to assembly, as a protein interaction domain for subunit-subunit interactions in the fiber, and potentially as a regulatory domain. The second subdomain (amino acids ~29 to 52), embedded in the C-terminal globular domain, is amphipathic and packs against the head domain. There is a curvature in the N-terminal helix that is thought to reduce intersubunit packing, and may also contribute to fiber assembly-disassembly dynamics by making transmembrane domains more flexible (Giltner et al., 2012).

The hypervariable C-terminus of pilins have a conserved structural element known as the disulfide-bonded loop (D-region). Two Cys residues near the C-terminus form a disulfide bond that staples the C-terminus of the protein to the β-sheet. The D region has a structural role, as mutations that lead to a loss of C-terminal disulfide bond formation prevent pilus assembly (Giltner et al., 2012).

Crystallographic and cryo-electron microscopy evidence suggests that the pilin monomers are arranged in a 3-start left-handed helix with a predicted 3.6 subunits per turn (Figure 1.13B). The fiber has an outer diameter of ~60 Å, consistent with the ~65 Å-diameter opening of the T4aP secretin through which the fiber passes to the cell's exterior. The fiber is stabilized by hydrophobic and electrostatic subunit-subunit interactions

between the N-terminal α-helices, which form the central core of the pilus (Craig et al., 2006).
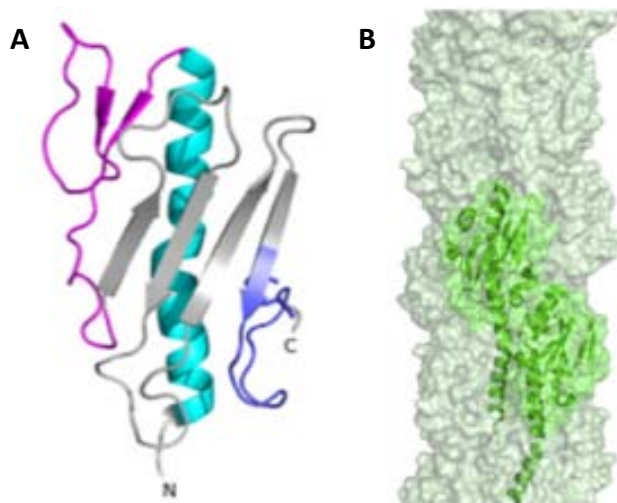
Figure 1.13. Structure of type IVa pilin of *Pseudomonas aeruginosa* PAK and model of type IVa pili. A. Structure of *P. aeruginosa* PAK PilA (PDB accession no. 1DZO). N-terminal α-helice (cyan) connected to a β-sheet (gray) by the αβ-loop (magenta) and D region (blue). B. Fiber models of the T4aP can be viewed as a one-start or four-start right-handed helix or a three-start left-handed helix (one strand is pictured, with two subunits shown as a green cartoon), with a diameter of approximately 60 Å. this figure is adapted from Giltner et al., (2012)

The N-terminal α-helices form staggered helical bundles that spiral along the length of the pilus. The C-terminal domains form subunit-subunit contacts, mainly in loop regions. The outer surface of the pilus, formed by the C-terminal domains of the subunits, is characterized by deep grooves separating the head groups Combined, these features allow for a highly flexible structure that can withstand the piconewton-scale forces required for twitching motility (Craig et al., 2006).

1.13 Regulation of pilin expression

In *P. aeruginosa*, the major pilin subunit is expressed from a σ^{54}-dependent promoter (Hobbs et al., 1993). The complex formed by RpoN and RNA polymerase binds the promoter but remains stalled in a closed configuration. The conversion of the complex to a transcription-competent open form requires PilR, the response regulator of the two-component PilR-PilS system.

In *P. aeruginosa,* PilS is a membrane-bound histidine kinase that is autophosphorylated in response to an as yet unknown signal and, in turn, phosphorylates PilR, promoting pilin expression. PilS levels are critical to the control of PilA levels, as both lack of PilS and its

overexpression had a detrimental effect on pilin gene transcription. Boyd and Lory (1996) suggested that PilS over-expression titrated a crucial interaction partner, leading to decreased pilin gene transcription.

Pilin mutants of *P. aeruginosa* had increased activity from a *pilA* transcriptional reporter (Bertrand et al., 2010), suggesting that PilA negatively regulates its own expression, possibly through direct interactions with PilS. In mutants lacking the retraction ATPase PilT, the reporter was similarly activated. In a *pilT* background pilin levels in the cytoplasmic membrane would presumably be low due to the inability of the cells to retract their pili. In contrast, *P. aeruginosa* mutants lacking the assembly ATPase PilB had low levels of reporter activity (Bertrand et al., 2010). Such mutants accumulated pilins in the inner membrane due to the blockade in pilus assembly. The phenotypes of assembly and retraction mutants are consistent with the hypothesis that high levels of PilA in the inner membrane suppress *pilA* transcription.

The sensor kinase PilS, which has an atypical 6 transmembrane domains but no periplasmic sensor domain, could monitor levels of pilin through direct intramembrane interactions with the conserved hydrophobic N-terminus of PilA. High PilA levels could drive PilA-PilS interactions that, depending on the particular species, could either promote or prevent kinase activity. As pilin levels decrease, PilS might adopt an alternate conformation and act as a PilR phosphatase.

1.13.1. Regulation through pilin stability

Pilins are quite stable, and protein levels remain elevated even after cells are transferred to growth conditions that do not promote *pilA* expression. A reduced rate of pilin turnover is consistent with an energy conservation strategy in which pilins are retained in the inner membrane under conditions where twitching motility is not optimal but are ready for deployment if conditions become favorable. The amount of motility was related to the number of surface pili rather than to total PilA levels. In *P. aeruginosa* using an arabinose-inducible promoter it was seen that without inducer, the low level of PilA synthesized was insufficient for motility. When 0.1% arabinose or more was used, motility was restored to wild-type levels (Harvey et al., 2009).

1.14 Flagellum structure

Flagellum is primarily a motility organelle that enables movement and chemotaxis although, depending on bacterial species and life cycle, can also participate in biofilm formation, protein export, and adhesion. The bacterial flagellum is a rotary structure driven from a motor at the base, with the filament acting as a propeller. Over 60 structural and regulatory proteins are required for flagellum assembly and function. Flagellum consists of a cytoplasmic export apparatus, a basal body embedded in the cell membrane, a hook that connects the basal body to the filament, and a filament that functions as a propeller (Figure 1.14) (Chevance and Hughes, 2008).
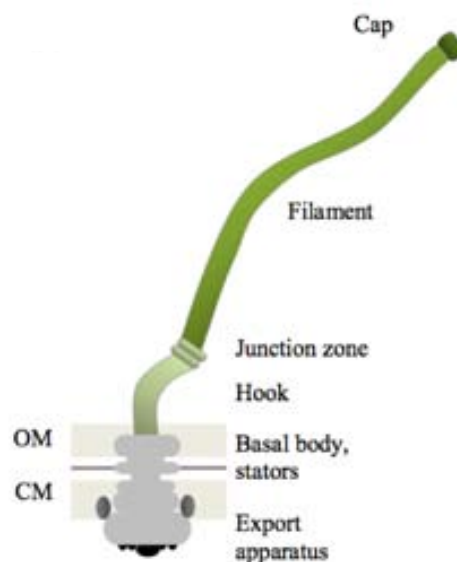
Figure 1.14. Schematic presentation of bacterial flagellum structure. The cap, the filament of flagellin subunits, the junction zone and the hook represent extracellular parts (green shades) of the flagellum. The basal body (grey) in the cell wall consists of a centrally located hollow rod embedded in the outer membrane (OM), the peptidoglycan layer, and the cytoplasmic membrane (CM). Stator complexes (dark grey), composed of membrane proteins MotA and MotB provide motility-required energy. The cytoplasmic export machinery (black) that secretes the extracellular subunits is located within the cytoplasmic ring. This figure is adapted from Haiko and Westerlund-Wikstrom (2013).

Flagellar assembly starts with the cell membrane-associated components of the basal body and the secretion apparatus, through which the other flagellar proteins are then secreted, first the remaining basal body components, then the hook and the hook-filament junction proteins. Filament assembly starts after the hook has been completed and the filament capping proteins have been positioned. The filament is composed of about 20,000 flagellin (FliC) proteins that are incorporated below the distal pentameric FliD cap, which

functions as a plug and is required for assembly of nascent monomeric flagellin (Chevance and Hughes, 2008).

FliC from *P. aeruginosa* strain consists of three domains (D0, D1, and D2). The crystal structure of FliC containing D1 and D2 domains (Song and Yoon, 2014) showed that D1 domain is folded into a rod-shaped structure while D2 domain exhibits a unique structure of two β-sheets and one α-helix that has not been found in other flagellins (Figure 1.15A). An *in silico* construction of a flagellar filament based on the packing of FliC in the crystal (Figure 1.15B) suggests that D2 domain would be exposed to solution and could play an important role in immunogenicity (Song and Yoon, 2014).
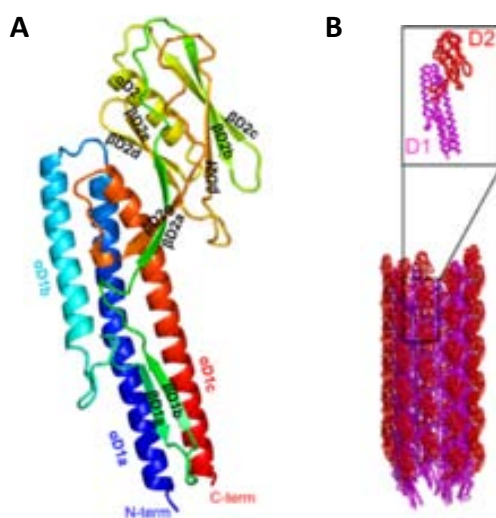
Figure 1.15. Structure of D1 and D2 domains of *Pseudomonas aeruginosa* flagellin and a model of flagellar filament. A. Flagellin structure in rainbow ribbons from the N-terminus in blue to the C-terminus in red. Secondary structures are labeled. B. Flagellin monomer structure (top) D1 and D2 domains are colored in magenta and red, respectively. A model of the filament (bottom). This figure is adapted from Song and Yoon (2014).

A two-component signaling cascade involving chemotaxis-related proteins affects flagellar rotation, which is facilitated by the engine consisting of the basal body-associated stator proteins and the basal body, which functions as a rotor (Sourjik and Wingreen, 2012).

1.15 Motilities in *Pseudomonas aeruginosa*

Bacterial motility is a crucial component of fitness in most habitats enabling cells to move toward sources of nutrition (Wei et al., 2011), move away from potential causes of harm such as toxins and predation (Chet and Mitchell, 1976) and avoid competition with clone mates (Taylor and Buckling, 2010). It is also particularly important in host colonization

success of opportunistic pathogens (Haiko and Westerlund-Wikstrom, 2013), and indirectly influences biofilm development (Harmsen et al., 2010) and bacteriophage attachment (Samuel et al., 1999). The cost of motility is significant considering the metabolic burden of synthesizing and assembling various flagellar and pili components and the energetic expense of fueling flagellar and presumably pili motors. Thus, the synthesis and control of movement through the chemotactic signal transduction system of the motility apparatus are subject to strict control mechanisms that are usually redundant and multilayered (Jimenez et al., 2012).

P. aeruginosa expresses two surface organelles that aid motility: a single polar flagellum and a retractable polar T4P. In addition, it is also able to secrete a lipid-based biosurfactant, rhamnolipids (Kohler et al., 2000). This enables the bacterium to adopt pilus-based twitching on hard surfaces (Burrows, 2012), flagella-mediated swimming in aqueous environments (Bai et al., 2007), and a coupled pili-flagella action combined with rhamnolipid production known as swarming on semi-soft surfaces (Kohler et al., 2000).

1.15.1. Swimming motility

Swimming on a surface takes place when the fluid film is sufficiently thick and the micromorphological pattern is unorganized. Monotrichous flagellated bacteria such as *P. aeruginosa*, with a single flagellum at the pole of the cell body, swims in run and reverse patterns. A reversal of the flagellar rotation direction from counterclockwise to clockwise causes a change in direction from forward to backward swimming (Vater et al., 2014).

1.15.2. Swarming motility

Swarming motility is a multicellular phenomenon involving the coordinated and rapid movement of a bacterial population across a semisolid surface (Fraser and Hughes, 1999). During swarming, either the number of flagella at the pole increases and/or the single polar flagellum is augmented by an alternative swarming-specific stator complex (Toutain et al., 2005). *P. aeruginosa* swarming depends on amino acids as a nitrogen source and the secretion of rhamnolipids and 3-(3-hydroxyalkanoyloxy)alkanoic acids as wetting agents and chemotactic-like stimuli (Deziel et al., 2003; Kohler et al., 2000). Under specific nutrient conditions, swarming motility requires T4P, but provision of specific carbon sources including glutamate, glucose, or succinate restores swarming of pilin mutants. It is proposed

that T4P could assist the flagellum in surface propagation or, alternatively, the pili could be involved in sensing the viscosity of the surface and sending a signal for initiation of swarming (Deziel et al., 2003; Kohler et al., 2000).

P. aeruginosa gene expression during swarming is substantially different from that observed during growth under lower viscosities (swimming/ planktonic conditions). A large number of virulence-related genes are up-regulated such as those encoding the type III secretion system and its effectors, those encoding extracellular proteases, and those associated with iron transport. In addition, swarming cells exhibit adaptive antibiotic resistance against polymyxin B, gentamicin, and ciprofloxacin (Overhage et al., 2008).

Studies in mutant strains have shown that swarming in *P. aeruginosa* could be regulated by the *las* and *rhl* cell-to-cell systems and by the pilus-related chemotaxis system (Deziel et al., 2003; Kohler et al., 2000). Additionally, proteins PvdQ and elastase LasB play important roles in swarming behavior. PvdQ degrades the QS signal 3-oxododecanoyl-homoserine lactone and a specific concentration of this molecule seems to be important for swarming motility (Overhage et al., 2008). The mechanisms of regulation mediated by Las B are still unknown.

1.15.3. Twitching motility

Twitching motility results from the repeated extension, tethering, and retraction of T4P (Bradley, 1980). Twitching occurs on moist surfaces of moderate viscosity, equivalent to that of 1% agar. Although individual cells are capable of movement, it is common to see them moving in rafts, groups of cells preferentially aligned along their long axes (Semmler et al., 1999). The exact distance traveled by twitching cells depends on the medium composition, viscosity, and hydrophobicity of the surface, and intrinsic factors like the amount of pili produced and their retraction rates, and the production of surface-tension-reducing surfactants (Burrows, 2012). Even though pili can be retracted individually, the cooperative retraction of several bundled fibers generates more substantial forces and allows cells to travel distances that exceed the length of individual pili (Burrows, 2012).

Twitching motility has been considered an important virulence factor in *P. aeruginosa* in different *in vivo* models. T4P are deployed in the early or acute phase of infection but are frequently lost owing to downregulation or mutation in chronic infections such as cystic fibrosis (Mahenthiralingam et al., 1994). Pilus retraction was important for cytotoxicity

against epithelial cells and for virulence *in vivo* (Comolli et al., 1999). Zolfaghar et al., (2003) showed that mutants lacking PilT or PilU were impaired in corneal infection models. They subsequently demonstrated *in vitro* that twitching motility mutants were impaired in translocation across the epithelium and in escape from infected cells (Alarcon et al., 2009)

In *P. aeruginosa,* twitching is controlled by a number of regulatory systems that sense external signals, most of which are unknown, and transduce them to modulate pilus extension and retraction. Some of these systems included QS that indirectly affects twitching through effects on Vfr, a homolog of *E. coli* cyclic AMP receptor protein (Beatson et al., 2002). The RhlI-RhlR QS system and iron levels regulate the expression of RhlAB, enzymes involved in production of the biosurfactant rhamnolipid (Glick et al., 2010) and the Pil-Chp chemotaxis system has cyclic AMP-dependent and cyclic AMP-independent effects on T4P function (Fulcher et al., 2010). The protein Ppk (polyphosphate kinase) is required for full twitching motility (Rashid and Kornberg, 2000).

1.16 Biofilm formation in *Pseudomonas aeruginosa*

Microbial biofilms have been subject to intense study for the scientific interest to understand how bacteria form and live in multicellular communities. Additionally, biofilms causes considerable problems in medical and industrial settings, because bacteria in biofilms can resist antibiotic treatment, host immune responses, and biocide treatment. The mechanisms that contribute to tolerance in *P. aeruginosa* biofilms include restricted antimicrobial diffusion, differential physiological activity, and induction of specific tolerance mechanisms.

The initial step of biofilm formation is the attachment of *P. aeruginosa* to surfaces. This process is influenced by a variety of components including flagella, T4P (O'Toole and Kolter, 1998), extracellular DNA (Whitchurch et al., 2002), production of polysaccharides (Ma et al., 2006) and rhamnolipids (Glick et al., 2010)and Cup fimbria (Vallet et al., 2001).

Surface-associated motility is an integrated part of *P. aeruginosa* biofilm formation. The formation of the "mushroom" caps in *P. aeruginosa* biofilms requires both T4P and flagella (Barken et al., 2008; Klausen et al., 2003). Cap formation requires flagellum-driven motility, whereas the dependence of cap formation on T4P may be due to the binding of these pili to extracellular DNA that is abundant on the microcolonies that become colonized during

mushroom-structure formation (Barken et al., 2008; Whitchurch et al., 2002). Alternatively, a kind of migration that requires both flagella and T4P might be occurring in the later phase of *P. aeruginosa* biofilm formation (Kohler et al., 2000).

The extracellular polymeric substance (EPS) matrix serves as the "house for biofilm cells". The *P. aeruginosa* EPS matrix contains polysaccharides including the glucose-rich Pel polysaccharide, the mannose-rich Psl polysaccharide and (Friedman and Kolter, 2004; Whitchurch et al., 2002). The Psl and Pel polysaccharides mediate cell-to-surface and cell-to-cell interactions, essential for *P. aeruginosa* biofilm formation and maintenance (Ma et al., 2006). Alginate is an acetylated polysaccharide, which has been shown to act as a physical barrier for antibiotics (Hatch and Schiller, 1998). Another major component of the *P. aeruginosa* EPS matrix is extracellular DNA that is similar to chromosomal DNA and appears to have a biofilm-stabilizing effect (Allesen-Holm et al., 2006).

The extent of dispersion of *P. aeruginosa* biofilms is dependent on the carbon source and is associated with increased expression of flagella and downregulation of twitching motility (Sauer et al., 2004). It appears that in general, high intracellular levels of the second messenger cyclic diguanosine-5´-monophosphate (c-di-GMP) upregulate matrix production and biofilm formation, whereas low intracellular c-di-GMP levels downregulate matrix production and induce a planktonic lifestyle. Carbon starvation and nitric oxide signaling were shown to induce the dispersal of *P. aeruginosa* biofilms via the induction of phosphodiesterase activity, causing decreased intracellular c-di-GMP levels (Barraud et al., 2009). QS also impacts several components of *P. aeruginosa* biofilm formation including the production of rhamnolipid, extracellular DNA and Pel polysaccharide and the expression of Cup fimbria proteins (Harmsen et al., 2010).

1.17 Proteases in *Pseudomonas aeruginosa*

P. aeruginosa contains about 5568 predicted genes, 155 genes (2.8% of the genome) of which are predicted proteases and deposited in the MEROPS database (Rawlings et al., 2014). Several of these proteases, including elastase B, elastase A, protease IV and alkaline protease, are secreted to the extracellular medium and play a role in the pathogenic interaction between bacterium and host (Hoge, 2010).

On the other hand, intracellular proteases play essential roles in orchestrating cellular activities via regulation of the levels of chaperones, labile regulators, and stress-related proteins, as well as by degrading misfolded proteins. In *P. aeruginosa*, the most studied are the ATP-dependent proteases that act like unconventional regulators presumably by modulating the stability of more conventional transcriptional regulators (Fernandez et al., 2012). ATP-dependent proteases like Lon belong to the family of AAA+ proteases. Lon is an 87-kDa cytoplasmic serine protease that associates into hexameric rings and exhibits a Ser679-Lys722 dyad, which is responsible for its catalytic activity. This protease contains an N-terminal domain for substrate recognition, an ATP-binding domain, and a proteolytically active C-terminal domain for cleavage. There is only restricted access to the active site, and only unfolded substrates can enter for cleavage (Rotanova et al., 2006).

There is an increasing body of evidence that these intracellular proteases play crucial roles in the coordination of antibiotic resistance, motility, biofilm formation, virulence, and stress adaptation in this microorganism besides their known role in protein quality control (Sauer and Baker, 2011). Mutations in the genes encoding the Lon and ClpP proteases have been identified to exhibit increased susceptibility to ciprofloxacin (Fernandez et al., 2012). Furthermore, these mutants show defects in motility (swarming, swimming, twitching) and biofilm formation (Fernandez et al., 2012). Apart from the virulence-related defects, *lon* mutants were consistently less virulent in several in vivo models (Breidenstein et al., 2012b). It is hypothesized that the Lon protease regulates and coordinates resistance to ciprofloxacin and key virulence determinants of *P. aeruginosa* by modulating the stability of key transcriptional regulators (Breidenstein et al., 2012a). AsrA (aminoglycoside-induced stress response ATP-dependent protease) has 60% similarity to the Lon protease and mediates the adaptive response to the aminoglycoside tobramycin by controlling heat shock responses (Kindrachuk et al., 2011).

1.18 Caenorhabditis elegans model to study Pseudomonas aeruginosa virulence

The establishment of a nematode, *C. elegans*, as a model for genetic study in 1960s has brought the possibility to study this animal in different biological arenas including that of pathogenic bacteria. The worm is genetically and physiologically well-understood.

Fascinatingly, this worm shares similarities to mammalian signaling cascades in innate immunity in response to pathogen invasion (Irazoqui et al., 2010). This fact, along with the susceptibility to human pathogens, has led to the widespread employment of *C. elegans* as an infection model by simply feeding the worm with the pathogen of interest.

Extensive studies on the effect of *P. aeruginosa* on *C. elegans* have resulted in four different infection models that each apply slightly different media and incubation conditions. In each of these models the lethality is depending on a different subset of virulence factors, which are also important to mammals, as summarized in Table 1.1.

Table 1.1 Resemblance of P. aeruginosa pathogenesis in C. elegans with infection in mammals. This table is adapted from Utari and Quax (2013).

Assay	Comparison with infection in mammals	References
Slow killing (agar-based)	As the killing mechanism based on infection-like process involves establishment and proliferation of pathogen in the intestine along with biofilm formation, study of inhibition of bacterial pathogenesis such as quorum quenching can be performed.	(Papaioannou et al., 2009)
Fast killing (agar-based)	Toxicity of phenazine-1-carboxylic acid under acidic conditions, suggests this toxin plays role in infection of cystic fibrosis patients or during the inflammation stage of acute wounds where the acidity increases.	(Cezairliyan et al., 2013)
Lethal paralysis (agar-based)	Cyanogenesis of *P. aeruginosa* during lung infection in cystic fibrosis patient is a complex network involving quorum sensing, GacA and RsmA regulation. Cyanogenesis is also suggested as a mechanism of *P. aeruginosa* dominance and persistence in the site of infection.	(Anderson et al., 2010)
Red death (agar-based)	This assay mimics *P. aeruginosa* infection during intestinal phosphate depletion after major surgery in mammals.	(Zaborin et al., 2009)
Liquid killing assay	This assay possibly mimics the conditions in cystic fibrosis patients where an oxygen gradient is observed in the thick mucus, inducing the hypoxic response which leads to deleterious epithelial airway inflammation. In addition, this assay tests for several virulence factors stimulated by pyoverdin, such as exotoxin ToxA and PrpL protease.	(Legendre et al., 2011)

In the "slow killing" assay, which takes place in minimal media (nematode growth media, NGM), intestinal colonization and QS are essential for the killing mechanism (Papaioannou et al., 2009). In high-osmolarity peptone glucose sorbitol media ("fast killing"), the worms are killed within several hours by intoxication by phenazine (Cezairliyan et al., 2013). Upon growth in Brain Heart Infusion media, *P. aeruginosa* produces hydrogen cyanide, a neurotoxin which causes lethal paralysis of the worm within hours (Gallagher and Manoil, 2001). In the assay called "red death", the lethality is induced by utilization of phosphate-depleted media. The global virulence systems involved in this response included

phosphate signaling (PhoB), the MvfR-PQS pathway of QS, and the pyoverdin iron acquisition system. Activation of all 3 systems was required to form a red colored PQS+Fe(3+) complex which conferred a lethal phenotype in this model (Zaborin et al., 2009).

In addition to the four agar-based assays mentioned above, recently a novel high-throughput liquid-based pathogenesis assay has been developed by the pioneers of the *C. elegans* infection model (Kirienko et al., 2013). Interestingly, in this liquid killing assay the virulence factors considered responsible for killing in classic assays, mostly are not lethally involved. In liquid killing the pyoverdin siderophore plays an important role in inducing a hypoxic response followed by death of the worms (Kirienko et al., 2013).

Discrete responses of *C. elegans* in reaction to *P. aeruginosa* in different pathogenesis assays demonstrate its excellent feature as a model to study host–pathogen interactions. As each assay probes a different subset of virulence factors, they can complement each other to unravel and better understand the exact virulence mechanisms. In addition, a combination of *C. elegans* assays can be used to study pathogenesis inhibition by novel antibacterial therapies urgently needed to combat antibiotic resistance. Non-conventional drug targets, such as the QS pathway, can now be studied by observation of pathogen attenuation under specific assay conditions with the read out based on the survival of the worms (Papaioannou et al., 2009). The high-throughput nature of the *C. elegans* assays render these into an invaluable preliminary method to screen compound libraries in relatively short time. The selected compounds can subsequently be tested in a higher animal model, such as rodent, to provide better pharmacological data in a complex mammalian system.

Objectives

The general objective of this thesis is to gain more insight into the structure, enzymatic properties and function of bacterial and human metallocarboxypeptidases of M14D subfamily. This is a novel type of proteases and we have selected two variants of them, quite apart phylogenetically to achieve such goal.

Specific objectives:

Chapter 1

- To express the cytosolic carboxypeptidase from *Pseudomonas aeruginosa* PAO1 (PaCCP) in *E.coli* and to purify the recombinant product.
- To derive and describe the three-dimensional structure of PaCCP.
- To evaluate the enzymatic activity and mechanism of PaCCP.
- To analize the interaction of PaCCP with M14A metallocarboxypeptidase inhibitors.
- To model the amino and carboxy domains of human cytosolic carboxypeptidases using the three-dimensional structure of PaCCP.

Chapter 2

- To unveiled and identify the functional role and mechanisms of PaCCP through the characterization of microbiological phenotypes of a *Pseudomonas aeruginosa* PAO1 strain deficient in PaCCP ($\Delta paccp$).
- To identify potential interacting partners of PaCCP.

Chapter 3

- To express the human cytosolic carboxypeptidase 6 (hCCP6) and a catalytic inactive mutant in mammalian cells and to partially purify both recombinant products.
- To analize the enzymatic activity of hCCP6 toward protein substrates.
- To determine the susceptibility of hCCP6 to M14A metallocarboxypeptidase inhibitors.
- To identify potential interacting partners of hCCP6.

2. Experimental Section

2.1. Materials and Methods used in Chapter 1

2.1.1. Sequence alignments

The amino terminal and carboxypeptidase domains (N- and CP- domains, respectively) of M14D peptidases *Pseudomonas aeruginosa* CCP (UniProt ID: Q9I012), *Burkholderia mallei* CCP (UniProt ID: Q62IR3), Shewanella denitrificans CCP (UniProt ID: Q12MY8) and humans CCP1 (UniProt ID: Q9UPW5), CCP2 (UniProt ID: Q5U5Z8), CCP3 (UniProt ID: Q8NEM8), CCP4 (UniProt ID: Q96MI9), CCP5 (UniProt ID: Q8NDL9) and CCP6 (UniProt ID: Q5VU57) were aligned using program T-coffee (Notredame et al., 2000). The alignment was manually corrected based on a) the PsiPred (McGuffin et al., 2000) secondary structure prediction of M14D peptidases, b) the conservation of zinc ligands and catalytic residues of M14 peptidases (Rodriguez de la Vega et al., 2007), and c) the experimentally solved structures of bacterial M14D and human M14A and M14B peptidases.

2.1.2. Cloning and DNA sequencing

The PA2831 gene, encoding PaCCP, was amplified by PCR from the genomic DNA of *P. aeruginosa* MPAO1 strain purchased from the Transposon Mutant Collection, University of Washington, USA (Jacobs et al., 2003). PCR was performed using FastStart TaqDNA Polymerase (Roche Diagnostics, GmbH) with the following primers: PaCCPfw 5´-AGT CTT CGA AAA GGG TGC AAT GCA GAT CCG CGC CGA TTT-3´ and PaCCPrv 5´-GTA CGA GCT CTC AGC GGA GTT CGT CAA CCA-3´. The amplified product and vector pTriEx6 (Novagen) were digested with restriction enzymes BstBI and SacI (Thermo Scientific). The digested fragments were purified from agarose gel 1% (w/v) with the GeneJET Gel Extraction Kit (Thermo Scientific) and ligated with T4 ligase (New England Biolabs) following the manufacturer's instructions. Cloning design allows the expression of N-terminal *Strep*-tag II (Schmidt and Skerra, 2007) fusion protein. Ligation mix was transformed into chemically competent *Escherichia coli* (*E. coli*) XL1-Blue cells and positive clones were selected in Luria Bertani (LB) agar plates containing 50 µg/ml ampicillin. The PaCCP Y64A mutant was generated by the primer extension method first described by Ho et al., (1989). Primers to amplify the two

fragments of the gene at both sides of the mutation point were the following, for fragment N terminal: Y64A3fw 5´-GCA GTT CGA AAA GGG TGC AAT GCA GAT CC-3´and Y64A3fusionrv 5´-CAT GGC TGG CGG CCG ACT GGC CGG CGT TG-3´ and for fragment C-terminal: Y64A3fusionfw 5´-GTC GGC CGC CAG CCA TGC CTG GAG CGG CTA C-3´ and PaCCPrv described above. PCRs were performed with the Phusion High-Fidelity DNA polymerase (Thermo Scientific) using the *paccp* wild-type (wt) gene cloned into vector pTriEx6 (*paccp*-pTriEx6) as template, following the manufacturer´s instructions. The final PCR product was cloned into vector pTriEx6 between BstBI and SacI restriction sites as described above. The PaCCP C320R mutant was generated by the Exsite method (Hemsley et al., 1989) with the following primers: C320Rfw 5'-CGC AAC TTC GTC GGG CAG AC-3' and C320Rrv 5'-GGC CAG CGC CAG GTT GG-3'. PCR was performed with the Phusion High-Fidelity DNA polymerase (Thermo Scientific) using *paccp*-pTriEx6 construct as template. The resultant plasmid was ligated and transformed into *E. coli* XL1-Blue competent cells as described above.

In each case, ten colonies resulting from transformation in *E. coli* XLI-Blue competent cells were selected for plasmid purification with GeneJET Plasmid Miniprep Kit (Thermo Scientific) and sequenced bidirectionally using T7 forward and TriExdown reverse primers in the Genomic Service, Institute of Biotechnology and Biomedicine (IBB), UAB. All primers were synthesized by Invitrogen, Life Technologies.

2.1.3. Polyacrylamide gels electrophoresis

Proteins were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) following the classical method (Laemmli, 1970). Samples were resuspended in reducing sample buffer 5X (0.25 M Tris/HCl, pH 8.0; 25% glycerol (v/v); 7.5% SDS (w/v), 0.25 mg/ml bromophenol blue; 12.5% mercaptoethanol (v/v)) and boiled for 5 min prior to electrophoresis. The resulting protein samples were resolved on a 10% SDS-PAGE minigel in a Mini-protean Tetra Electrophoresis system (BioRad). The reagents: Acrylamide / Bisacrylamide 37.5:1.0 (BioRad: 1610148), SDS, glycine, ammonium persulfate and N,N,N',N'-Tetramethylethylenediamine (TEMED) were purchased from Sigma y GE Healthcare. The buffer 125 mM Tris-HCl, pH 6.8 was used for stacking gel; 375 mM Tris-HCl, pH 8.8 for separating gel and 25 mM Tris supplemented with 192 mM glycine pH 8.3 (Tris-Gly pH 8.3) for running buffer. All buffers contained 0,1% SDS (w/v). EZ-Run™

Prestained Rec Protein Ladder (Fisher) was used as protein marker and gels were visualized using Coomassie blue dye.

Non-denaturing PAGE followed the protocol previously described except that SDS was omitted from all buffers and samples were not boiled nor treated with β-mercapthoethanol. Sample buffer 6X contained 0.25 M Tris/HCl, pH 6.8; 25% glycerol (v/v) and 0.25 mg/ml bromophenol blue. The protein samples were resolved on a minigel in a Mini-protean Tetra Electrophoresis system (BioRad) using NativeMark™ (Invitrogen) and visualized using Coomassie blue dye.

2.1.4. Protein production and purification

The production and purification of PaCCP wt, PaCCP Y64A and PaCCP C320R mutants followed the same procedure. Expression plasmids were transformed into competent *E. coli* BL21 (DE3) cells following the current protocol (Maniatis, 1982). Transformed cells were grown at 37 °C, with shaking at 250 rpm, in LB growth medium supplemented with 50 μg/ml ampicillin to an optical density (O.D.) of 0.6 at λ = 600 nm. Protein expression was induced by addition of 0.5 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) to the culture medium, followed by incubation at 37 °C for 16 h. Cells were harvested by centrifugation at 10,000g for 20 min, resuspended in 0.1 M Tris-HCl, pH 8.0 supplemented with 0.15 M NaCl (buffer W) and 1 mM phenylmethanesulfonyl fluoride (PMFS, Sigma) and disrupted by sonication. Insoluble material was removed by centrifugation at 20,000g for 20 min. The cell lysate was loaded onto a Strep-Tactin Macro-Prep column (IBA GmbH) connected to an Äkta Purifier system (GE Healthcare). The column was washed with 5 column volumes of buffer W, and the protein was eluted from the resin with buffer W supplemented with 2.5 mM desthiobiotin (IBA GmbH). Fractions contained the PaCCP protein were pooled and loaded onto a HiLoad Superdex 75 26/60 column (GE Healthcare) previously equilibrated with 20 mM Tris-HCl, pH 8.0 supplemented with 50 mM NaCl. Eluted fractions were analyzed by SDS-PAGE, and the purest samples containing the protein were pooled.

The oligomerization state of PaCCP was further analyzed by gel filtration in a Superdex 200 10/300 GL column (Tricorn) with the following buffers 0.02 M Tris-HCl pH 8.0 supplemented with 0.5 M NaCl; 0.02 M Tris-HCl pH 6.5 supplemented with 0.15 M NaCl and 0.02 M MES pH 5.5 supplemented with 0.15 M NaCl. The column was previously calibrated

with the Kit for Molecular Weights 12,000-200,000 Da (Sigma) and manufacturer instructions were followed to determine the molecular weight.

2.1.5. *In vitro* enzymatic activity assay

2.1.5.1. Chromogenic substrates

Enzymatic activity of the recombinant PaCCP was tested against specific synthetic substrates of CPs. Carboxypeptidase A (CPA)-like substrates N-(4-methoxyphenylazoformyl)-l-phenylalanine (AzoF) (Mock et al., 1996) and N-benzoyl-Gly-Phe (Hippuryl-L-phe) (Bergmeyer, 1974); Carboxypeptidase B (CPB)-like substrates furyl-acryloyl-l-alanine-l-lysine (FAAK) (Plummer and Kimmel, 1980) and N-(4-methoxyphenylazoformyl)-l-arginine (AzoR) (Mock and Stanford, 2002) were purchased from Sigma and Bachem. Carboxypeptidase O (CPO)-like substrate N-benzoyl-l-alanyl-l-glutamic acid (BzAE) was kindly provided by Dr. Michael Edge (AstraZeneca, Wilmington, DE, USA). Dimethyl sulfoxide (DMSO, Sigma) was used as cosolvent in the assay buffer at a final concentration less than 2% (v/v). Enzymatic reactions were developed in a final volume of 1 ml of aqueous buffer containing 20 mM Tris pH 7.5 (pH 8.0 in assay with BzAE) supplemented with 150 mM NaCl. The final concentrations of substrates in the reaction were 0.1 mM for AzoF and AzoR, 2 mM for Hippuryl-L-phe, 0.2 mM for FAAK, and 0.75 mM for BzE. The rate of hydrolysis of each substrate was continuously monitored spectrophotometrically at 25 °C by measuring the absorbance at 350 nm (AzoR and AzoF), 330 nm (FAAK), or 254 nm (Hippuryl-L-phe and BzAE). The final concentration of PaCCP in the assays was 27 μM.

The activity of PaCCP against AzoF was also measured in 96 well plates to test buffers of different composition. The reaction mix contained 0.1% of detergent 3-[(3-cholamidopropyl) dimethylammonio]-1-propanesulfonate (CHAPS, Sigma) (w/v) and 1% DMSO (Sigma). The following buffers, at concentration 0.1 M were tested: Bistris pH 5.2, Bistris propane pH 6.0, MES pH 6.0, Tris-HCl pH 7.0; pH 7.5; pH 8.0 and pH 8.5 and Bicine pH 9.0 supplemented with 0.3 M or 1 M NaCl. Final concentrations of PaCCP and AzoF in the assay were 20 nM and 0.1 mM, respectively. The reaction was monitored in a plate reader ELISA iEMS *Reader* MF (Labsystems) at λ = 350 nm and 25 °C. The PaCCP Y64A mutant was assayed against the AzoF, AzoR and BzAE in the same conditions described above. The final concentration of the protein in the assays was 30 μM.

2.1.5.2. Synthetic peptide substrates

Protein PaCCP activity was also tested against synthetic peptide substrates with different C-terminal amino acids purchased from American Peptide Company (Sunnyvale, USA). The following peptides were used: Dynorphin B (1-9) (Tyr-Gly-Gly-Phe-Leu-Arg-Arg-Gln-Phe); Dynorphin A (1-9) porcine (Tyr-Gly-Gly-Phe-Leu-Arg-Arg-Ile-Arg); ACTH (4-11) human (Met-Glu-His-Phe-Arg-Trp-Gly-Lys); Ser-Tyr-Leu-Gln-Asp-Ser-Val-Pro-Asp-Ser-Phe-Gln-Asp; β-Amyloid (1-11) (Asp-Ala-Glu-Phe-Arg-His-Asp-Ser-Gly-Tyr-Glu) and IRL-1720 (Ac-Asp-Lys-Glu-Ala-Val-Tyr-Phe-Ala-His-Leu-Asp-Ile-Ile-Trp). Reactions were performed in 50 mM Tris-HCl pH 7.5 supplemented with 5 mM NaCl at 37 °C using peptide and enzyme amounts of 2 μmoles and 10 nmoles, respectively. Reactions with bCPA (Sigma), CPB (Sigma) and CPO (kindly gift by Dr. Sebastian M. Tanco, Enzymology and Proteomic Laboratory, IBB) were used as positive controls. Reactions were stopped at 7 h by lowering the pH to 4.0 using 1% trifluoroacetic acid (Sigma). Reaction products were desalted using a C18 ZipTip (Millipore Ibérica), following the manufacturer's protocol. Peptides were eluted with 1 μl of 70% acetonitrile, 0.1% trifluoroacetic acid mixed with an equal volume of a saturated solution of α-hydroxycinnamic acid (Bruker) in the same solvent. Analysis was performed in a Bruker Daltonics Ultraflex matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometer.

Others synthetic peptide substrates tested, with the general formula N-acetyl-Gly-X (X = Tyr; Lys; Glu or Asp) were synthesized at the Combinatorial Chemistry Unit, Barcelona Science Park, Barcelona University. The reactions took place overnight, at 37 °C, in the buffer 50 mM Tris-HCl pH 6.5 supplemented with 100 mM NaCl in the presence of 2 nM of PaCCP or PaCCP Y64A and 10mM of peptide. The reactions were analyzed using the matrix 2',4',-Dihydroxyacetophenone (Sigma) in a Bruker Daltonics Ultraflex MALDI-TOF mass spectrometer.

2.1.5.3. Purified porcine tubulin

The enzymatic reaction with purified porcine brain tubulin (Cytoskeleton, Inc) was used to evaluate the ability of the PaCCP C320R mutant to hydrolyze polyglutamates chains. The reactions took place in the activity buffer, 80 mM PIPES pH 6.8 supplemented with 1 mM MgCl2 and 1/500 EDTA-free protease inhibitor cocktail Set III (Calbiochem). The reactions

contained 0.625 µg of purified tubulin and 40 µg PaCCP C320R. Human CCP6 (0.2 µg) was used as control of a positive reaction and no enzyme was added in negative controls. After 5 h or overnight incubation at 37 °C, the reactions were stopped with sample buffer 5X and boiled for 5 min prior to electrophoresis. The resulting protein samples were resolved on a 10% SDS-PAGE, as describe above, before Western blot analysis.

2.1.6. Western blot

After SDS-PAGE, proteins samples were transferred onto polyvinylidene diflouride membranes (PVDF) (Immobilon-P, Millipore) (Matsudaira, 1987). Transference was made in buffer Tris-Gly pH 8.3, 10% methanol at 4 °C in a BioRad Mini Trans-Blot Cell at 100 V, for 1.5 h. The membrane was blocked with 5% skim milk (w/v) in phosphate buffered saline containing 0,05% Tween 20 (v/v) (PBS-T), 1 h at room temperature (RT) followed by 4 washing steps of 10 min at RT with PBS-T. Primary antibody rabbit anti-polyglutamilated tubulin 1:5,000 (polyE long chain, gift of Dr. M.A. Gorovsky, Departament of Biology, University of Rochester) was incubated with the membrane for 1 h at RT. After four washes with PBS-T, the filters were incubated with HRP-labeled donkey anti-rabbit 1:5,000 (GE Healthcare) 1 h at RT. After four rinses, immunoreactive bands were visualized with Luminata Forte Western HRP substrate (Millipore) in a Molecular Imager VersadocTM (BioRad).

2.1.7. Interaction of PaCCP with carboxypeptidases inhibitors

The formation of complexes between PaCCP and carboxypeptidases (CPs) inhibitors latexin, TCI (Tick Carboxypeptidase Inhibitor), NvCI (*Nerita versicolor* Carboxypeptidase Inhibitor) and PCI (Potato Carboxypeptidase Inhibitor) was assessed by non-denaturing PAGE. The inhibitors were recombinantly produced at Enzymology and Proteomic Laboratory, IBB. The protein and inhibitors were mixed in a molar ratio of 4:1 (PaCCP:inhibitor) and incubated in the buffer 20 mM Tris-HCl pH 7.5 supplemented with 50 mM NaCl at RT, in a final volume of 50 µl. After 1 h, the formation of complexes was examined by 10% non-denaturing PAGE and visualized with Coomasie blue dye.

2.1.8. Crystallization and data collection

Crystals of the PaCCP were obtained at 18 °C by sitting-drop vapor diffusion method. The reservoir solution contained 0.2 M magnesium formate and 20% PEG3350. Single crystals appeared after 2 days from equal volumes of protein solution (5 mg/ml in 5 mM Tris, pH 8.0 supplemented with 25 mM NaCl) and reservoir solution. Crystals were cryo-protected in reservoir buffer containing 12% glycerol and flash-frozen in liquid nitrogen prior to diffraction analysis. Diffraction data were recorded from cryo-cooled crystals (100 K) at Grenoble beamline ID23-2 (European Molecular Biology Laboratory, Grenoble, France). Data were integrated and merged using XDS (Kabsch, 2010), and scaled, reduced, and further analyzed using CCP4 (Winn et al., 2011). To obtain the complexes with CP inhibitors: guanidinoethylmercaptosuccinic acid (GEMSA; Sigma) and benzylsuccinic acid (BZS; Sigma), PaCCP crystals were soaked for 3 days in reservoir solution containing 10 mM of each compound.

Table 2.1. Data collection and refinement statistics.

Statistic	PaCCP	PaCCP -BZS	PaCCP -GEMSA
Data Collection			
Space group	$P2_1$	$P2_1$	$P2_1$
Cell dimensions			
a, b, c (Å)	41.89, 83.85, 107.46	42.10, 84.02, 107.69	41.20, 83.73, 105,58
α, β, γ (°)	90.00, 96.37, 90.00	90.00, 96.62, 90.00	90.00, 100.78, 90.00
Resolution (Å)	25 - 1.55 (1.66 - 1.55)	53.5 – 2.0 (2.1 – 2.0)	44.10 – 1.5 (1.58 – 1.5)
Rmerge†	0.118 (0.447)	0.105 (0.564)	0.071 (0.233)
I∕σI	9.5 (2.4)	7.7 (1.5)	7.2 (2.5)
Completeness (%)	89.7 (88.5)	98.0 (88.9)	88.4 (72.1)
Redundancy	6.9 (4.1)	3.2 (2.5)	2.4 (1.8)
Refinement			
Resolution (Å)	25 - 1.6	25 – 2.0	25 – 1.6
No. reflections	87768	49724	84517
Rwork ∕Rfree ‡	19.89 / 21.36	23.26 / 27.22	21.20 / 23.04
No. Atoms	6602	6299	6354
Protein	751	751	751

Water	557	246	277
R.m.s deviations			
Bond lengths (Å)	0.0053	0.0066	0.0049
Bond angles (°)	1.436	1.362	1.345

*Statistic for highest resolution shell is shown in parentheses.

†Rmerge 1⁄4 ∑ jIi − < I > j⁄ ∑ Ii , where Ii is the ith measurement of the intensity of an individual reflection or its symmetry-equivalent reflections and <I > is the average intensity of that reflection and its symmetry-equivalent reflections.

‡Rwork 1⁄4 ∑ jjFobsj − jFcalcjj⁄∑ jFobsj for all reflections and Rfree 1⁄4 ∑ jjFobsj − jFcalcjj⁄∑ jFobsj, calculated on the 5% of data excluded from refinement.

2.1.9. Structure determination and refinement

The structure for the PaCCP was determined from the X-ray data at 1.6 Å by molecular replacement using the Protein Data Bank (PDB; http://www.wwpdb.org) information from *Shewanella denitrificans* (3l2n) as a model. The electron density map produced from molecular replacement was improved to build up a model for the PaCCP using the program Coot (Emsley et al., 2010). Refinement utilized CNS (Brunger, 2007) and Phenix (Adams et al., 2010). Crystals of the GEMSA and BZS complexes with PaCCP were determined with molecular replacement using the refined PaCCP structure. Ramachandran analysis shows 95.58% of residues (714 residues) are in preferred regions, 3.88% of residues (29 residues) are in allowed regions, and 0.54% of residues (4 residues) are in outlier regions, for both molecules in the asymmetric unit for the wt and complex structures. Refinement and data statistics are provided in Table 2.1. The structure representations were prepared with PyMOL (Schrodinger, 2010).

2.1.10. Accession codes

Protein Data Bank (PDB) coordinates and structure factors from the 3 structures were deposited in the PDB archive with accession codes 4a37, 4a38, and 4a39.

2.1.11. Comparative modeling of human CCPs

The CP- and N-domains of the six human CCPs (CCPs 1,2,3,4,5 and 6;) were modeled by comparative modeling using the PaCCP structure as template. The boundaries of these domains were defined based on sequence alignment and PsiPred secondary structure predictions (McGuffin et al., 2000). An initial alignment of the sequence of PaCCP and the

sequences to be modeled was performed using T-coffee program (Notredame et al., 2000). This alignment was manually modified based on 3D structure information, secondary structure predictions, and the conservation of catalytic residues and motifs of the M14D subfamily (Rodriguez de la Vega et al., 2007). The models were obtained using the I-Tasser modeling procedure (Roy et al., 2010). The confidence level of the models was evaluated by the C-score and the TM-score (Roy et al., 2010). The structure representations were prepared with PyMOL (Schrodinger, 2010).

2.1.12. Database Searches

An *in silico* search for proteins with acidic C-terminals in the proteome of *P. aeruginosa* MPAO1 was performed using ScanProsite (de Castro et al., 2006) and the pattern [ED](n)>, where n represents the number of acidic residues at a protein's C-terminus which was allowed to vary from 1 to 5. The pattern search was performed in the *P. aeruginosa* MPAO1 (taxon ID: 1131757) of the 2014_03 UniProtKB/Swiss-Prot database, using the default settings.

2.2 Materials and Methods used in Chapter 2.

2.2.1. Bacterial strains and media

The bacterial strains used in this study are listed in Table 2.2. *P. aeruginosa* PAO1 wild-type (MPAO1) strain and the negative mutants deficient in *paccp*, *pilA* and *fliC* genes were purchased from the distributor of the PAO1 mutant library of the University of Washington (Jacobs et al., 2003). The identity of the mutant deficient in *paccp* gene was checked by PCR, as recommended by the authors of the PAO1 mutant library (Jacobs et al., 2003). Unless stated otherwise, bacterial strains were grown at 37 °C in liquid cultures in LB broth or on solid media by addition of 1.5% agar. Antibiotic gentamicin was supplemented as follows: 10 µg/ml for *E. coli* or 40 µg/ml for *P. aeruginosa*. All strains were maintained as glycerol stocks at -80 °C after one passage on LB broth. For growth curves, strains were cultured in 50 mL of LB broth and grown at 37 °C with shaking at 250 rpm to stationary phase. Aliquots of 1 ml were taken and O.D. at $\lambda = 600$ nm was measured in a spectrophotometer.

Table 2.2. Strains used in this study and their relevant features.

Strains	Relevant characteristic (s)	Source or Reference
MPAO1	PAO1 wild-type, laboratory strain	(Jacobs et al., 2003)
Δ*paccp*	MPAO1 transposon mutant PW5744, Tn5ISphoA-hah insertion in *paccp* (PA2831)	(Jacobs et al., 2003)
Δ*pilA*	MPAO1 transposon mutant PW8622, Tn5ISphoA-hah insertion in *pilA* (PA4525)	(Jacobs et al., 2003)
Δ*fliC*	MPAO1 transposon mutant PW2971, Tn5ISphoA-hah insertion in *fliC* (PA1092)	(Jacobs et al., 2003)
MPAO1/pBBR5	PAO1 wild type transformed with the empty vector pBBR1-MSC5	This study
Δ*paccp* /pBBR5	Δ*paccp* mutant strain transformed with the empty vector pBBR1-MSC5	This study
MPAO1/pBBR5-*paccp*	PAO1 wild type strain transformed with the vector *paccp*-pBBR1-MSC5	This study
Δ*paccp*/pBBR5-*paccp*	Δ*paccp* mutant strain complemented with the vector *paccp*-pBBR1-MSC5	This study

2.2.2. Genetic construct for complementation

For expression of the gene encoding PaCCP under its own promoter, the *paccp* coding sequence, including 397 bp upstream and 238 bp downstream, was amplified from genomic DNA of *P. aeruginosa* MPAO1 (purchase from PAO1 mutant library (Jacobs et al., 2003)). PANGEA High Fidelity DNA Polymerase Master Mix (Canvax) was used for PCR following the instructions of the manufacturer. The primers: Compfw 5'-GCA GAG CTC CCA GCA TCA TCG TCA TGT GGT TCT C-3' and Comprv2-5' GGT AGT CGA TGG TGA CCA GCA AGC C ACT AGT GCG-3', were used for amplification. To clone into vector pBBR1-MSC5 (Kovach et al., 1995), the amplification product and vector were digested with restriction enzymes *Sac*I and *Spe*I (Thermo Fisher Scientific Biosciences GmbH) and purified from 1% agarose gel (w/v) with GeneJET Gel Extraction Kit (Thermo Scientific). Both digested DNA fragments were ligated with T4 ligase (New England Biolabs) following the manufacturer's instructions. The ligation mix was transformed into chemically competent *E. coli* XL1-Blue cells and plated on LB agar plates containing 10 μg/ml gentamicin. Five colonies were selected for plasmids purification with GeneJET Plasmid Miniprep Kit (Thermo Scientific)and sequenced. Given the length of the insert, sequencing was performed with primers M13fw and M13rv and the internal oligonucleotides: PaCCPintermfw 5'-TAC TTC GAG CCC TAC AGC-3' and PaCCPintermrv 5'-

CTG AAG CTG TGG GTC ATC-3'. All primers were synthesized by Invitrogen, Life Technologies, and DNA inserts sequenced at the Genomic Service Facility (IBB, UAB).

2.2.3. Transformation of *P. aeruginosa* MPAO1

The empty vector pBBR1-MSC5 and *paccp*-pBBR1-MSC5 construct were transformed into MPAO1 and Δ*paccp* strains by electroporation following the protocol of Choi et al. (2006) with some modifications. Briefly, electro-competent cells were prepared by washing two times the cell pellet, obtained from 1 ml overnight culture, with 1 ml of ice-cold 300 mM sucrose and resuspended in a final volume of 100 µl of 300 mM sucrose. The purified plasmid DNA (1 µg) was mixed with 100 µl of electro-competent cells and the mixture was transferred to a 2 mm gap-width electroporation cuvette. After applying a pulse (settings: 25 µF; 200 Ω; 2.5 kV on a Bio-Rad GenePulserXcell™, Bio-Rad), 1 ml of RT LB medium was added and cells were transferred to an eppendorf tube for 1 h of additional incubation at 37 °C with rotatory shaking (250 rpm). The cells were then harvested by centrifugation at 10,000 g, 2 min and the cell pellet was resuspended in 100 µl of residual medium. The entire mixture was then plated onto LB plates containing 40 µg/ml gentamicin. The plates were incubated at 37 °C until colonies appeared. The presence of the plasmids, in gentamicin resistance colonies, was checked by PCR with primers M13fw and M13rv.

2.2.4. Expression of *paccp* gene in *P. aeruginosa* MPAO1

Cultures of MPAO1 and Δ*paccp* strains, transformed with empty pBBR1-MSC5 plasmid or *paccp*-pBBR1-MSC5, were grown overnight at 37 °C with rotatory shaking at 250 rpm, in LB broth containing gentamicin 40 µg/ml. Identical fresh medium was inoculated with a dilution 1:100 from the overnight cultures (starting O.D. 600 nm = 0.1) and incubated for 4 h, at 37 °C and rotatory shaking at 250 rpm. Samples for SDS-PAGE were prepared by taking a volume of each culture corresponding to 1/(O.D. at 600 nm), to ensure equivalent amount of cells. Then, the cells were harvested by centrifugation 2 min at 10,000 g, resuspended in 100 µl of reducing sample buffer 1X and boiled for 5 min before loading onto 10% SDS-PAGE. The electrophoresis was carried out as described in section 2.1.3. After SDS-PAGE, the proteins were transferred to PVDF membranes (Immobilon-P, Millipore) and Western blot was performed with primary mouse polyclonal antibodies raised against PaCCP, 1:1,000 dilution (obtained at Antibody Production's Facility, IBB, Barcelona, Spain) and secondary

antibody HRP-labeled goat anti-mouse 1:5,000 (Biorad). Western blot protocol is detailed in section 2.1.6.

2.2.5. Minimum inhibitory concentration to antibiotics

Antimicrobial susceptibility to a variety of drugs including ceftazidime, imipenem, ciprofloxacin, amikacin, piperacillin, tobramicin and colistin was evaluated by the Etest method (bioMérieux Clinical Diagnostics) as follows. Cells of MPAO1 or Δ*paccp* were grown overnight in LB medium at 37 °C with rotatory shaking at 250 rpm. The cellular pellets were resuspended in 2 ml of 0.9% NaCl to O.D. at 600 nm = 0.1 and Mueller-Hinton agar (Oxoid) plates were inoculated by swabbing the bacterial suspension onto the entire agar surface. One Etest strip was placed on each plate and incubated overnight at 37 °C after which the Minimum Inhibitory Concentration (MIC) was read. Three replicas were done for each strain. The MIC was the value where the growth inhibition ellipse intersected the strip. Interpretive breakpoints used were those recommended by the Clinical and Laboratory Standards Institute (CLSI) guidelines (CLSI, 2012).

2.2.6. Casein Hydrolysis Test

Protease activity was evaluated using a casein hydrolysis test (Murray, 1999). In brief, the MPAO1 or Δ*paccp* strains cultures were grown in LB broth at 37 °C, with rotatory shaking at 250 rpm to O.D. at 600 nm = 1 and 2 μl of this suspension were spotted onto Mueller-Hinton agar (Oxoid) plates supplemented with 2% skim milk (w/v) (Difco). The plates were incubated overnight at 37 °C. The presence of a transparent zone around the inoculums spot indicated a positive test. Five replicas were done for each strain. Graphic and statistics were performed in GraphPad Prism 4.00.

2.2.7. Biofilm formation

Biofilm formation was measured in a microtiter plate assay following the protocol of O'Toole, et al. (2011) with some modifications. The MPAO1 or Δ*paccp* strains, cultured overnight, in LB broth at 37 °C, with rotatory shaking at 250 rpm were resuspended in fresh LB to O.D. at 600 nm = 0.1. A volumen of 200 μL per well of the dilution was added to untreated 96-well polystyrene plates (BrandTech 781662), and incubated 24 h at 37°C. Then, the liquid was shaken out, the plate dried 15 min at 65 °C and 250 μL of 0.1% crystal

violet (CV) were added to each well followed by incubation for additional 15 min at RT. After washing with distillated water, 200 μL of ethanol 95% were added to each well of the microtiter plate. Finally, 150 μL of the solubilized CV were transfer to a new microtiter plate and the absorbance was measured in a plate reader at 620 nm. Eight replicas were done per strain and LB broth was used as control. Graphic and statistics were performed in GraphPad Prism 4.00.

2.2.8. Motility Assays

2.2.8.1. Swimming motility

The protocol proposed by Rashid and Kornberg (2000) was followed with some modifications. Media used for assay was LB broth supplemented with 0.25% Agar-Noble (w/v) (Difco). The MPAO1 or Δ*paccp* strains cells grown overnight at 37 °C in 1.5% LB agar (w/v) plates, were resuspended in LB media to O.D. at 600nm = 0.1 and 5 μl of the suspension inoculated in the center of the swim plate (three plates per strain). After incubation overnight at 30 °C, the diameter of the growth area was measured. Three replicas were done per strain. Graphic and statistics were performed in GraphPad Prism 4.00.

2.2.8.2. Swarming motility

For evaluation of swarming motility, the protocol in Xavier et al. (2011) was followed. The minimal medium for swarming plates was prepared using the following recipe: 800 ml of milliQ water with 6.25% agar (Difco), 200 ml of 5X stock phosphate buffer, 1 ml of 1 M Mg_2SO_4, 0.1 ml of 1 M $CaCl_2$, 25 ml of 200 g/L solution of casamino acids (BactoTM from BD, Sparks, MD). One litre of 5X stock phosphate buffer was prepared by dissolving 12 g of Na_2HPO_4, 15 of KH_2PO_4 and 2.5 g of NaCl into 1 L of miliQ water. The final pH of medium was 6.7. Each swarming plate was prepared by pouring exactly 20 ml of medium onto a Petri dish and allowed to cool upright for 30 min. The plates were then turned upside down and left at RT to dry for 15 h. The inocula were prepared from 1 ml of overnight cultures washed twice with phosphate buffered saline 1X (PBS). Plates inoculation was carried out by spotting a 2 μl drop at the center of the swarming plate and allowed to dry. Plates were then placed upside down at 37 °C for 24 h. Three replicas were done per strain and photographs of the resulting swarming phenotypes were taken.

2.2.8.3. Twitching motility

Twitching motility was assessed using the stab assay as described by Rashid and Kornberg (2000). Briefly, twitch plates containing LB broth solidified with 1% Bacto-agar (w/v) (Difco) were stab inoculated with a sharp toothpick to the bottom of the Petri dish from an overnight-grown 1.5% LB agar (w/v) plate. After incubation at 37 °C for 24 h, the resulting zones of twitching motility were visualized by carefully removing the agar, dried 15 min at 60 °C and staining the bacteria, adhering to the petri dish, with 1% CV (w/v) for 15 min at RT. Finally, a brief rinse with tap water was done to remove unbound dye. Five replicas were done per strain and photographs taken of the resulting twitching zones.

2.2.9. Transmission electron microscopy

Bacteria from MPAO1 or Δ*paccp* strains were examined for the presence of pili by transmission electron microscopy in the Microscopy Facility, UAB, Barcelona. Bacterial strains were grown in LB plates overnight at 37 °C. Colonies of each plate were gently suspended in PBS and 5 µl of this solution were placed on the corresponding carbon-coated grid. After 5 min, the grids were dried with filter paper from the edge and stained with 2.5 µl of 2% uranyl acetate for 45 s. Visualization was made via a JEOL JEM 1400 (Japan) microscope operated at 80 kV.

2.2.10. Analysis of sheared surface proteins

Cell surface appendages (flagella and pili) were isolated following the protocol in Asikyan *et al*., (2008). Bacteria were streaked to cover all the surface of LB agar plates or LB agar plates supplemented with gentamicin 40 µg/ml when appropriate and incubated overnight at 37 °C. The cells were gently removed from the agar surface with an inoculation loop and resuspended in 2 ml sterile PBS, pH 7.4 per sample. The number of cell used was normalized by dilution of the samples to the same optical density at 600 nm. Surface proteins were sheared by vigorous vortexing for 1 min and the suspension centrifuged for 5 min at maximum speed to pellet the cells. The supernatant was transferred to a new tube and centrifuged for an additional 25 min at maximum speed, at RT to remove any remaining cells. To precipitate the sheared proteins, 1/10 volumes each of 5 M NaCl and 30% polyethylene glycol (molecular weight range, 8,000) were added to the supernatant and the

samples incubated on ice for 1 h. Samples were centrifuged at maximum speed for 25 min at 4 °C and the pellets were resuspended in SDS-PAGE sample buffer and boiled for 5 min. The resulting protein samples were resolved on a 15% SDS-PAGE. The identity of the gel band corresponding to PilA was confirmed by MALDI-TOF mass spectrometry analysis carried out in the Proteomics Facility at UAB, Barcelona.

2.2.11. Analysis of pilins in whole-cell lysates

After being vortexed to remove surface proteins, the harvested cell pellet was resuspended in sterile PBS to a final O.D. at 600 nm of 0.5. A 200 µl aliquot of the cell suspension was centrifuged at maximum speed for 5 min to harvested the cells. The supernatant was removed and the cell pellet was resuspended in 100 µl of SDS-PAGE sample buffer and boiled for 5 min, and 10 µl per sample was separated on a 15% SDS-PAGE minigel. After separation, the proteins were transferred to PVDF membrane (Immobilon-P, Millipore) and Western blot analysis was performed as described in section 2.1.6. PilA was detected with rabbit polyclonal antibodies raised against PilA, 1:10,000 dilution (kindly donated by Dr. Barbara I. Kazmierczak, Departments of Medicine & Microbial Pathogenesis, Yale University School of Medicine) followed by donkey anti-rabbit, HRP-labeled antibody 1:5,000 dilution (GE Healthcare).

2.2.12. *Caenorhabditis elegans* Slow-Killing Assay

Virulence of MPAO1 or Δ*paccp* strains was determined in *Caenorhabditis elegans* based on the "Slow-Killing" method (Tan et al., 1999). *C. elegans* sterile temperature-depending strain CF512 (fer-15(b26) II; fem-1(hc17) IV) was provided by the Caenorhabditis Genetics Center, University of Minnesota. The worms were routinely maintained in NGM plates (1.7% Agar, 50 mM NaCl, 0.25% peptone, 1 mM CaCl2, 5 µg/ml cholesterol, 25 mM KH2PO4, 1 mM MgSO4) seeded with *E. coli* OP50, at 16 °C. *P. aeruginosa* MPAO1 and Δ*paccp* strains were grown in LB broth, overnight at 37 °C and 100 µl of each culture were spread on a 5.5 cm diameter NGM agar plate and incubated at 37 °C for 24 h. Next day, each plate was seeded with 15-20 adult hermaphrodite CF512 worms. Plates were incubated at 25 °C and scored for live worms every 24 h. *E. coli* OP50 was used as a negative control. A worm was considered dead when it no longer responded to touch. Two replicates per strain were carried out.

2.2.13. Interactomic assays using magnetic beads

2.2.13.1. Preparation of MPAO1 cell lysate

The soluble fraction of MPAO1 cells, used in interactomic assays was prepared as follow. Overnight cultures grown at 37 °C with rotatory shaking at 250 rpm were diluted 1:100 to inoculate fresh medium. The fresh cultures grown under the same conditions for 4 h after which the pellet was harvested by centrifugation at 10,000g, 10 min and resuspended in buffer W (100 mM Tris-HCL pH 8.0 containing 150 mM NaCl) supplemented with the cocktail protease inhibitor EDTA-free complete 1X (Roche). The cells were disrupted by sonication and cleared by centrifugation at 20,000 g for 20 min. The concentration of the resulting protein extract was estimated by measuring the absorbance at 280 nm and adjusted to 5 mg/ml.

2.2.13.2. Isolation of protein complexes of recombinant PaCCP

Magnetic beads containing surface epoxy groups (Dynabeads M-270 Epoxy, Invitrogen) were used to isolate proteins interacting with recombinant PaCCP. The workflow is depicted in Figure 2.1A. All buffer changes and washing steps were performed in a magnetic rack, following manufacturer´s instructions. The beads were weighed (5 mg per sample), resuspended in buffer 0.1 M sodium phosphate pH 7.4 and incubated 10 min at RT with tilting/rotation. For coupling, 100 μg of recombinant purified PaCCP (in PBS) were added to the beads, per sample, and incubated overnight, at 37 °C, with tilting/rotation, in the presence of 1 M ammonium sulfate. The coupling solution was discarded and the beads were washed with 1 ml PBS. After washing, 1 ml of MPAO1 cell lysate was added per sample and incubated overnight, at 4 °C with tilting/rotation. Samples were washed four times with PBS before proteins elution from beads with 20 μl of 1% formic acid followed by nano-ESI-MS/MS analysis. Samples of background control were prepared following the same protocol except that no PaCCP was coupled to the beads.

2.2.13.3. Isolation of protein complexes of endogenous PaCCP

The epoxy beads (Dynabeads M-270 Epoxy, Invitrogen) were also used to immunoprecipitate endogenous PaCCP and associate protein complexes (Figure 2.1B). The above-described protocol was used but, in this case, the epoxy beads were coated with 50

µg of purified polyclonal serum raised against PaCCP (obtained at Antibody Production's Facility, IBB). As background control, beads coupled with 50 µg of irrelevant purified IgGs from human serum (kindly gift by Camilla Lund, Laboratory for proteome analysis and protein characterization, University of Aarhus, Denmark). The washing steps were carried out with PBS-T. In the same way, beads were incubated with MPAO1 lysate and proteins were eluted from beads with 20 µl of 1% formic acid followed by nano-ESI-MS/MS analysis.
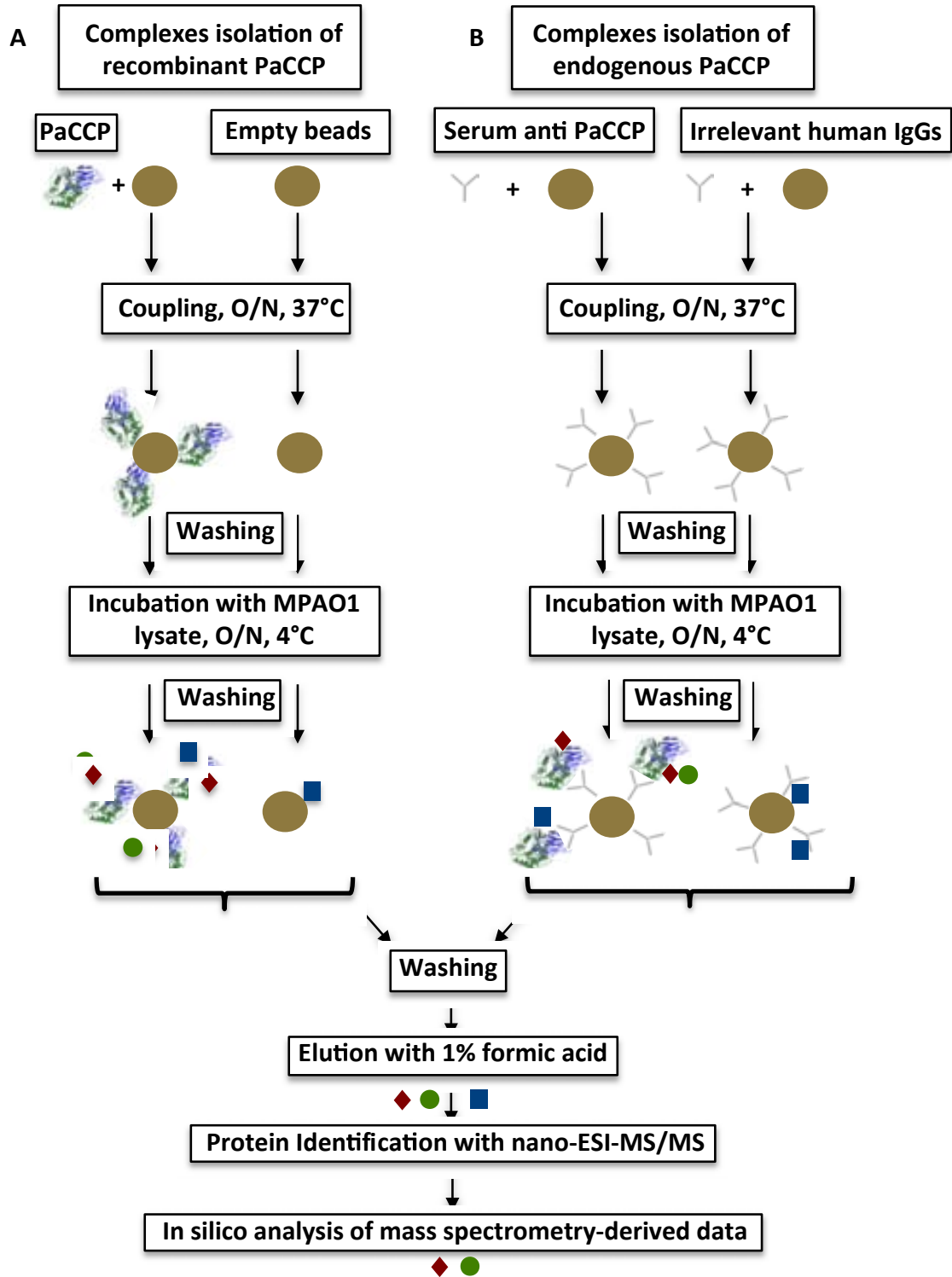
Figure 2.1. Workflow for the isolation and identification of interacting partners of recombinat (A) or endogenous (B) PaCCP using magnetic beads and nano-ESI-MS/MS. O/N: overnight.

2.2.13.4. Protein identification with nano-Esi-MS/MS

Proteins samples for nano-Esi-MS/MS were denatured with 6 M Urea; reduced with 5 mM dithiothreitol, for 30 min at RT; alkylated with iodoacetamide 15 mM, for 20 min at RT and digested overnight, at 37 °C with trypsin (1 μg trypsin (Sigma)/50 μg of sample proteins). Tryptic peptides were micro-purified using C18 stage tips (Proxeon, Thermo Scientific). Nano-ESI-MS/MS analyses were performed on an EASY-nLC II system (ThermoScientific) connected to a TripleTOF 5600 mass spectrometer (AB Sciex) equipped with a NanoSpray III source (AB Sciex) and operated under Analyst TF 1.5.1 control. The collected MS files were converted to Mascot generic format (MGF) files using the AB SCIEX MS Data Converter beta 1.1 (AB SCIEX) and the "proteinpilot MGF" parameters. The generated peak lists were searched against the NCBI database using an in-house Mascot search engine (Matrix Science). Spectra were searched with trypsin enzyme settings, allowing one missed cleavage. Carbamidomethylation of cysteines was set as a fixed modification, and mass tolerance of the precursor ion and ion fragment were set to 10 ppm and 0.3 Da respectively. Each experiment was performed independently, at least three times and proteins identified in the three replicas are presented. Protein score is -10*Log(P), where P is the probability that the observed match is a random event. Protein scores greater than 100 are significant ($p < 0.05$).

2.2.14. Statistical analysis

The results are given as mean ± standard error of the mean. Student's t test analysis was performed and differences were considered significant when $p < 0.05$. Survival data were analyzed using Kaplan-Meier survival curves with GraphPad Prism 4.00 (GraphPad Software Inc.) and significance was tested by Log-rank (Mantel-Cox) analysis.

2.3. Materials and Methods used in Chapter 3CCP6

2.3.1. Cloning and DNA sequencing

cDNA encoding human CCP6 was purchased from the Biological Resource Center (NBRC, National Institute of Technology and Evaluation, Japan, code: AGBL4cDNA-AK094826-FLJ37507-BRAWH2018206-clon2). A shorter version of CCP6 (lacking 37 amino acids from the C-terminus) was amplified by two consecutive PCRs for cloning into the

pTriEx6 vector (Novagen). As a result of this cloning design, both *Strep*-tag II (Schmidt and Skerra, 2007) and hemagglutinin epitope (HA) tag were introduced at the protein N-terminus. The following primers were used: CCP6fw1, 5´-CGA TGT TCC AGA TTA CGC TGG TGC AAT GGC GGA GGG GAG CCA GTC-3´; CCP6fw2, 5´-GCA GTT CGA AAA GGG TGC ATA CCC ATA CGA TGT TCC AGA TTA CGC TGG TG-3´ and CCP6rv2, 5´-TTA AGG TAC CCT AAA AAG GGG TTG AAG GGT CCT TGT GG-3´ (synthesized by Invitrogen, Life Technologies). The PCR reactions were performed with Phusion High-Fidelity DNA Polymerase (Thermo Scientific) following the manufacturer´s instructions. The final amplification product and the pTriEx6 vector were digested with BstBI/KpnI (Thermo Scientific) purified from 1% agarose gel (w/v) using GeneJET Gel Extraction Kit (Thermo Scientific) and ligated with T4 ligase (New England Biolabs). The ligation mix was transformed into chemically competent *E. coli* XL1-Blue cells and plated on LB agar plates containing 50 µg/ml ampicillin. The CCP6 (E425Q) mutant (equivalent to E270Q of bovine CPA) was generated by the primer extension method (Ho et al., 1989). Primers for amplification of fragment 1 (upstream of the point mutation) were: hCCP6E270Qfw 5´-CAG TTC GAA AAG GGT GCA TAC CCA TAC GAT G-3´and hCCP6E270Qfusionrv 5´-AGA CCT GTA GGG TGT AGC AAT AGG AAG TGT G-3´. For amplification of fragment 2 (downstream of the point mutation) the following primers were used: hCCP6E270Qfusionfw 5´-CTA CAC CCT ACA GGT CTC CTT CTA CAG CTA C-3´ and CCP6rv2 (described above). Mutagenesis PCR reactions were carried out using PANGEA High Fidelity DNA Polymerase Master Mix (Canvax) following the manufacturer's instructions. The amplification product was cloned into pTriEx6 using the restriction enzymes BstBI and KpnI. In each case, ten colonies were selected for plasmid purification with GeneJET Plasmid Miniprep Kit (Thermo Scientific) and sequenced bidirectionally using T7 forward and TriExdown reverse primers in the Genomic Facility IBB, UAB.

The high mobility group protein B3 (HMGB3, UniProt ID: O15347) and TRAF-type zinc finger domain-containing protein 1 (TRAFD1, UniProt ID: O14545) were obtained in our research group for co-expression experiments. Both proteins were cloned into the pTriEx6 vector for expression of N-terminal *Strep-tag* II (Schmidt and Skerra, 2007) and hemagglutinin epitope (HA) fusion proteins (Tanco et al., 2014, submitted).

2.3.2. Cell Culture and transfection

HEK 293T cells (American Type Culture Collection (ATCC), Barcelona, Spain) were cultured in Dulbecco's Modified Eagle's Medium with 4500 mg/L glucose, 2 mM L-alanyl-L-glutamine (GlutaMAX–I) and 1 mM pyruvate supplemented with 10% fetal calf serum (Invitrogen, Carlsbad, CA, USA). Cells were grown in a constant temperature incubator set at 37 °C and 10% CO2. HEK 293F cells (Invitrogen) were cultured in FreeStyle 293 expression medium (Invitrogen). Cells were grown in flasks on a rotary shaker (120 rpm) at 37 °C in 8% CO_2 and 70% relative humidity atmosphere.

DNA transfections were carried out using polyethylenimine (PEI), linear, 25kDa (PolySciences). Briefly, DNA was mixed with PEI at a ratio of 1:3 and incubated in presence of serum-free medium for 15 min at room temperature (RT). HEK 293T, at 40-50% confluence, were exposed for 72 h to the transfection complex, using 1.4 μg DNA per ml of culture. For HEK 293F cells, DNA was mixed with PEI at a ratio of 1:3 and incubated in FreeStyle 293 expression medium for 15 min at RT. HEK 293F, at $1x10^6$ cells per ml, were exposed for 72 h to the transfection complex, using 1 μg DNA per ml of culture.

2.3.3. Preparation of clarified protein extracts

Transfected HEK 293F or HEK 293T cells were collected by centrifugation at 400 g for 5 min. Cells were washed in PBS and resuspended at 1/10 of the original culture volume in buffer W (100 mM Tris-HCl, pH 8.0, 150 mM NaCl supplemented with 1/500 of EDTA-free protease inhibitor cocktail Set III (Calbiochem) and 0.1% (v/v) of Nonidet-P40 detergent (Sigma). Cells were then subjected to three rounds of freeze-thaw lysis and cleared by centrifugation at 13,000 g for 5 min. The protein concentration of lysates was measured by the method of Bradford using bovine serum albumin (BSA, Sigma) as standard (Bradford, 1976).

2.3.4. Detection of posttranslational modifications of tubulins

The detection of posttranslational modifications of tubulins was performed by Western blot as follow, equivalent amounts each of lysate or fraction were separated by SDS-PAGE 10% as described in section 2.1.3. Proteins were transferred onto PVDF filters (Immobilon-P, Millipore) and Western blot was performed as described in section 2.1.6. Primary antibodies

used were: rabbit anti-Δ2 tubulin 1:500 (Millipore), rabbit anti-polyglutamylated tubulin 1:5,000 (polyE long chain, gift of Dr. M.A. Gorovsky, Departament of Biology, University of Rochester, UK), mouse anti-HA 1:5,000 (clone HA-7, Sigma), rat anti-tyrosine tubulin 1:2,000 (clone YL1/2, Millipore), mouse anti-αtubulin 1:3,000 (Sigma), or mouse anti-β-actin 1:50,000 (Sigma) antibodies. After washing with PBS-T, the filters were incubated with HRP-labeled donkey anti-rabbit 1:5,000 (GE Healthcare), goat anti-mouse 1:5,000 (Biorad) or goat anti-rat 1:5,000 (GE Healthcare) for 1 h. After four rinses, immunoreactive bands were visualized with Luminata Forte Western HRP substrate (Millipore) in a Molecular Imager VersadocTM (BioRad).

2.3.5. Immunocytochemistry

For immunofluorescence, HEK 293T cells were seeded onto sterile glass coverslips in 6-well plates (Nunc), incubated overnight and transfected for 48 h either with the *hccp6*-pTriEx6 construct or empty pTriEx6 as a control. Cells were washed once with pre-warmed PBS and then fixed with cold methanol for 15 min at 20 °C. Fixed cells were rehydrated in PBS-T and incubated 8 min with cold PBS containing 0.1% Triton X-100 (v/v) (Roche) at 4 °C for permeabilization. Samples were blocked using 5% BSA (Sigma) in PBS-T for 2 h and then rinsed once with PBS-T before overnight incubation at 4 °C with primary antibodies. Primary antibodies used were mouse anti-HA 1:1,000, rabbit anti-Δ2-tubulin 1:100 and rabbit anti-polyglutamylated tubulin 1:500 (polyE long chain). Cells were washed three times with PBS-T and incubated 2 h at RT with secondary antibodies: Alexa Fluor 647 goat anti-rabbit 1:500 (Invitrogen), Alexa Fluor 488 goat anti-mouse 1:500 (Invitrogen) and 2,4-diamidino-2-phenylindole 1:500 (DAPI; Sigma) in PBS-T and 2% BSA. After rinsing with PBS-T and PBS, the coverslips were mounted onto slide glasses with Fluoprep (BioMérieux). Immunofluorescence images were acquired at room temperature on a Leica TCS SP5 confocal microscope using an X63 (1.4) oil-immersion objective (Leica Microsystems, Wetzlar, Germany). Images were edited in Adobe Photoshop CS 8.0.1.

2.3.6. *In vitro* substrate validation

The proteins hCCP6, HMGB3 and TRAFD1 were affinity purified with a Strep-Tactin affinity column according to the protocol of the manufacturer (IBA GmbH, Göttingen, Germany). Cleared lysates of HEK 293F cells overexpressing TRAFD1, HMGB3 or hCCP6 were

prepared as indicated above. Strep-Tactin column (IBA GmbH), equilibrated with buffer W, was loaded with the clear lysate and the column was washed with 5 column volumes (CV) of Buffer W. Strep-tagged proteins were eluted and fractionated using buffer E (buffer W supplemented with 2.5 mM desthiobiotin). The fractions with the highest amount of protein, determined by SDS-PAGE, were selected for immunoblot and *in vitro* enzymatic assay. The reaction mixture for the *in vitro* enzymatic assay contained 20 µl of hCCP6 fraction and was incubated overnight at 37 °C with affinity purified fractions of TRAFD1 (20 µl) or HMGB3 (10 µl) in the activity buffer (80 mM PIPES pH 6.8, 1 mM MgCl2) supplemented with 1/500 EDTA-free protease inhibitor cocktail Set III (Calbiochem). The same reactions were performed in the presence of the carboxypeptidase inhibitor ortho-phenanthroline 10 mM (Sigma). An aliquot of the reaction mixtures was separated by SDS-PAGE 10 % and transferred onto PVDF filter for Western blot as described in section 2.1.6. Primary antibodies used were rabbit anti-polyglutamylated tubulin 1:5,000 for reactions with TRAFD1 and mouse anti-HA 1:5,000 for reactions with HMGB3.

2.3.7. Inhibition Tests

The interaction of hCCP6 with CP inhibitors was assayed *in vitro* using purified porcine brain tubulin (Cytoskeleton, Inc) as substrate. Reactions took place in activity buffer (80mM PIPES pH6.8, 1mM MgCl2) supplemented with 1/500 EDTA-free protease inhibitor cocktail Set III (Calbiochem). All reaction mixtures contained 0.625 µg of purified tubulin, 5 µl of hCCP6, purified by Strep-tag affinity chromatography, and one of the following inhibitors: 10 mM EDTA (Sigma), 10 mM ortho-phenantroline (Sigma), 0.01 mM latexin, 0.1 mM TCI (Tick Carboxypeptidase Inhibitor), 0.01 mM ACI (Ascaris Carboxypeptidase Inhibitor), 0.033 mM LCI (Leech Carboxypeptidase Inhibitor), 0.1 mM NvCI (*Nerita versicolor* Carboxypeptidase Inhibitor) and 0.1 mM PCI (Potato Carboxypeptidase Inhibitor). Control reactions (no inhibitor), with or without hCCP6 were carried out in parallel. All reactions were first incubated on ice, 30 min before adding the substrate, *i.e.* tubulin. Reactions were incubated for 5 h at 37 °C, stopped by addition of SDS PAGE sample loading buffer 5X and 5 min boiling, followed by separation on SDS-PAGE 10%. The proteins were then transferred onto a PVDF filter for Western blot, as described in section 2.1.6, and detected using the anti-polyglutamylated tubulin (polyE long chain). The recombinant inhibitors TCI, ACI, LCI, NvCI,

PCI and latexin were produced in the Protein Engineering and Proteomics laboratory, IBB. All reactions were performed in duplicate.

2.3.8. Interactomic assays using magnetic beads

The search for interacting partners of hCCP6 was performed by immunoprecipitation using Dynabeads® Protein G (Protein G-beads, Invitrogen) following the manufacturer's instructions. The antibody anti-HA (Sigma) was used to capture overexpressed HA-tagged hCCP6 in HEK 293F cells. Soluble extracts of HEK 293F cells, transfected with the *hccp6-pTriEx6* construct or the empty vector, were prepared as indicated above (preparation of cleared lysates). The concentration of the resulting protein extracts was estimated by measuring the absorbance at 280 nm and adjusted to 10 mg/ml. Two methods**,** direct (Figure 2.2A) and indirect (Figure 2.2B), were used for immunoprecipitation experiments. All buffer changes and washing steps were performed in a magnetic rack.

2.3.8.1. Direct method

Protein G-beads (50 µl per sample) were mixed with the anti-HA antibody (2 µg per sample) resuspended in PBS supplemented with 0.01% Tween 20 (v/v) (PBS-T 0.01%) and incubated for 10 min at RT with tilting/rotation. Beads-antibody complexes were washed with PBS-t 0.01% and the antibodies cross-linked to the beads with 5mM of bis(sulfosuccinimidyl)suberate (BS$_3$, Thermo Scientific, Pierce) following the manufacturer´s instructions. The coated beads were washed with PBS-T 0.01% and incubated overnight at 4 °C with tilting/rotation, with 200 µl of soluble extracts of HEK 293F cells overexpressing hCCP6 or transfected with the empty vector. Samples were washed 5 times with 200 µl PBS-T 0.01% and resuspended in 100 µl of PBS before elution. After removing PBS solution, proteins were eluted from beads with 20 µl of 1% formic acid and samples were prepared for nano-ESI-MS/MS analysis.
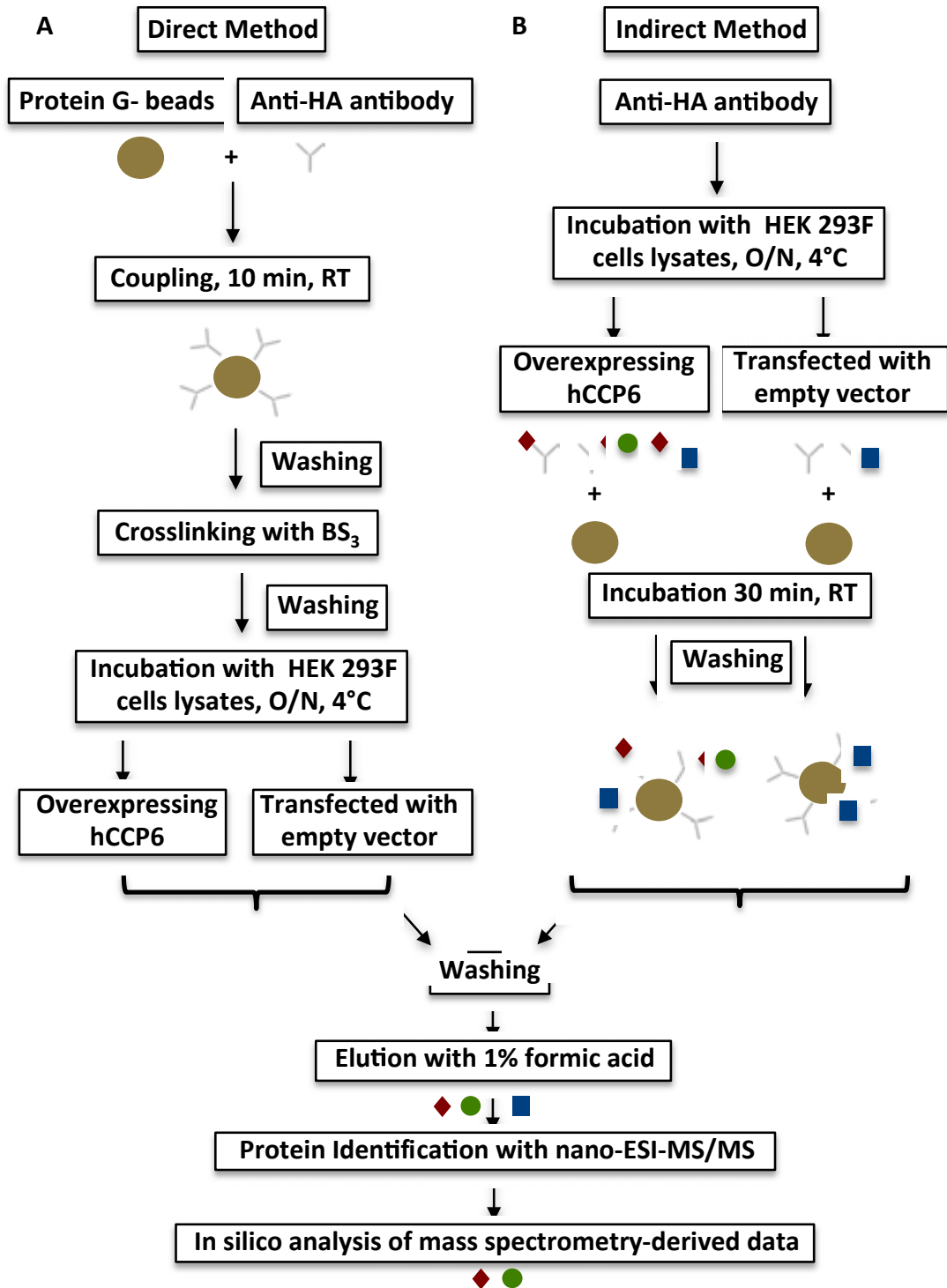
Figure 2.2. Workflow for the isolation and identification of interacting partners of hCCP6 by the direct (A) or indirect (B) method using magnetic beads and nanoESI-MS/MS. O/N: overnight; BS₃: bis(sulfosuccinimidyl)suberate; RT: room temperature.

2.3.8.2. Indirect method

In the indirect method, the anti-HA-tag antibody (3μg per sample) was first incubated overnight, at 4 °C, with tilting/rotation, with 500 μl of HEK 293F cellular extracts,

overexpressing hCCP6 or transfected with the empty vector. Then, protein G-beads (50 μl per sample) were added to mixture and incubated for 30 min, at RT, with tilting/rotation. After the coupling, the beads were washed 5 times with PBS-T 0.01% and finally resuspended in 100 μl of PBS. Proteins were eluted from beads using 20 μl of formic acid 1% followed by identification with nano-ESI-MS/MS.

2.3.8.3. Protein identification with nano-Esi-MS/MS

Protein identifications with nano-Esi-MS/MS were performed as described in the section 2.2.13.4.

3. Results and Discussion

3.1. Chapter 1. Structure of the cytosolic carboxypeptidase of *Pseudomonas aeruginosa*, a model for mammalian cytosolic carboxypeptidases.

The group of Protein Engineering and Proteomic, IBB has extensive experience in the study of M14A/B metallocarboxypeptidases (MCPs). In recent years, our group has participated in the initial description and characterization of the novel M14D subfamily members (Rodriguez de la Vega et al., 2007; Rodriguez de la Vega Otazo et al., 2012; Tanco et al., 2014, submitted). Attempts made to clone, overexpress, and analyze structurally and functionally CCPs from different species (i.e., human) clashed with the recurrent difficulty of getting these forms or their functional domains soluble and in sufficient quantities (Rodriguez de la Vega et al., 2007). For this reason, a prokaryotic CCP, in this case from *Pseudomonas aeruginosa*, was selected for structural and functional characterization with the aim to provide a model for the subfamily. Bacterial CCPs are shorter, have simpler domain architectures and are expected to produce best in the *E.coli* expression system due to evolutionary proximity.

3.1.1. Recombinant protein production of PaCCP

The PA2831 gene was identified previously (Rodriguez de la Vega et al., 2007) as a gene encoding an M14D MCP from *P. aeruginosa* MPAO1. Its protein product is henceforth referred to as PaCCP here. The protein contains 375 residues and has been overexpressed in *E. coli* with a strep-tag at the N-terminus for purification and detection (Otero et al., 2012). PaCCP was brought to homogeneity in a two-step purification protocol that includes affinity and gel filtration chromatography. The theoretical molecular weight of PaCCP is 44.3 kDa (based in its gene structure) and the protein eluted from the size exclusion chromatography showed an apparent molecular mass of 43.7 kDa as calculated with a previously determined calibration curve. Buffers with a pH range from 5.5 to 8.0 or higher ionic strength (0.5 M NaCl) did not change the elution pattern. These results together with the mobility of the purified protein in non-denaturing PAGE (Figure 3.1) allowed us to conclude that the protein

behaves as a monomer in solution. Nevertheless, other members of the bacterial subfamily of CCPs form oligomers in solution. The assembly observed in the crystal structures of the CCP of *Burkholderia mallei* ATCC 23344 (BmCCP; PDB entry: 3k2k) and the CCP of *Shewanella denitrificans* OS-217 (SdCCP; PDB entry 3l2n) (Joint Center for Structural Genomics) suggests that these proteins can be in tetrameric form. Interestingly, it has been demonstrated, in our lab, that BmCCP behaves as a homo-tetramer in solution by gel filtration chromatography experiments and visualization in negative staining electron microscopy (unreported results). The same conclusion was achieved by Rimsa et al. (2014) for the CCP of *Burkholderia cenocepacia* J2315 (BcCCP) whose structure was recently determined (PDB entry 4b6z).
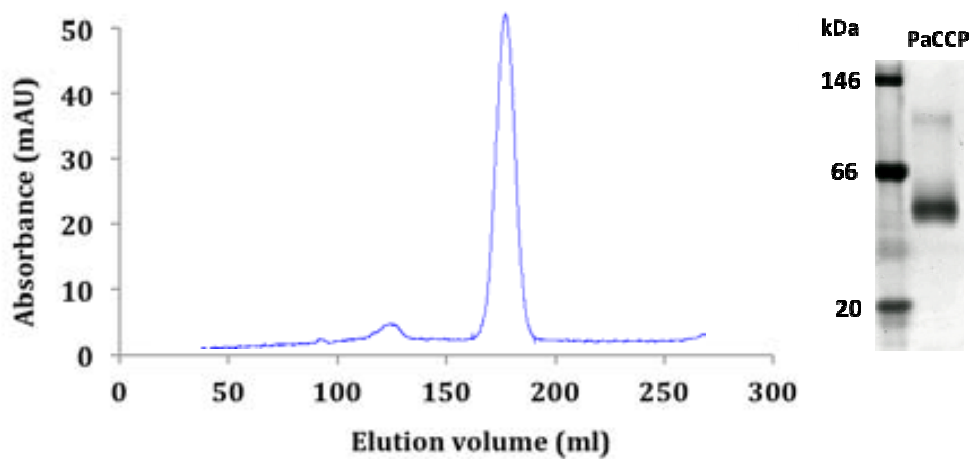
Figure 3.1. PaCCP behaves as a monomer in solution. A gel-filtration chromatogram for PaCCP is displayed on the left. The protein eluted with an apparent molecular mass of 43.7 kDa. Native PAGE analysis of purified PaCCP is shown on the right with molecular-mass standards on the left.

3.1.2. Overall structure of PaCCP

The crystal structure of PaCCP was determined by crystallization and X-ray diffraction at high resolution (1.6 Å). The asymmetric unit of the crystal contains 2 polypeptide chains. The two chains are almost identical, showing a root mean square (RMS) deviation of 0.21 Å for 375 aligned residues but, taking into account previous gel filtration experiments, there is no indication that they form a biological unit. The crystal structure of each polypeptidic chain contains 2 distinct continuous domains, a novel β-rich N-domain (Met1 to Glu111), which creates an extended interface that interacts with a canonical CP domain (Pro112 to Arg375), containing the catalytic Zn^{2+} and the potential substrate binding residues (Figure 3.2). The structure does not contain any disulfide bridge, in contrast to the other MCP

members from M14A and M14B subfamilies, which are normally secreted outside the cell and contain several disulfide bridges (Arolas et al., 2007).
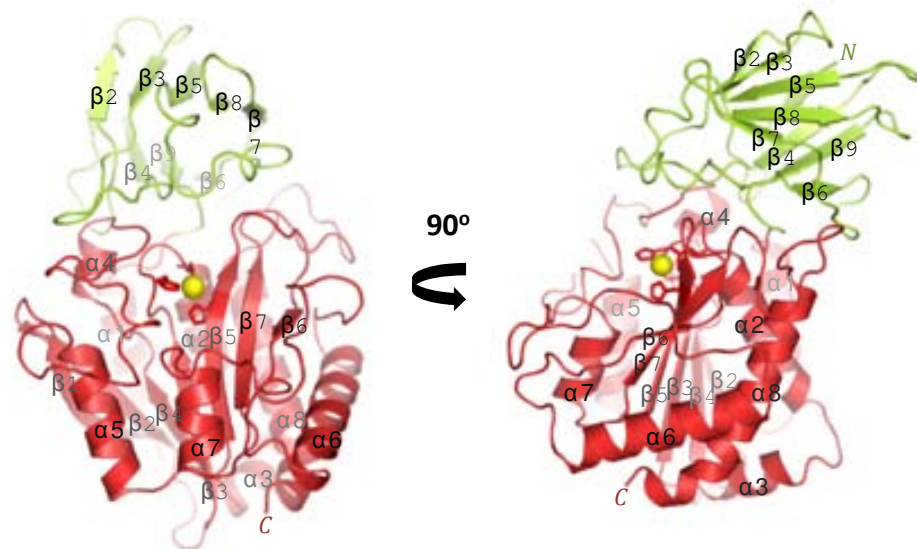
Figure 3.2. Crystal structure of the PaCCP. Two views of the PaCCP, shown in ribbon representation. Secondary structure elements are numbered with either α (helixes) or β (sheets). The N-domain is in green; the CP domain is in red. The zinc binding residues are depicted in stick representation near the zinc atom, represented as a yellow sphere. The figure was prepared with PyMOL.

The overall structure of PaCCP is similar to the derived data from the structures of proteobacterial SdCCP, BmCCP and BcCCP. These proteins are clustered into the M14D1b clade of the M14D subfamily (Otero et al., 2012; Rodriguez de la Vega et al., 2007). Structural comparison displays a global RMS deviation of 1.1 Å for 374 aligned residues for 3l2n, 1.3 Å for 370 aligned residues for 3k2k and 1.03 Å for 364 aligned residues for BcCCP. All proteins have ~40% sequence identity with respect to PaCCP. On the other hand, the PaCCP sequence shows ~21 to 25% identity with all 6 members of the human CCP subfamily (CCP1-6; Figure 3.3). Human CCPs display a variety of length sizes, ranging from 500 residues for CCP6 to 1226 residues for CCP1. Despite this variety, however, all members contain the N- and CP-domains, the zinc binding motifs and the catalytic or substrate binding residues (Figure 3.3).

```
        2    D.S.N..                    20        R..D..L..A..        H.. W..40
PaCCP   -QIRADFDSGNIQVIDASD----------------------PRRIRLAIRPDLAS-----QHFQWFHFKVEGMAPATEHCFTLVN
BmCCP   LSITSNFDAGAIDVVSCER----------------------ADAIRLRVRGDNRS-----EFAQWFYFRLTGARG-ERCVMTFEN
BoCCP   LSITSNFDAGAIDVVSCER----------------------PDAIRLRVRGDNRS-----EFAQWFYFRLTGARG-ERCVMTFEN
SdCCP   -RISANFDGGNIETISLAN----------------------PDDIQLAIRPDAGG-----EFYQWFNFRFEATIG-KTYTLNILN
hCCP1   LKFNSKFESGNLRKVIQIR----------------------KNEYDLILNSDINSN----HYHQWFYFEVSGMRPGVAYRFNIIN
hCCP2   LLFESRFESGNLQKAVRVD----------------------TYEYELTLRTDLYTN----KHTQWFYFRVQNTRKDATYRFTIVN
hCCP3   LMFEARFESGNLQKVVKVA----------------------EYKYQLTVRPDLFTN----KHTQWYYFQVTNMRAGIVYRFTIVN
hCCP4   LRFFSKFESGNLRKAIQVR----------------------EFEYDLLVNADVNST----QHQQWFYFKVSGMQAAIPYHFNIIN
hCCP5   LLFSSRFDSGNLAHVEKVESLSSDGEGVGGGASALTSGIASSPDYEFNVWTRPDCAETEFENGNRSWFYFSVRGGMPGKLIKINIMN
hCCP6   LIFDACFESGNLGRVDQVS----------------------EFEYDLFIRPDTCNP----RFRVWFNFTVENVKESQRVIFNIVN

        60   Y..H..A..Y..            80W..R..                              100      Y..F..
PaCCP   AGQ--SAYSHAWSGYQAVAS-----YDGERWFRVPSQY----------------------DADGLHFQLE-PEESEVRFAYFEPYSR
BmCCP   AND--CAYPAGWRDYRAVAS-----YDRVNWFRVPTSY----------------------DGQMLTIDHT-PEFDSIHYAYFEPYSE
BoCCP   AAE--CAYPSGWRNYRAVAS-----YDRVDWFRVPTTF----------------------DGRTMTIDHT-PEFDSIYYAYFEPYSE
SdCCP   AGG--ASYLKGWEDYQAVAS-----YDRQTWFRLPTEY----------------------KDGKLSISVE-LDCEAIQIAYFTPYSY
hCCP1   CEKSNSQFNYGM--QPLMYSVQEALNARPWWIRMGTDICYYKNHFSRSSVAAGGQKGKSYYTITFTVNFP-HKDDVCYFAYHYPYTY
hCCP2   LLKPKSLYTVGM--KPLLYSQLDANTRNIGWRREGNEIKYYKNNT-------DDGQQPFYCLTWTIQFP-YDQDTCFFAHFYPYTY
hCCP3   FTKPASLYSRGM--RPLFYSEKEAKANHIGWQRIGDQIKYYRNNP--------GQDGRHYFSLTWTFQFP-HNKDTCYFAHCYPYTY
hCCP4   CEKPNSQFNYGM--QPTLYSVKEALLGKPTWIRTGHEICYYKNHYRQSTAVAGGASGKCYYTLTFAVTFP-HSEDVCYLAYHYPYTY
hCCP5   MNKQSKLYSQGM--APFVRT---LPTRPPWWERIRDRPTFEM-------------TETQFVLSFVHRFVEGRGATTFFAFCYPFSY
hCCP6   FSKTKSLYRDGM--APMVKS-----TSRPFKWQRLPPKNVYYYRCP---------DHRKNYVMSFAFCFD-REEDIYQFAYCYPYTY

        120                          140                                             160
PaCCP   ERHARLVERALGIEG---------------VERLAVGTSVQGRDIELLRVRRHP-----------------------DSHLKLWVI
BmCCP   ERHSEFLGAVQQMPQ---------------ASVVELGRTVEGRPMSLVVLGT-PDEA-----------------------GAAKKKVWII
BoCCP   ERHAAFLGAVQQLPQ---------------ASVVELGRTVEGRPMSLLTLGT-PETD-----------------------DAPKKKVWII
SdCCP   ERHLDLISAVQLHPL---------------VSTEHLGLTLDGRDMTLVKVGDDD----------------------PSKKSIWIT
hCCP1   STLQMHLQKLESAHNP---------QQIYFRKDVLCETLSGNSCPLVTITAMPESNYYE--------------HICHFRNRPYVFLS
hCCP2   TDLQCYLLSVANNPI----------QSQFCKLQTLCRSLAGNTVYLLTIT-NPSQTP-----------------QEAAAKKAVVLS
hCCP3   TNLQEYLSGINNDPVR---------SKFCKIRVLCHTLARNMVYILTIT-TPLKN------------------SDSRKRKAVILT
hCCP4   SDCQELLDILEKSVNL---------KEVYFRQDVLCQTLGGNPCPLVTITAMPESNSDE--------------HLEQFRHRPYQVIT
hCCP5   TALMTHLNQLDQRFPENHPTHSSPLDTIYYHRELLCYSLDGLRVDLLTITSCHGLREDREPRLEQLFPDTSTPRPFRFAGKRIFFLS
hCCP6   TRFQHYLDSLQKRNMD-----------YFFREQLGQSVQQRKLDLLTIT-SPDNLR-------------------EGAEQKVVFIT

        H..X..X..E..X.. 180             200 N..P..D..G..  R..      N..R..      240
PaCCP   AQQNHPGEBMAEWFMEGLIERLQR----PDDTEMQRLLEKADLYLVPNMNPDGAFHGNLRTNAAGQDLNRAWLEPSAERSPEVWFVQQ
BmCCP   ARQNHPGESMAEWFIEGLVKRLVGWGDWSGDPVARKLYDHATFYIVPNMNPDGSVHGNLRTNAAGANLNREWMEPDAERSPEVLVVRD
BoCCP   ARQNHPGESMAEWFVEGLVKRLAGWGDWAGDPVARKLYDRVTFYIVPNMNPDGSVHGNLRTNAAGANLNREWMAPDAERSPEVLAVRD
SdCCP   ARQNHPGETMAEWLVEGLLNQLLD----NDCPTSKALLDKANFYIVPNMNPDGSVRGHLRTNAVGANLNREWQTPSLERSPEVYYVVN
hCCP1   ARVNHPGETNASWVMKGTLEYLMS-----NNPTAQSLRESYIFKIVPMLNPDGVINGNHRCSLSGEDLNRQWQSPSPDLHPTIYHAKG
hCCP2   ARVNHPGESNGSWVMKGFLDFILS-----NSPDAQLLRDIFVFKVLPMLNPDGVIVGNYRCSLAGRDLNRHYKTILKESFPCIWYTRN
hCCP3   ARVNHPGETNSSWVMKGFLDYILG-----NSSDAQLLRDTFVFKVVPMLNPDGVIVGNYRCSLAGRDLNRNYTSLLKESFPSVWYTRN
hCCP4   ARVNHPGESNASWVMKGTLEFLVS-----SDPVARLLRENFIFKIIPMLNPDGVINGNHRCSLSGEDLNRQWLSPSAHLQPTIYHAKG
hCCP5   SRVNHPGETPSSFVFNGFLDFILRP----DDPRAQTLRRLFVFKLIPMLNPDGVVRGHYRTDSRGVNLNRQYLKPDAVLNRPAIYGAKA
hCCP6   GRVNHPGETPSSFVCQGIIDDFLVS-----QNHPIACVLREYLVFKIAPMLNPDGVYLGNYRCSLMGFDLNRHWLDPSPWVHPTLHGVKQ

        X..H..X..X..X..X..                280                        300
PaCCP   EMKR------------HGVDLFLDIHGDEEIPHVFAAGCEGNPGY----------TPPRLERLEQRF-REELMARGEFQIRHGYPR
BmCCP   AIHA------------IGCDLFFDIHGDEDLPYVFAAGSEMLPGF----------TEQQRVEQSAFIDSFKRASPDFQDEHGYPP
BoCCP   AIHA------------IGCDMFFDIHGDEDLPYVFVAGSEMLPSF----------TEQQGKEQAAFIDAFKVASPDFQTEHGYAA
SdCCP   KMHE------------TGVDLFYDVHGDEGLPYVFLAGCEGIPNY----------SDKLASLQQDFVAALSLASADFQTEFGYDK
hCCP1   LLQYLAAV--------KRLPLVYCDYHGHSRKKNVFMYGCSIKETVWNTNDNATSCDVVEDTGYRTLPKILSHIAP----AFCMSS
hCCP2   MIKRLLE---------EREVLLYCDFHGHSRKNNIFLYGCNNNNRK----------YWLHERVFPLMLCKNAPD----KFSFHS
hCCP3   MVNHRLME--------KREVILYCDLHGHSRKENIFMYGCDGSDRSKT---------LYLQQRIFPLMLSKNCPD----KFSFSA
hCCP4   LLYHLSSI--------GRSPVVFCDFHGHSQKKNVFLYGCSIKETLWQAACTVGTSTILEEVNYRTLPKILDKLAP----AFTMSS
hCCP5   VLLYHHVH---//---KESGVAYYVDLHGHASKRGCFMYGNSFSDES------------TQVENMLYPKLISLNSA----HFDFQG
hCCP6   LIVQMYND--------PKTSLEFYIDIHAHSTMMNGFMYGNIFEDEE-----------RFQRQAIFPKLLCQNAE----DFSYSS

        I..                 E..   340                               360
PaCCP   SAPGQA-----------NLALACNFVGQT-YDCL-AFTIEMPFKDHDDNPEP---------------------GTGWSGARSKRL
BmCCP   GKYRED-----------AFKLASKYIGHR-FGCL-SLTLEMPFKDNANLPDE----------------------HIGWNGARSASL
BoCCP   SKYKED-----------ALKLASKYIGHQ-FGCL-SLTLEMPFKDNANLPDE----------------------RVGWNGERSAAL
SdCCP   DEPGKA-----------NLTVACNWVANT-FKCL-SNTLEMPFKDNANLADP----------------------FQGWSPERSVYF
hCCP1   CSFVVE-----------KSKESTAR-VVVWREIGVQRSYTMESTLCGCDQ-GKYK----------------------GLQIGTRELEEM
hCCP2   CNFKVQ-----------KCKEGTGR-VVMWR-MGILNSYTMESTFGGSTL-GNKR----------------------DTHFTIEDLKSL
hCCP3   CKFNVQ-----------KSKEGTGR-VVMWK-MGIRNSFTMEATFCGSTL-GNKR----------------------GTHFSTKDLESM
hCCP4   CSFLVE-----------KSRASTAR-VVVWREMGVSRSYTMESSYCGCNQ-GPYQCTQRLLERTKNERAHPVDGLQGLQFGTRELEEM
hCCP5   CNFSEKNMYARDRRDGQSKEGSGR-VAIYKASGIIHSYTLECNYNTGRSVNSIPAAC----HDNGRASPPPPPAFPSRYTVELFEQV
hCCP6   TSFNRD-----------AVKAGTGRRFLGGLLDHTSYCYTLEVSFYSYIISGTTA-----------------------AVPYTEKAYMKL

        375
PaCCP   GQDVLSTLAVLVDELR-
BmCCP   GAAMLGAILEHVRAFA-
BoCCP   GAAMLAAILVNVDTFA-
SdCCP   GEASLIAMRAVIDKIGQ
hCCP1   GAKFCVGLLRLKRLTSP
hCCP2   GYHVCDTLLDFCDPDQM
hCCP3   GYHFCDSLLDYCDPDRT
hCCP4   GAMFCLGLLILELKSAS
hCCP5   GRAMAIAALDMAECNPW
hCCP6   GRNVARTFLDYYRLNPV
```

Figure 3.3. Sequential alignment of bacterial and human CCPs. Multiple alignment of *Pseudomonas aeruginosa* **CCP (PaCCP),** *Burkholderia mallei* **CCP (BmCCP),** *Burkholderia* *cenocepacia* **CCP (BcCCP) and** *Shewanella dentrificans* **CCP (SdCCP) and human CCPs (CCP1-6). Gaps are denoted by dashes. Numbering above the sequence alignment corresponds to the amino acid position in PaCCP. The residues of PaCCP in direct contact with the CP domain are indicated in green. The zinc binding motifs ($H_{69}X_{70}X_{71}E_{72}X_{73}$), ($X_{195}H_{196}X_{197}X_{198}X_{199}X_{200}$), NPDG signature ($N_{112}P_{113}D_{114}G_{115}$) and the catalytic or substrate binding residues (N_{144}, R_{145}, R_{127}, E_{270}) are indicated (bovine CPA1 numbering). Under the alignment is shown the scheme of conserved secondary structures (in green N-terminal and red CP domain) in the M14D subfamily. α-helixes are represented as boxes and β-Strands as arrows.**

3.1.3. Structure of the N-domain

The novel N-domain of 111 residues of PaCCP (Met1 to Glu111) essentially folds a β-sandwich structure, essentially formed by 9 strands (named β1–β9 in Figure 3.2) arranged in two opposite mixed β-sheets. A first β-sheet, formed by 4 antiparallel strands (β1-β4-β9-β6), is stacked against a second β-sheet, formed by 5 mixed strands (β2-β3-β5-β8-β7), sandwiching a hydrophobic core in the center of this domain (Figure 3.2).

The N-domain constitutes a novel structure feature among the M14 MCP family members, typical of the novel M14D subfamily. Although it is also located N-terminally to the CP domain, sequential and structural comparative analyses show that this domain does not display any homology with the inhibitory pro-segment typical of the M14A zymogen subfamily (Arolas et al., 2007). The N-domain is also different from the β-barrel transthyretin-like domain located at the C-terminus of the mammalian M14B subfamily (Arolas et al., 2007). The top structural homologues found are the other members of the same bacterial CCP subfamily (PDB codes 3k2k and 3l2n, both with an RMS deviation of 1.0Å and PDB code 4b6z with an RMS deviation of 0.89 Å for 109 aligned residues with 40% sequence identity).

The N-domain resembles, due to topology and strand connectivity, the C-terminal domain of heparinase II (PDB 3e80), a depolymerizing enzyme of heparin and heparan sulfate (Shaya et al., 2010). Structural comparative servers display an RMS deviation of 2.6 Å for 76 aligned residues with 14% sequence identity. Such a C-terminal domain of heparinase II is involved in the homodimer formation necessary for the arrangement of the biological unit of this enzyme. Although gel-filtration experiments indicate that the biological unit of PaCCP is a monomer, it cannot be discarded the potential contribution of this domain in protein–protein interactions, or hindering/modulation actions with substrates of the

enzyme.

Noteworthy, the N-domain of PaCCP buries a quite extended region of the CP domain, basically by establishing polar, hydrophobic and Van der Waals interactions with the upper face of the CP domain, next to the active site groove. An overall interface of 2490 Å^2 is buried between the two domains. A total of 17 residues from the N-domain are in direct contact with the CP domain (Figure 3.3). Of particular interest is the hydrophobic core created by Trp38, Trp82, and Phe110, located in the center of the interface between both domains and establishing many contacts with residues of the CP domain. Besides this hydrophobic interface, Arg28, Asp30, and Arg84 also establish hydrogen bonds with the main chain of the CP domain. Interestingly, all these residues are highly conserved in all members of the bacterial and human CCP subfamily, highlighting the structural similarities between such evolutionary distant members (Figure 3.3).

As mentioned above, the N-domain is located in a different orientation compared to the inhibitory pro-segment domain of the M14A subfamily (Figure 3.4). The active site groove of PaCCP is more accessible than in the M14A zymogens although is not fully free. The orientation of the N-domain with respect to the CP moiety partially blocks the entrance of substrates to the catalytic cleft. In particular, the side chain of Tyr64 of the N-domain, which is conserved in human and bacterial CCP subfamilies (Figure 3.3), directly clashes with the P2 subsite in a canonical orientation for the entrance of substrates into the active-site of MCPs (Figure 3.4).

The regulatory properties of the human CCP subfamily might reside in the novel interaction of the N-domain with the conserved CP domain observed in the crystal structure of PaCCP. In the case of CPs of the subfamily M14A the function is regulated by enzymatic activation. The pro-region that covers the active-site is removed, mainly by trypsin, and starts with a cleavage between the activation segment and the CP moiety (Arolas et al., 2007). In this regard, no connection loop exists between the N- and CP domains of PaCCP that can be accessible for endopeptidases. PaCCP was subjected to limited digestion with trypsin, proteinase K, and subtilisin and it was only detected a discrete cleavage with trypsin that affected the integrity of the CP domain. Nevertheless, it cannot be discarded the possibility of activation of CCPs by the proteolytic hydrolysis by another endopeptidase.
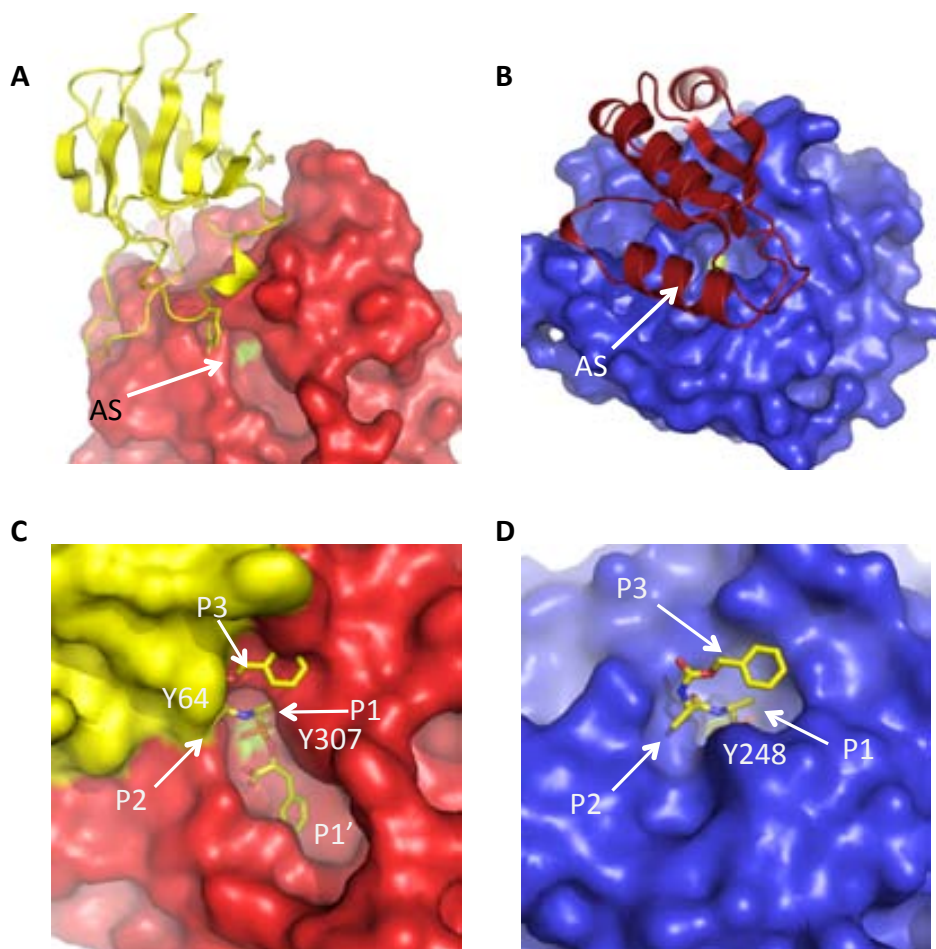
Figure 3.4. Analysis of the N-domain of PaCCP. A. Ribbon representation of the N-domain of PaCCP (yellow) and surface representation of the CP domain of the PaCCP (red). Tyr64 is shown in stick representation. B. Ribbon representation of the pro domain of human PCPA2 (red, PDB: 1aye) and surface representation of the CP domain of the PCPA2 (blue). Arrows indicate the location of the active site (AS) groove (A, B). C. Surface representation of the active site groove of PaCCP modeled with the inhibitor Phosphonate ((((N-phenylmethoxycarbonyl-L-alanyl) amino)ethyl)hydroxyphosphono)-L-benzylacetic acid). N-terminal and carboxypeptidase domain are colored yellow and red, respectively. Phosphonate is shown in stick representation (yellow). Residues in contact with phosphonate and active site subsite are labeled. D. Surface representation of the active site groove of bovine CPA (PDB 6CPA) in complex with the phosphonate inhibitor. Graphics prepared with PYMOL.

The activity of PaCCP might also be regulated by a structural rearrangement of the molecule upon binding of the biological substrate or by an allosteric molecule. This would alleviate the steric contacts with the P1, P3, and, particularly, P2 subsites and facilitate the correct positioning and cleavage of the C-terminal residue of the substrate. A potential role of the N-domain could be similar to the C-terminal β-sheet domain of the enzyme aspartoacylase (ASPA) (Bitto et al., 2007) from the AspA/AstE family of MCPs. This protein is formed by a catalytic N-terminal domain, with a protein fold similar to the catalytic domain of M14A MCPs, and a C-terminal domain positioned in a way that prevents productive

binding of polypetides in the active site. Thus, ASPA has a strict specificity for Nα-acetyl-L-aspartate (Bitto et al., 2007).

3.1.4. Structure of the CP-like domain

The spherical CP-like domain of PaCCP exhibits the classic α/β hydrolase fold typical of the M14 family of MCPs (Arolas et al., 2007). Its topology is conserved and is formed by a central 7-stranded β-sheet flanked by 8 α-helices, which basically run antiparallel to the sheet strands. The central sheet consists of a regular flat core of 4 parallel strands (β4- β3- β5- β7), containing the Zn^{2+} binding residues, flanked by 3 antiparallel strands (β1- β2- β6). In all other M14 members, another short strand is formed in contact with β6, between β6 and β7, but is absent in the structure of PaCCP. The CP domain is smaller in comparison to the other members of the M14 family (~270 residues vs. 300) due to the presence of a smaller loop between α4 and α5, located at the entrance of the active site (Figure 3.5). The hydrophobic core of PaCCP is basically composed by the side chains at both sides of the central β-sheet and by residues from the internal side of the flanking α-helices.

An optimal superposition of the bacterial CP domain with bovine CPA (a M14A subfamily representative) and human CPN (a M14B subfamily representative) exhibits a similar topology of secondary structure elements, despite the low sequence identity (Figure 3.5). Comparative structural analysis with PaCCP shows RMS deviations of 2.5 Å and 2.8 Å for 238 and 239 aligned residues with 18 and 20% sequence identity for both, respectively. The overlapping is strong in the conserved secondary structure elements, which shapes the active site groove and the hydrophobic core of the protein, but is poor in the connecting loops between helices and strands. In particular, connecting loops involved in the interface with the N-domain, namely β3-α2, α4-α5, β5-β6, and β7- α8, do not exhibit structural homology with the two major subfamilies of MCPs (M14A and M14B).
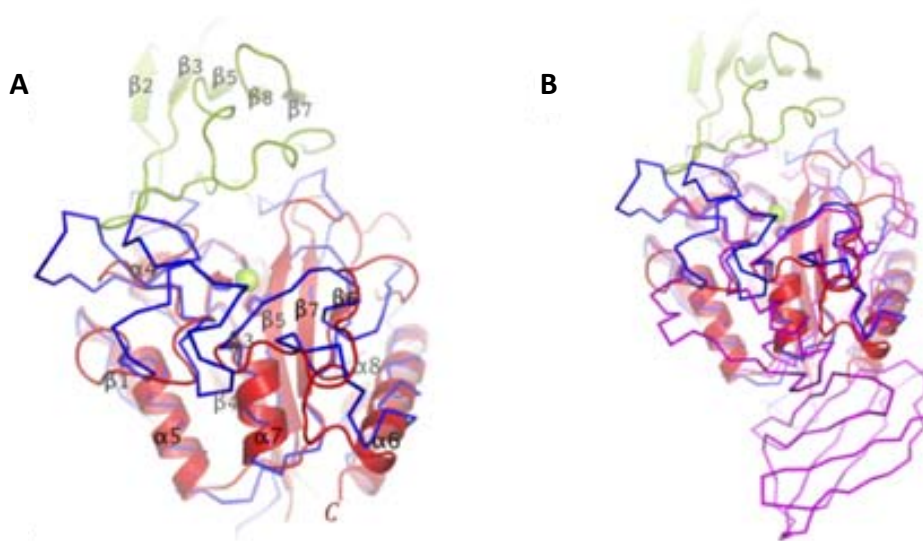
Figure 3.5. Structural comparison of the *P. aeruginosa* CCP with canonical carboxypeptidases. A. Superposition of PaCCP and bovine CPA (PDB 6CPA) structures in ribbon representation. PaCCP is colored red and green, for the N-terminal and CP domain, respectively, and bovine CPA is colored blue. Zinc binding residues are shown in stick representation. The zinc atom represented as a yellow sphere. B. Superposition of PaCCP with bovine CPA (PDB 6CPA) and human CPN (PDB 2NSM) in ribbon representation. CPA is colored blue and CPN is colored magenta. Graphics prepared with PYMOL.

The loop connecting α-helices α4 and α5 is minimal in length for the bacterial CCP and fundamentally contains the conserved basic patch formed by Arg218, Asn227, and Arg228 (the essential Arg127, Asn144, and Arg145 in the canonical M14A member, bovine CPA1, that represents the CP form after removal of the pre and pro regions). Despite the low structural similarity in this loop between these different types of CPs, the formation of the active site groove is not affected and the basic patch seems to be well structured and evolutionary conserved for the catalytic reaction of the enzyme.

Connecting loops between strand β6 and helix α6, and helixes α6 and α7, also display different conformations when compared to bovine CPA (M14A) and human CPN (M14B) (Figure 3.5). In PaCCP, these two loops are shorter, particularly in comparison to human CPN. The different conformation of the loop between helices α6 and α7 affects the location of the active site residue Tyr306 (equivalent to the Tyr248 in the M14A). This residue is ~2 Å displaced with respect to the canonical Tyr248 in the M14A and M14B families and could influence its catalytic properties.

The only extended loop in PaCCP is formed between strands β5 and β6, with 4 extra residues, and is located next to the active site groove and in contact with the N-domain.

Based on previous substrate-enzyme structures of MCPs (Kim and Lipscomb, 1990), this loop would clash with the P1 and P3 subsites of potential peptide substrates (Figure 3.4) and would require a structural rearrangement on substrate binding to allow CP activity.

3.1.5. Active-site groove

In PaCCP the active-site is located in the same location as in the other M14 members, essentially at the basis of a deep groove formed by the C-terminal ends of strands β3 and β5, and by the N-terminal end of helix α2. The wall of the groove is shaped with residues involved in two multiple turn loops, α4–α5 and α6–α7, which contribute with essential catalytic residues and with the formation of the S1′ subsite pocket to accommodate the side chain of the C-terminal residue of the substrate (P1′ subsite). In contrast to the other M14A and M14B members, no disulfide bridges that stabilize these multiple turn loops are found in the structure of the bacterial PaCCP, despite the presence of four free cysteine residues.

The catalytic zinc ion is located at the center of the active site groove, coordinated by the imidazole δ1 nitrogen atoms of His167 and His260, and by the two carboxylate oxygen atoms of Glu170 (His69, Glu72, and His196 in canonical M14A numbering) (Figure 3.6A). A fifth coordinating position for the zinc ion is occupied by water in the two molecules of the asymmetric unit, which would correspond to the position for the carbonyl oxygen of the scissile peptide bond during the catalytic reaction. The relative orientation of the catalytic residues in the active site is highly conserved with the other members of the M14 family. The basic patch formed by Arg218, Asn227, and Arg228 (Arg127, Asn144, and Arg145 in M14A numbering) are identically placed in the active site cleft of PaCCP (Figures 3.6A, B), as well as Glu335 (Glu270 in M14A), which directly participates in the catalytic reaction. The only active site residue located in a different position with respect to the canonical orientation is Tyr306 (the "flippy" Tyr248 in bovine CPA), which is displaced around 2 Å and would need structural rearrangement to be able to participate in a standard CP reaction.

In a first instance, the structural arrangement of the active-site residues in PaCCP seems optimal to carry out the expected C-terminal protease activity. Nevertheless, all attempts to detect enzymatic activity against different types of substrates, have failed. The enzymatic activity of PaCCP was evaluated against synthetic chromogenic substrates typically used for M14A and M14B MCPs. Hippuryl-L-X derivatives (Bz-Gly-X-COOH derivatives), furylacryloyl derivatives, and N-(4-methoxyphenylazo-formyl)-X derivatives

with different C-terminal residues (Phe, Lys, Arg and Glu) were tested unsuccessfully. Apparently, these are not good substrates for CCPs as CCP1 also failed to metabolize a wide variety of standard MCP substrates (Wu et al., 2012). Instead, CCP1 hydrolyzed synthetic substrates with two or more glutamates at the C-terminal indicating a preference for polyglutamates-containing substrates (Wu et al., 2012).

The activity of PaCCP against synthetic peptides that have been used for the characterization of substrate specificity of M14A MCPs (Lyons et al., 2008; Tanco et al., 2010) was also evaluated. It was not detected cleavage of peptides with aromatic, basic or acid C-terminal residues.
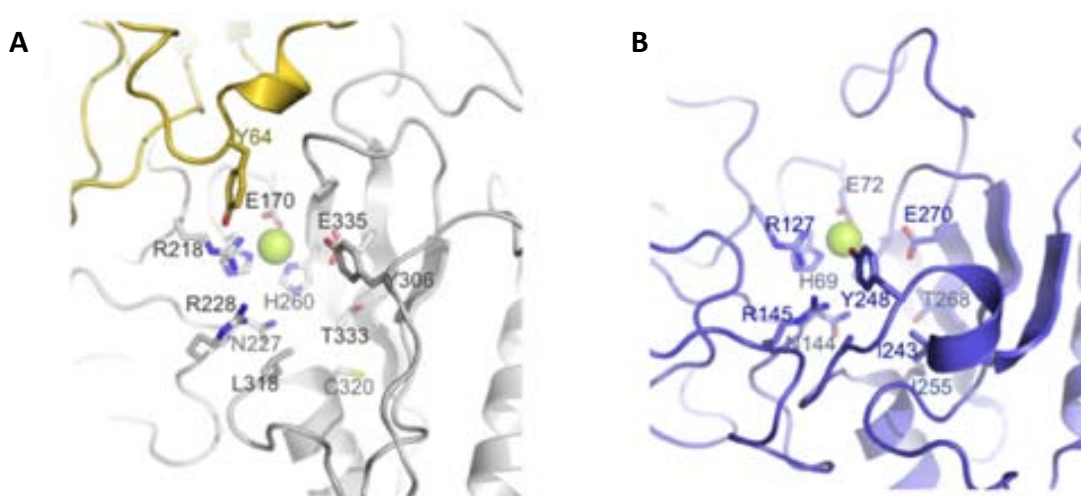
Figure 3.6. Active Site Groove of the *P. aureginosa* CCP. A. Ribbon representation of the active site groove of PaCCP. N-terminal domain is colored yellow, and the CP domain is colored grey. B. Ribbon representation of the active site groove of bovine CPA (PDB 6CPA). A and B: Residues involved in the catalytic reaction are labeled and shown in stick representation. Zinc atom is diplayed as a yellow sphere.

According to the PaCCP structure, enough space might allow small substrates to access the active-site. Nonetheless, the channel is partially occluded by the conserved Tyr64, which fills the S2 recognition site of classic CPs (Figure 3.4A,C) and could prevent the correct positioning of the substrate. To analyze the influence of Tyr64 in the accessibility to the active-site, dipeptide substrates with an acetyl group at the N-terminus (Ac-Gly-X series) were synthesized and assayed. The acetyl group should be positioned at the S2 recognition pocket of PaCCP when bound in a classical way. It is predicted that this N-terminal capping, having a much smaller size, could alleviate the potential clash with Tyr64. Such a series of dipeptides also failed to be cleaved by PaCCP. Another approach to reduce the steric

interference of Tyr64 was the production of the mutant PaCCP Tyr64Ala. The activity of the mutant was evaluated against chromogenic or peptide substrates but, once again, no cleavage was detected.

The apparent lack of activity of PaCCP could be influenced by another featured described above, which is the midway position of Tyr306 (Tyr248 in bCPA). This residue has been traditionally considered important for peptide recognition and catalysis in its "up-and-down" positioning (Kim and Lipscomb, 1990). The correct positioning of both Tyr64 and Tyr306 could depend on binding of the substrate(s) or of another molecular component that affects the conformation and environments of these residues. The control of PaCCP activity by binding to its substrate(s) and/or an allosteric molecule may be one of the mechanisms regulating its activity in the intracellular environment.

Another peculiar feature of PaCCP, potentially related to the difficulty of evaluating its functionality by substrates, is the occurrence of a free-cysteine residue in the position equivalent to Ile255 (bCPA), something very unusual in mammalian MCPs. According to the structural alignment in figure 3.3, SdCCP also has a Cys in this position while BmCCP and BcCCP have a Ser. The CPA from *Helicoverpa armigera* also has a Ser in the position equivalent to Ile255 (bCPA) resulting in a larger and more polar pocket and explains its broad substrate specificity (Estebanez-Perpina et al., 2001). According to this, it could be predicted that some bacterial CCPs have broad substrate specificity. Unlike bacterial CCPs, all human CCPs have an Arg in the position equivalent to Ile255 (Figure 3.3) which is the same case of the CPO, a M14A MCP member specific for acidic amino acids (Lyons and Fricker, 2011). In agreement with this, CCP1, 4, 5 and 6 have specificity for C-terminal glutamates (Rogowski et al., 2010) and also CCP2 and 3, according to results from our laboratory (Tort et al., unpublished results). In CCP1 it was demonstrated that Arg1078 (equivalent to Ile255) was essential for the activity *in vivo* (Wu et al., 2012), pointing out the importance of this residue for the catalysis.

The Cys320 in PaCCP, equivalent to Ile255 (bCPA), was substituted by Arg mimicking the human CCPs, with the aim of detecting enzymatic activity towards poly-glutamylated tubulin. Unfortunately, it was not detected activity evidencing that the single substitution of residues is not enough to render PaCCP active and reinforcing the idea of a structural rearrangement on substrate binding to allow CP activity.

3.1.6. Binding of CP inhibitors in the active site of PaCCP

Cleavage of any peptidic substrate by PaCCP have not been observe up to now, a fact that led us to the question whether it has capability to bind them properly. To characterize the substrate binding, co-crystallization experiments of PaCCP with two compounds that block the active-site cleft were carried out. The first compound is guanidinoethylmercaptosuccinic acid (GEMSA), which was selected because of its strong inhibitory specificity for B-type MCP. The other compound is benzylsuccinic acid (BZS), which is a specific inhibitor for A-type MCPs. In both cases, electron density maps indicated the presence of the co-crystallized compounds in one of the two monomers that formed the asymmetric unit of the crystal, indicating major accessibility to one of the active sites (Figure 3.7). The carboxylate groups of GEMSA and BZS are clearly observed in the electron density maps in the same orientation as previously reported co-crystal structures of MCPs (Aloy et al., 2001; Garcia-Saez et al., 1997). The carboxylic oxygens O1 and O2 of BZS and GEMSA are coordinated in a symmetric bidentate form with the zinc ion, approximately at 2.5 Å distance. The inhibitors are also bound to the enzyme by its other carboxylic oxygens, O3 and O4, which establish hydrogen bonds with the side chains of the basic patch of the CP active site (Arg218, Arg228, and Asn227). However, the other half of the compounds, namely the benzyl group and the guanidinoethylmercapto group of BZS and GEMSA, respectively, cannot be observed in the electron density maps and display high temperature factors in both crystal structures (Figure 3.7). This observation indicates that the compounds only partially bind to the expected positions of the enzyme and are not fixed in the specificity pocket of the active site of PaCCP. In contrast, previous co-crystal structures made with those compounds displayed a clear binding in the pocket (Aloy et al., 2001).

The interaction of PaCCP with proteinaceous exogenous inhibitors PCI, NvCI, and TCI and the endogenous inhibitor latexin was evaluated by denaturing PAGE. No stable complexes were formed between PaCCP and CP inhibitors of M14A subfamily MCPs. It is possible that, as in the case of GEMSA and BZS, the exogenous inhibitors are not efficiently fixed in the active-site of PaCCP preventing a stable interaction. Although the inhibition mechanism of latexin is different, it is clear that the interaction with PaCCP is not lasting.
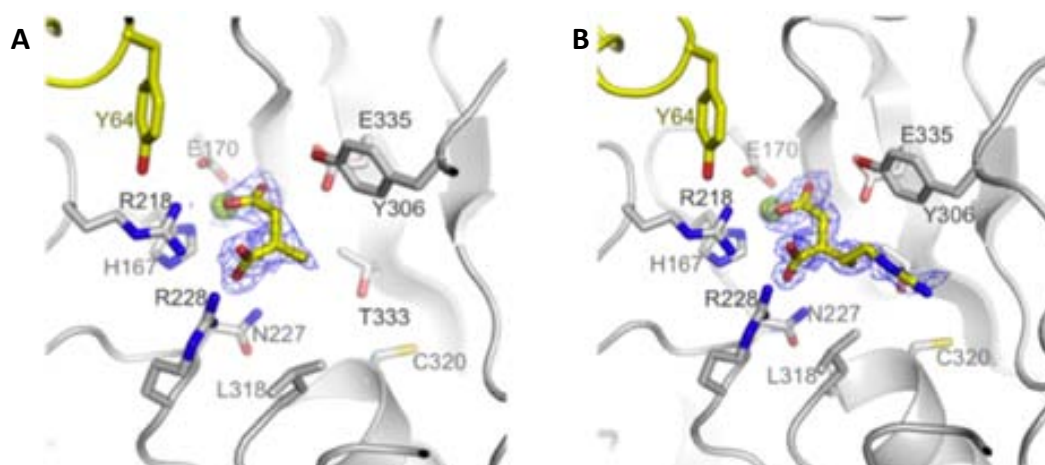
Figure 3.7. Analysis of the active site groove of the *P. aeruginosa* CCP in complex with CP inhibitors. A. Ribbon representation of the active site groove of PaCCP in complex with benzylsuccinic (BZS) B. Ribbon representation of the active site groove of PaCCP in complex with gunidinoethylmercaptosuccinic (GEMSA). BZS and GEMSA molecules are displayed in stick representation around its electron density map (2Fo-Fc map). Graphics prepared with PyMOL.

The interaction of the above-mentioned inhibitors into the active site groove with both the zinc ion and the arginine-rich patch suggests that this enzyme might only possess, in certain conditions, C-terminal hydrolase activity. Besides, the protease does not seem to display preference for the side chains of either aromatic or basic residues, as shown by the lack of electron density of BSZ and GEMSA compounds into the S1 specificity pocket. This means that bacterial CCPs could have the same substrate specificity as the mammalian counterparts.

3.1.7. Comparative modeling of the human CCP subfamily

Sequential and structural comparisons of PaCCP with all six members of the human CCP subfamily show in all of them a similar overall conformation with identically arranged structures for the novel N-domain and for the conserved CP domain (Figure 3.8). Three-dimensional modeling approaches using the PaCCP structure here described could generate structures for the six members of the human CCP family with good scores in comparison to the target structure. The global fold for the N- and CP domains is conserved, and the structural models displayed reasonable C- score, TM-score, and RMS deviation for 361 to 365 aligned residues, with a sequence identity from 21.7 to 25.7% for the six members of the human CCP family (Figure 3.8).

A

B

Human CCP	% Identity	C-score	TM-score	rms deviation (Å)
CCP1	21.7	0.1	0.73 ± 0.11	6.8 ± 4.0
CCP2	23.2	0.5	0.78 ± 0.10	5.8 ± 3.6
CCP3	25.7	0.48	0.78 ± 0.10	5.8 ± 3.6
CCP4	24.0	0.04	0.72 ± 0.11	7.0 ± 4.1
CCP5	21.3	-2.17	0.46 ± 0.15	12.9 ± 4.2
CCP6	23.5	0.95	0.84 ± 0.08	4.8 ± 3.1

Figure 3.8. Homology-based protein structure models of the six human CCPs. A. Two views of the models of the six human CCPs (red) superimposed to the PaCCP structure (blue) in ribbon representation. The zinc atom is displayed as a blue sphere. Graphics prepared with PyMOL. b. Pairwise sequence identity of human CCPs and PaCCP and C-score, TM-score and rms deviation values determined for the CCP models. The C-score (range [-5 to 2]) is a confidence score for estimating the quality of predicted models: the higher value, the higher confidence. TM-score and rms deviation are standards for measuring structural similarity between two structures and were calculated comparing models and PaCCP structure. The TM-score, unlike the rms deviation parameter, is insensitive to the local modeling error and was used to validate the model (TM-score >0.5 indicates a model of correct topology).

Although it is tentative to extract solid conclusions from the analysis with 3-dimensional modeling software, the results indicate that the novel fold for the N-domain and its relative spatial position regarding the CP domain in PaCCP are conserved for all members of human CCP subfamily. Interestingly, in the N-domain, in addition to the conserved secondary structure elements, most of the residues involved in the interface with the CP moiety are conserved between bacteria and human, indicating the validity of the structural model (Figure 3.3). Furthermore, the positions of the residues involved in the zinc binding and catalysis are highly conserved in all structures.

It is also tentative to establish functional relationships between bacterial and mammalian CCPs but is not easy between phylogenetically distant organisms. CCPs from higher organisms show specificity for polyglutamate chains either encoded in the main chain or posttranslationally added to the C-terminus of tubulins and other proteins (Berezniuk et al., 2013a; Rogowski et al., 2010; Wu et al., 2012). Protein FtsZ is the structural homologue of tubulins in bacteria although they have several qualitative changes like a limited sequence similarity (Lowe and Amos, 2009). The C-terminal of FtsZ, in contrast to the highly conserved C-terminal of tubulins, is not conserved across various species in terms of length and amino acid sequence (Buske and Levin, 2012). Therefore, if bacterial CCPs are functionally related to FtsZ they must recognize a wide array of C-terminal residues and not a conserved pattern as in the case of mammalian CCPs. Actually, the presence of proteins with acidic C-terminal in the proteome of *P. aeruginosa* MPAO1 is very scarce having only 4 proteins with 3 or more encoded acidic residues at the C-terminal identified by ScanProsite tool (de Castro et al., 2006). In the human proteome a similar search found 111 proteins (Tanco et al., 2014, submitted) since it has been shown that in higher organisms, these C-terminal acidic tails, provide a mechanism of regulation of protein-protein or DNA-protein interactions (Janke and Bulinski, 2011; Lee and Thomas, 2000). Furthermore, the posttranslational modification consisting in the addition of poly-glutamates chains to the C-terminal of proteins has not been described in bacteria until now (Cain et al., 2014). It appears that, the function of CCPs in eukaryotes might have adapted to the emergence of complex mechanisms of regulation making it difficult to predict their function in prokaryotes.

3.2. Chapter 2. Characterization of a *Pseudomonas aeruginosa* mutant strain deficient in PaCCP

The 3D structure of *P. aeruginosa* MPAO1 CCP (PaCCP) has been obtained and described, providing a good structural model for the conserved N- and CP domains of human CCPs. The zinc binding and catalytic residues of PaCCP seem to be well positioned to carry out the expected CP activity, but cleavage of synthetic substrates, peptides or proteins, has not been detected. As suggested before, it is possible that PaCCP only displays enzymatic activity upon binding its specific substrate(s). This has led us to explore the function(s) of PaCCP in the bacterium *P. aeruginosa*, with the aim to propose related functional mechanisms and eventually a suitable substrate to detect its enzymatic activity. Moreover, considering the biomedical relevance of this microorganism (Mesaros et al., 2007), phenotypes related with virulence and antibiotics resistance were selected to determine if PaCCP has a role in the pathogenesis of *P. aeruginosa.* The study was conducted by comparing *P. aeruginosa* PAO1 wild-type (MPAO1) and a *paccp* negative mutant (Δ*paccp*) strain, both purchased from PAO1 mutant library of the University of Washington (Jacobs et al., 2003).

3.2.1. Influence of *paccp* mutation in the growth of MPAO1

The growth of MPAO1 and Δ*paccp* strains was evaluated at 37 °C, in LB broth to stationary phase. The O.D. at 600 nm was monitored and both strains showed similar growth curves (Figure 3.9A). Additionally, the expression of gene *paccp* was detected by Western Blot with primary antibody anti-PaCCP for each point on the curve. The PaCCP protein was only detected in the wt strain and its concentration increases with culture time (Figure 3.9B). This means that the production of PaCCP depends on culture status and cell density and raises the question of whether or not the protein is related to Quorum Sensing (QS). QS is a well described regulatory mechanism in *P. aeruginosa* and many species of bacteria, controlling the synthesis of several virulence factors like motilities and biofilm formation in response to population density (Cornforth et al., 2014).
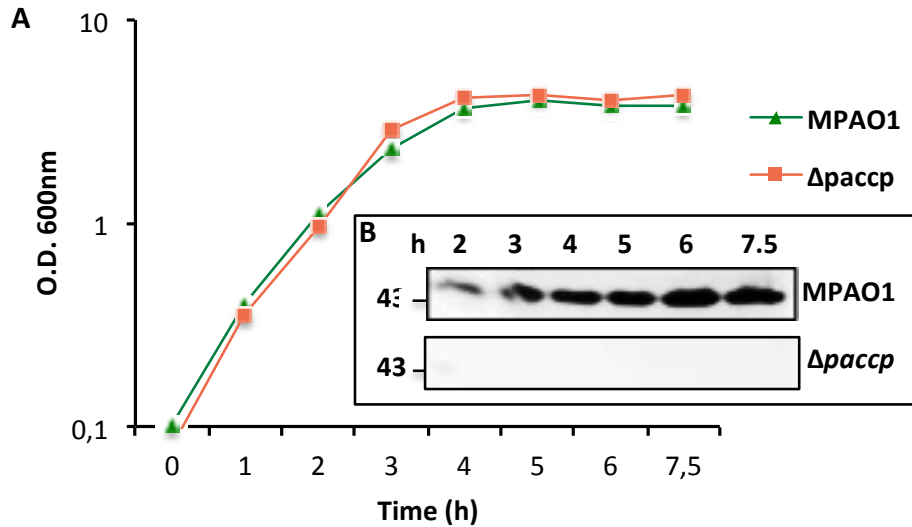
Figure 3.9. Growth curves of MPAO1 and Δ*paccp* strains and detection of protein PaCCP by Western Blot. A. MPAO1 and Δ*paccp* strains were grown in LB medium at 37 °C, 250 rpm. Bacteria were inoculated from overnight cultures grown in the same media. Samples were taken for O.D. 600nm measurements. B. Samples of MPAO1 and Δ*paccp* strains cultures, with normalized numbers of cells, were analyzed by SDS-PAGE followed by Western Blot using anti-PaCCP antibody. Protein molecular marker is indicated in the left in kDa and time of sample collection in hours is indicated above.

3.2.2. Antibiotic resistance

Antimicrobial resistance of *P. aeruginosa* presents a serious therapeutic challenge for treatment of both community-acquired and nosocomial infections. Practically all known mechanisms of resistance can be seen it and the picture is complicated by the development of multi-resistant strains (Morita et al., 2014; Poole, 2011). Some intracellular proteases have been related to antibiotic resistance mechanisms in *P. aeruginosa*. For instance, the disruption of the genes encoding the ATP-dependent proteases Lon and ClpP, independently leads to super-susceptibility to the fluoroquinolone ciprofloxacin (Breidenstein and Hancock, 2013). Another example is the network forms by the membrane protease FtsH and the cytoplasmic proteases HtpX and HslVU that confers robust protection from aminoglycosides (Hinz et al., 2011). Then, our aim was to explore if the PaCCP protease is involved in a mechanism of antimicrobial resistance in bacteria. The Minimum Inhibitory Concentrations (MICs) for front-line anti-pseudomonal drugs used nowadays were determined by Etest for the wt and the mutant strains. The results are shown in Table 3.1.

Table 3.1. MICs of selected antibiotics for MPAO1 and Δ*paccp* strains.

Antibiotic Category	Antibiotic	MIC (μg/ml)	
		MPAO1	Δ*paccp*
β-Lactams	Ceftazidime	0.75	1.5
	Piperacillin	1.5	3.0
	Imipenem	4.0	4.0
Fluoroquinolones	Ciprofloxacin	0.19	0.094
Aminoglycosides	Amikacin	3.0	4.0
	Tobramycin	0.5	0.5
Polymyxin Group	Colistine	3.0	1.5

In general, there are no marked differences in the MICs of either strain for the antibiotics tested. Nevertheless, the resistance pattern of the mutant differs from the wt strain in a slight increase of resistance to ceftazidime and piperacillin (two-fold ratio) and an increase of susceptibility to ciprofloxacin and colistine (two-fold ratio). Apparently, the protease PaCCP is not directly involved in the resistance mechanisms of the antibiotics analyzed here, but modifications in other phenotypes could account for the altered resistance pattern of the mutant.

3.2.3. Proteases secretion

The secretion of proteases to the extracellular medium is a well-described feature of *P aeruginosa,* important in the establishment and maintenance of infections (Hoge, 2010). The production of extracellular proteases in MPAO1 and Δ*paccp* strains was assessed by testing the hydrolysis of casein in agar plates supplemented with skim milk. Both strains showed the typical halo characteristic for casein hydrolysis, but it was highly statistically significant in the mutant strain (Figure 3.10).

The Δ*paccp* mutant not only retains but also has increase ability to degrade casein reflected in a higher hydrolysis zone. This indicates that PaCCP might be involved in the regulation of expression and/or secretion of one or more proteases export to the extracellular medium. Alternative, the enhanced extracellular proteolysis could be an indirect consequence of the alteration of other cellular process by the absence of PaCCP.
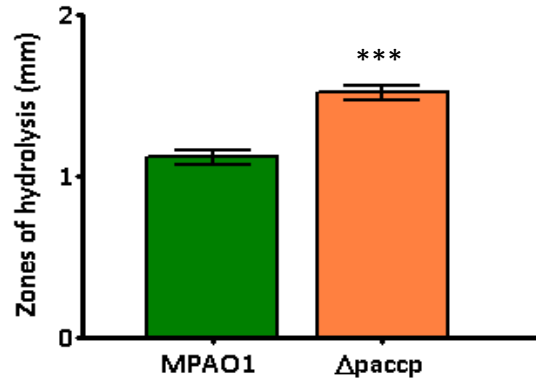
Figure 3.10. Hydrolysis of casein by MPAO1 and Δ*paccp* strains. The hydrolysis of casein determined with a casein hydrolysis test. Mueller-Hinton agar plates supplemented with 2% skim milk were pin-inoculated with bacterial cultures. The diameter of hydrolysis zones was measured for five replicas of each strain. Bars represent the mean values and standard deviations. *** p<0.05 by Student T-Test.

3.2.4. Biofilm formation

P. aeruginosa is able to transit from free-living (planktonic) bacteria into structured communities known as biofilms. In clinical settings, this feature is considered a key component of the adaptation and persistence of *P. aeruginosa*, and the high level of resistance to antimicrobials makes them extremely difficult to eradicate (Hoiby et al., 2011). The biofilm formation capacity of the Δ*paccp* strain was evaluated in a microtiter plate assay. The cells forming the biofilm were stained with crystal violet and the absorbance of the solubilized dye was quantified at 620 nm (Figure 3.11). The biofilm formation was significantly decreased in the Δ*paccp* strain.

Biofilm formation is a complex process that implies diverse physiological changes, and a large number of factors have been shown to contribute to the different stages of its formation. The initial attachment to a surface is often associated with flagella and pili (Klausen et al., 2003; O'Toole and Kolter, 1998) and polysaccharides, such as Pel and Psl, which play an important role in maintaining the integrity of the biofilm community (Friedman and Kolter, 2004). Additionally, biofilm formation is regulates by different systems like cyclic diguanosine-5'-monophosphate, small RNAs and QS controlling the production of individual components in response to environmental stimuli (Fazli et al., 2014).
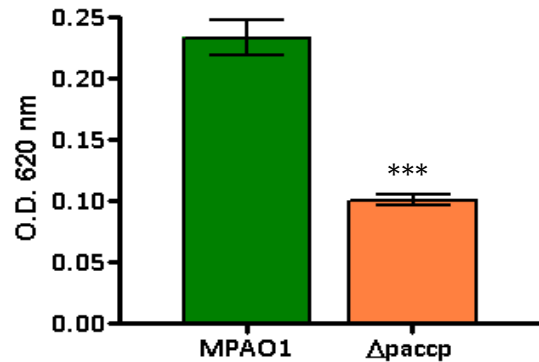
Figure 3.11. Biofilm formation of MPAO1 and Δ*paccp* strains determined with a microtiter plate assay. Cells from overnight cultures were diluted and incubated in a microtiter plate, 24 h at 37 °C. After removing the media, the cells were stained with crystal violet, and transferred to a new plate to measure the O.D. at 620 nm. Bars represent the media and standard deviations of eight replicas for each strain. LB broth was used as control. *** $p<0.05$ by Student T-Test.

Park et al, (2014) examined the influence of growth mode on the protein content of whole cell *P. aeruginosa*, providing a view of both planktonic and biofilm proteomes as they transition through various stages of development. PaCCP was produced in both conditions of growth with the same pattern; it was detected at 24h and 48h but not at 96h. The protein was not found at significantly different levels in neither of the two forms of growth indicating that is not determinant in the transition from planktonic to biofilm life style.

Then it was decided to analyze the effect of *paccp* mutation in the pili and flagella-dependent motilities displayed by *P. aeruginosa*, as important features for biofilm formation.

3.2.5. Motility Assays

P. aeruginosa exhibits different types of motility depending on the composition and viscosity of its environment. The three most characterized to date are flagellum-mediated swarming on semisolid surfaces and swimming in aqueous environments and pilus-mediated twitching on solid surfaces (Rashid and Kornberg, 2000).

Swarming motility was examined on plates containing 0.5% of agar. After incubation overnight, pictures were taken of swarming phenotypes (Figure 3.12).

MPAO1   **Δ*paccp***   **Δ*fliC***
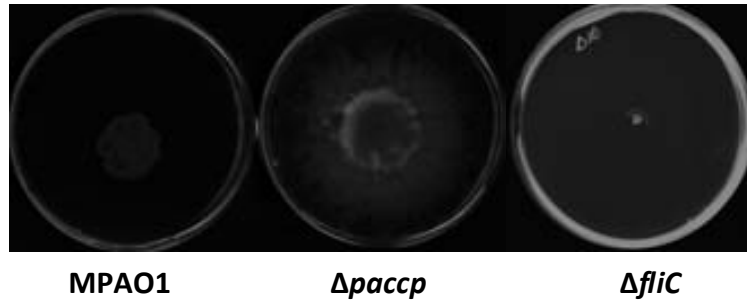
Figure 3.12. Swarming motility of MPAO1, Δ*paccp* and Δ*fliC* strains. Cells were inoculated onto swarm plates (Modified M9 minimal medium supplemented with 0.05% casamino acids plus 0.5% Agar) from overnight LB-agar (1.5%) plates. The resulting colonies were photographed after overnight incubation at 37 °C. Representative images of the swarming motility of three replicas of each strain.

Swarming motility was evidently enhanced in Δ*paccp* strain, occupying practically the entire surface of the plate at the time the picture was taken (Figure 3.12). This behavior cannot be attributed to different growth rate of the strains (Figure 3.9). Mutant Δ*fliC*, deficient in the structural protein of flagella, flagellin, was included as a reference of absence of swarming motility (Rashid and Kornberg, 2000). There is a clear deregulation of swarming motility that probably impacted the biofilm formation. It has been shown that unregulated motility during biofilm formation destabilizes its structure (Klausen et al., 2003; O'Toole and Kolter, 1998). In particular, cells that are enhanced for swarming motility form a flat and featureless biofilm (Shrout et al., 2006). The biofilm of strain Δ*paccp* could be unstructured resulting in fewer cells adhered to the surface of the microtiter plate.

The rotary flagella also account for *P. aeruginosa* cells swimming in aqueous environments (Rashid and Kornberg, 2000). The swimming motility of Δ*paccp* strain was examined in LB broth supplemented with 0.25% agar. After incubation overnight at 30 °C, the diameter of the growth was measured (Figure 3.13A).

Swimming motility is slightly increased in the Δ*paccp* strain by visual inspection, but the difference with the wt strain is not statistically significant (Figure 3.13B). Although swarming and swimming are flagella-dependent motilities, it has been shown that swarming is a more complex type of motility influenced by a large number of different genes (Overhage et al., 2008).
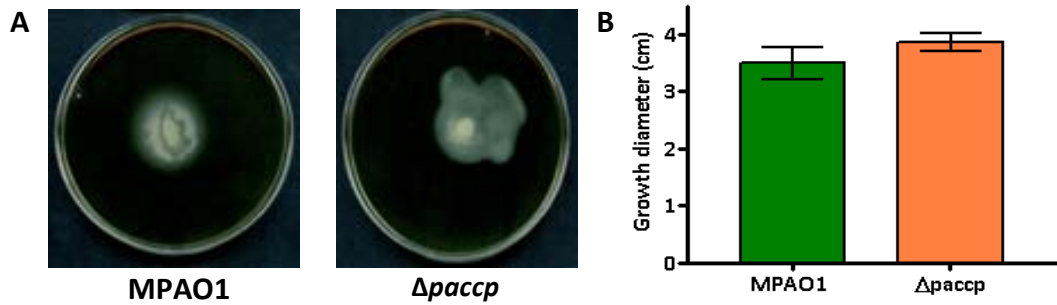
Figure 3.13. Swimming motility of MPAO1 Δ*paccp* on a semisolid agar plate. Cells were inoculated from an overnight LB agar plate onto a swim plate (LB broth plus 0.25% Agar-Noble). A. Representative images of the swimming motility of three replicas of the MPAO1 and Δ*paccp* strains. B. The diameter of the growth was measured for three replicas of each strain. Bars represent the mean values and standard deviations.

P. aeruginosa additionally also exhibit twitching motility, a form of translocation on a solid surface dependent on type IV pili (T4P) (Henrichsen, 1972). The twitching motility was examined in LB broth supplemented with 1% agar. The resulting zones of twitching motility of MPAO1 and mutants Δ*paccp* and Δ*pilA* were visualized by staining the bacteria with crystal violet (Figure 3.14).
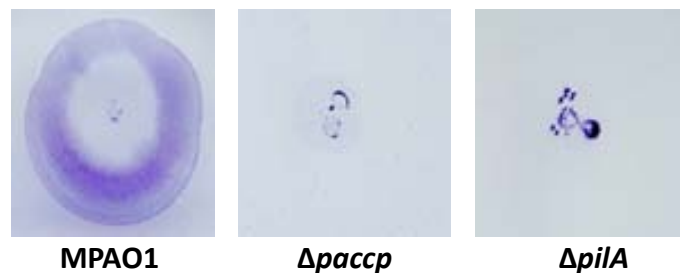
Figure 3.14. Twitching motility of MPAO1, Δ*paccp* and Δ*pilA* strains on semi-solid agar plates. Cells from an overnight LB agar (1.5%), plate were inoculated with a toothpick to the bottom of the LB agar (1%) plate and incubated for 24 h at 37 °C. After removed the agar, bacteria adhered to the petri dish were stained with violet crystal. Representative images of the twitching motilities zones of five replicas of each strain.

The Δ*pilA* strain, deficient in the T4P structural protein, was included in the study as a reference of total deficiency in twitching motility (Bradley, 1980). Interestingly the twitching motility was completely abolished in the Δ*paccp* strain showing the same phenotype of the Δ*pilA* strain (Figure 3.14). The effect observed cannot be explained for a growth defect, as both strains have similar growth curves under rich-medium growth conditions (Figure 3.9).

It has been shown that twitching motility mutants, lacked functional T4P, were either apiliated or hyperpiliated (Bradley, 1980). To characterize the surface morphology of the

Δ*paccp* strain regardless of pili production, several negative stained cells were analyzed by transmission electron microscopy (TEM) (Figure 3.15).
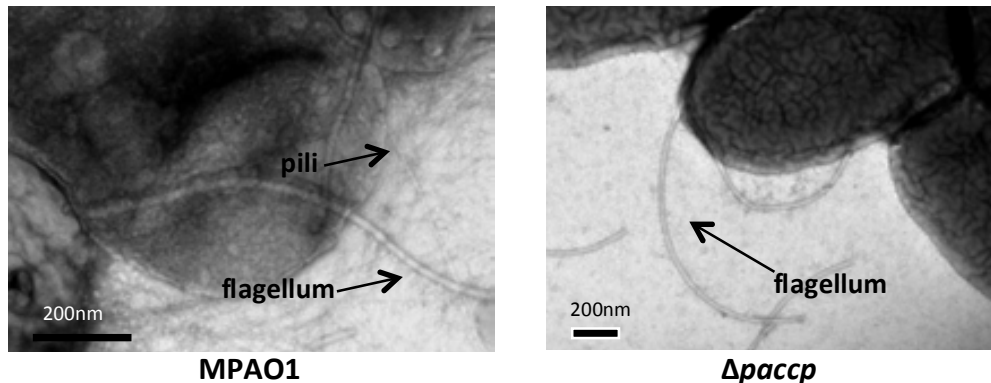
Figure 3.15. TEM analysis of MPAO1 and Δ*paccp* cells. Bacteria were harvested from plates cultured overnight, negatively stained using 2% uracyl acetate, and visualized using JEOL JEM 1400 electronic microscopy. The images are representative bacterium of those observed for each strain.

The presence of T4P was only observed in the surface of wt strain cells. Conversely, the assembly of flagella was not affected in the Δ*paccp* cells. The absence of T4P and, therefore, of twitching motility probably also contribute to the reduced capacity of the Δ*paccp* strain to form biofilms. O´Toole and Kolter (1998) first reported that *P. aeruginosa* mutants lacking T4P are deficient in biofilm formation. In addition to promoting adherence and aggregation in biofilms, T4P participates in remodeling biofilm architecture through twitching motility (Klausen et al., 2003).

In the absence of PaCCP, the MPAO1 cells lost twitching motility due to a total deficient in T4P in the bacterial surface. To assess the rescue of the wt phenotypes, the mutant strain was complemented by synthesis of protein PaCCP from a plasmid. For that, the encoding and regulatory sequences of the *paccp* gene were cloned into the plasmid pBBR1-MSC5 for expression under its own promoter. The plasmid was transformed into Δ*paccp* and wt strains and the expression of the protein was first confirmed by Western Blot with primary antibody anti-PaCCP (Figure 3.16A). MPAO1 and Δ*paccp* strains, transformed with plasmid *paccp*-pBBR5, successfully produced the PaCCP protein but at much higher concentration than wt levels. Although the protein is expressed from the plasmid under its own promoter, the plasmid copy number might accounts for the higher production of PaCCP. The expression of the *paccp* gene from the plasmid is achieved, but twitching motility was not restored in the mutant strain after complementation with PaCCP wt (Figure 3.16B) nor the protease secretion, biofilm formation and swarming motility wt phenotypes were restored.
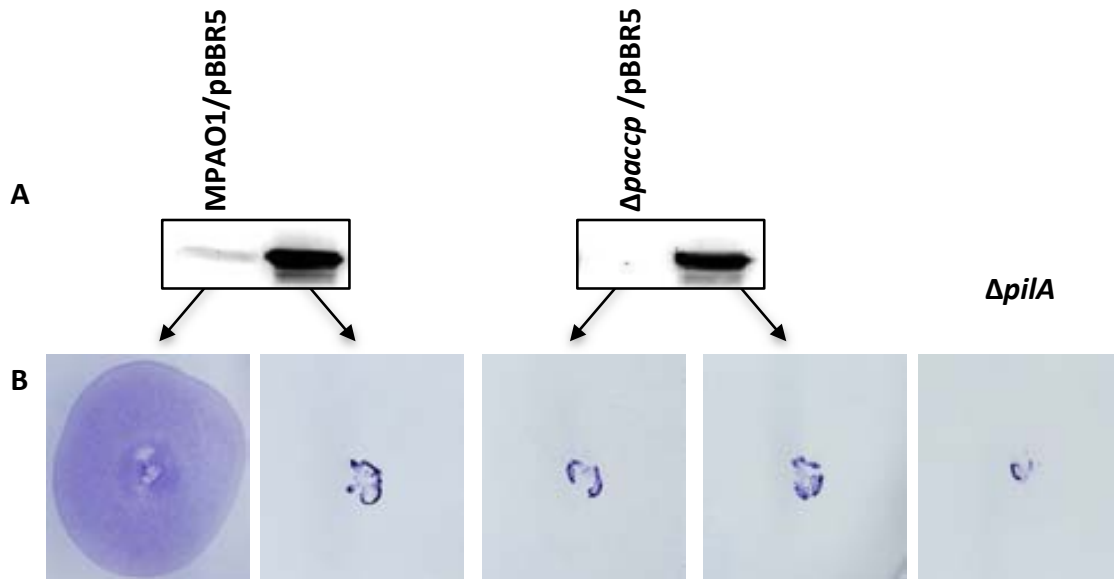
Figure 3.16. Production of PaCCP wt protein and twitching motility in MPAO1 and Δ*paccp* complemented strains and controls. A. PaCCP production was checked by Western Blot with anti-PaCCP antibody. All strains were cultured 4h, at 37°C in LB broth supplemented with gentamicin 40 μg/ml. B. Representative images of twitching motility zones for MPAO1, Δ*paccp* and Δ*pilA* strains.

Interestingly, the MPAO1 strain lost its twitching motility when higher intracellular concentration of PaCCP was supplemented in the wt background. This fact suggests that PaCCP activity is regulated by its concentration, and variations outside the optimal range, could prevent its functioning. A similar regulation was seen for the peptidoglycan-binding protein Fimv, a regulator of twitching motility that promotes the assembly of the *P. Aeruginosa* T4P secretin (Wehbi et al., 2011). Overexpression of FimV interfered with normal twitching motility both in the mutant and in wt strains and a specific level of FimV expression is required for normal motility (Semmler et al., 2000). The inability to complement Δ*paccp* phenotypes with plasmid pBBR5-*paccp,* coupled with the deficient phenotypes observed in the wt strain carrying the complementation plasmid could be explained in many different ways. For instance, a higher concentration of PaCCP could promote the formation of protein aggregates preventing the access of substrate(s) to the active site of PaCCP and/or blocking regions involved in protein-protein interactions. A second explanation may be that PaCCP sequesters a functional partner making it unavailable in the cell.

The absence of T4P in the surface of the Δ*paccp* strain suggests that PaCCP could be required for *pilA* expression, T4P assembly or T4P export. To address this question, the

surface sheared proteins of MPAO1 and Δ*paccp* strains, transformed with *paccp*-pBBR5 or the empty vector, were analyzed. Additionally, the Δ*pilA* and Δ*fliC* strains, deficient in the structural proteins of pili and flagellum, respectively, were included in the study. The cells were subjected to vigorous vortexing and the resulting sheared proteins were analyzed on SDS-PAGE (Figure 3.17, top).
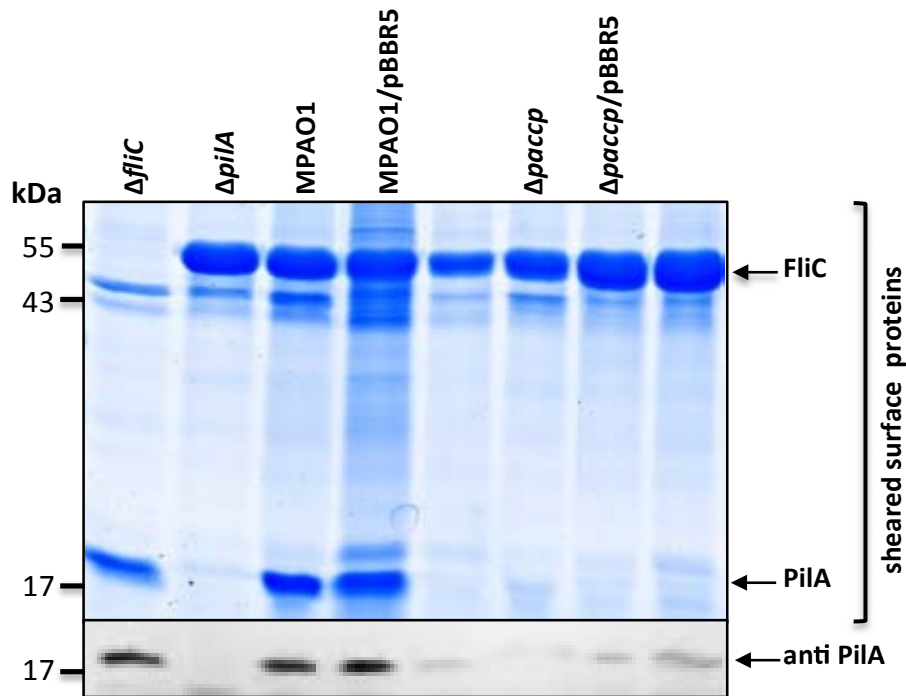
Figure 3.17. Analysis of sheared surface proteins by SDS-PAGE and the intracellular levels of PilA by Western Blot. (Top) Representative sheared-surface proteins preparation of MPAO1, Δpaccp, ΔfliC, ΔpilA strains; MPAO1 and Δpaccp strains transformed with empty vector pBBR1-MSC5 (MPAO1/pBBR5 and Δpaccp/pBBR5, respectively) and MPAO1 and Δpaccp strains transformed with the vector paccp-pBBR1-MSC5 (MPAO1/pBBR5-paccp and Δpaccp/pBBR5-paccp, respectively) visualized on Coomassie-stained SDS-PAGE 15%. (Bottom) Western Blot analysis of the whole-cell lysates of the sheared cells probed using an anti-PilA antibody. In both cases, the numbers of cells used were normalized by dilution of the samples to the same optical density at 600 nm.

The most abundant proteins in the surface of the *P. aeruginosa* MPAO1 cells are the appendages structural proteins pilin and type-B flagellin. The expression of flagellin is similar in Δ*paccp* and wt strains and does not seem to be a determinant factor for the differences observed in swarming motility (Figure 3.12). On the other hand, PilA cannot be detected in Δ*paccp* strain, which explains the absence of the pili observed in TEM and therefore twitching motility (Figures 3.14 and 3.15). As expected from twitching motility experiments, complementation of the mutant strain with the *paccp*-pBBR5 plasmid, failed to recover wt PilA levels. The PilA protein is undetectable in the extracellular medium of the wt strain

when PaCCP wt is supplemented in trans, confirming the lack of function of PaCCP at higher concentrations.

It can be stated now that the defect in twitching motility is not in the assembly of T4P but in the expression of *pilA* or T4P export. To elucidate between these two processes, the intracellular levels of PilA of sheared cells were assessed by Western Blot with anti-PilA antibody (Figure 3.17 bottom). The levels of intracellular PilA were almost undetectable in Δ*paccp* and the strains carrying the PaCCP wt protein expression plasmid. These results indicate that PaCCP is involved directly or indirectly in the regulation of the expression of the *pilA* gene.

It has been shown that the expression of gene *pilA* depends on the sigma factor RpoN together with the PilR-PilS two-component system (Hobbs et al., 1993). Autophosphorylation of PilS and subsequent phosphorylation of its response regulator, PilR, upregulate pilA transcription. The exact signal sensed by PilS has not yet been determined but is potentially PilA itself. In a *pilB* mutant, where unassembled pilins accumulate in the inner membrane, *pilA* transcription was reduced. In *pilT* mutants, where pilin pools were depleted because of unopposed assembly in the absence of retraction, *pilA* transcription was increased (Bertrand et al., 2010). PilS levels are also critical to the control of PilA levels, as both, lack of PilS and its overexpression, had a detrimental effect on pilin gene transcription. It has been proposed that PilA negatively regulates its own expression, through direct intramembrane interaction with the PilS (Bertrand et al., 2010). The levels of intracellular PilA might not be the only intracellular signal for PilS to regulate the expression of pilin genes. It was suggested that PilS overexpression titrated a crucial interaction partner, leading to decreased pilin gene transcription (Boyd and Lory, 1996). In general, PilR-PilS may regulate other components in the cell in response to alterations in the amount of pilin detected. Bioinformatic studies of *G. sulfurreducens* promoters suggested that PilR might regulate the production of pili, flagella, secretion components, and proteins involved in cell envelope biogenesis (Krushkal et al., 2010). We have identified a CCP whose absence causes a significant reduction in the levels of PilA but also affects other process like swarming motility, biofilm formation and proteases secretion. One option is that PaCCP is a signal itself or generates, via it enzymatic activity, new signals for PilR-PilS that in turn modulates the promoter activity of multiples genes.

3.2.6. Virulence *in C. elegans* slow-killing model

Bacterial motilities and biofilm formation have been identified as virulence factors in acute and chronic infections of *P. aeruginosa* (Balasubramanian et al., 2013). The altered phenotypes found in the Δ*paccp* strain probably influenced its pathogenicity. The virulence of the Δ*paccp* strain was evaluated in a *C. elegans* slow-killing model; a useful tool for the identification of known and novel virulence factors in *P.aeruginosa*, important for mammalian pathogenesis (Feinbaum et al., 2012; Tan et al., 1999). The Δ*paccp* and wt strains were grown in NGM agar plates, seeded with adult worms and live worms recorded every 24h. *E. coli* OP50 was used as negative control strain (Figure 3.18).
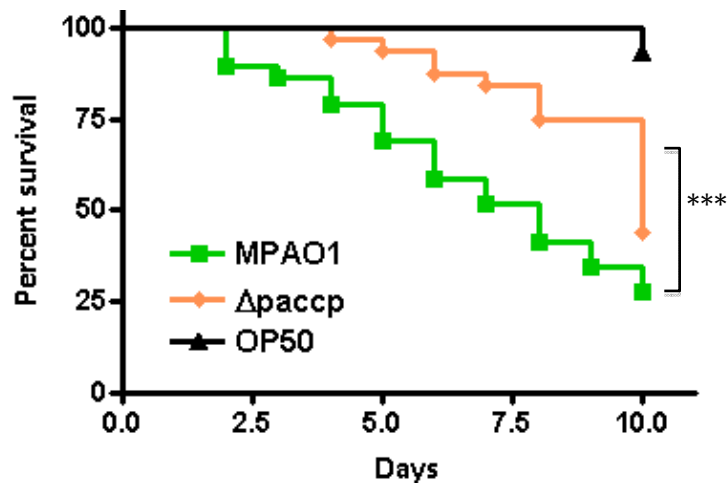
Figure 3.18. Virulence in *C. elegans* of Δ*paccp* strain compared with the parental MPAO1 in slow-killing assay. Strains grown in LB broth plates, overnight at 37 °C were spread on NGM agar plates. After incubation, 24 h at 37 °C, the plates were seeded with 15-20 adult worms, incubated at 25 °C and scored for live worms every 24h. *E. coli* OP50 was used as a negative control. Kaplan-Meier survival plots for MPAO1 (n=30), Δ*paccp* (n=30) and *E. coli* OP50 (n=30) strains. Log-rank analysis (Mantel-Cox) was used to compare Δ*paccp* with MPAO1 strains. ***p <0.05

The Δ*paccp* strain showed reduced virulence towards *C. elegans* in comparison with the wt strain. In this model, bacteria colonize the intestine of the worms and establish an infectious process (Tan et al., 1999). The deficiency in twitching motility and the reduced capacity to form biofilm probably contribute to the reduced virulence of the mutant strain. This is a very important result and further experiments must be performed in mammalian models to see if PaCCP also plays a role in its pathogenesis.

3.2.7. Interactomic studies

With the aim of gaining insight into the mechanisms involving PaCCP, an interatomic approach was performed to identify interacting partners of the protein. The study was carried out with magnetic beads activated with epoxy groups to which the recombinant purified PaCCP protein or antibodies that recognize PaCCP were coupled. The beads were incubated with the soluble fraction of the MPAO1 strain and proteins that remained bound, after several washing steps, were identified by nano-Esi-MS/MS. The proteins also identified in control experiments were eliminated from the final list of results. Table 3.2 shows the proteins identified in the pull-down using the recombinant purified PaCCP.

Table 3.2. List of proteins identified in pull-down experiments using recombinant purified PaCCP.* Protein score is -10*Log(P), where P is the probability that the observed match is a random event. Protein score for one of the replica is shown.

UniProt ID	Locus ID	Gene	Protein name	Protein Score*	PseudoCAP function or (annotation from other DB)
P72151	PA1092	*fliC*	B-type flagellin	2311	Ciliary or bacterial-type flagellar motility
Q9HVC5	PA4670	*prs*	Ribose-phosphate pyrophosphokinase	742	Cellular catabolic process, pentose-phosphate shunt
Q9S646	PA5242	*ppk*	Polyphosphate kinase	559	Nucleotide metabolic process
G3XD61	PA3148	*wbpl*	UDP-2,3-diacetamido-2,3-dideoxy-D-glucuronate 2-epimerase	474	O antigen biosynthetic process (UniProtKB-UniPathway)
Q9HV43	PA4761	*dnaK*	Chaperone protein DnaK	458	DNA metabolic process, protein folding
Q51342	PA3108	*purF*	Amidophosphoribosyl transferase	429	Glutamate metabolic process, purine nucleotide metabolic process
Q9HVJ1	PA4595		Probable ATP-binding component of ABC transporter	412	ATP catabolic process (GOC)
Q9I691	PA0422		Uncharacterized protein	406	Respiratory electron transport chain (InterPro)
Q9I1M1	PA2248	*bkdA2*	2-oxoisovalerate dehydrogenase subunit beta	365	Cellular amino acid metabolic process
P37798	PA4848	*accC*	Biotin carboxylase	334	Cellular lipid metabolic process
Q9HVU0	PA4481	*mreB*	Rod shape-determining protein MreB	327	Cytokinesis by binary fission
Q9I4Z5	PA0968		Uncharacterized protein	326	Metabolic process (GOC)
P08280	PA3517	*recA*	Protein RecA	274	DNA metabolic process
G3XD23	PA3158	*wbpB*	UDP-N-acetyl-2-amino-2-deoxy-D-glucuronate oxidase	270	O antigen biosynthetic process (UniProtKB-UniPathway)
Q9L7T2	PA0357	*mutM*	Formamidopyrimidine-DNA glycosylase	264	DNA metabolic process
Q9HXZ4	PA3637	*pyrG*	CTP synthase	262	Nucleotide metabolic process

Q9HUG0	PA5005		Probable carbamoyl transferase	242	Biosynthetic process (InterPro)
Q9I296	PA2015	*liuA*	Putative isovaleryl-CoA dehydrogenase	238	Cellular catabolic process
Q9I1M2	PA2247	*bkdA1*	2-oxoisovalerate dehydrogenase subunit alpha	223	Cellular amino acid metabolic process
Q9I299	PA2012	*liuD*	Methylcrotonyl-CoA carboxylase, alpha-subunit (Biotin-containing)	219	Cellular catabolic process
Q9I3D5	PA1583	*sdhA*	Succinate dehydrogenase (A subunit)	209	Generation of precursor metabolites and energy
Q9HY54	PA3563	*fruR*	Fructose transport system repressor FruR	206	Cellular catabolic process, regulation of transcription, DNA-templated, transport
G3XCW7	PA1462		Probable plasmid partitioning protein	205	na
Q9HZV0	PA2897		Probable transcriptional regulator	170	Transcription, DNA-templated (UniProtKB-KW)
Q9I5I4	PA0745		Probable enoyl-CoA hydratase/isomerase	165	na
Q9HT18	PA5556	*atpA*	ATP synthase subunit alpha	162	Generation of precursor metabolites and energy
G3XD01	PA3156	*wbpD*	UDP-2-acetamido-3-amino-2,3-dideoxy-d-glucuronic acid N-acetyltransferase	161	O antigen biosynthetic process (UniProtKB-UniPathway)
Q9I0J9	PA2639	*nuoC*	NADH-quinone oxidoreductase subunit C/D	159	Oxidative phosphorylation
Q9I1S2	PA2194	*hcnB*	Hydrogen cyanide synthase subunit HcnB	155	Metabolic process
Q9HXP8	PA3746	*ffh*	Signal recognition particle protein	145	Secretion
Q7DC82	PA1901	*phzC1*	Phenazine biosynthesis protein PhzC	142	Aromatic amino acid family biosynthetic process (InterPro)
Q9I3D4	PA1584	*sdhC*	Succinate dehydrogenase (B subunit)	138	Generation of precursor metabolites and energy
Q9HWP9	PA4132		Uncharacterized protein	137	Transcription, DNA-templated (UniProtKB-KW)
Q9HTP6	PA5308	*lrp*	Leucine-responsive regulatory protein	137	Metabolic process, regulation of transcription, DNA-templated
Q9HZE0	PA3068	*gdhB*	NAD-specific glutamate dehydrogenase	137	Cellular amino acid metabolic process
Q51366	PA5453	*gmd*	GDP-mannose 4,6-dehydratase	136	Lipopolysaccharide biosynthetic process (InterPro)
Q9HY63	PA3554	*arnA*	Bifunctional polymyxin resistance protein ArnA	132	Lipid A biosynthetic process, response to antibiotic (UniProtKB-KW)
Q9HTQ0	PA5304	*dadA1*	D-amino acid dehydrogenase 1 small subunit	132	L-phenylalanine metabolic process, generation of precursor metabolites and energy
Q9I3A5	PA1617		Probable AMP-binding protein	131	na
Q9HXJ4	PA3803	*ispG*	4-hydroxy-3-methylbut-2-en-1-yl diphosphate	126	Terpenoid biosynthetic process (UniProtKB-HAMAP)

			synthase		
Q9I4I1	PA1156	*nrdA*	Ribonucleotide-diphosphate reductase	122	Nucleotide metabolic process
Q9HZ76	PA3155	*wbpE*	UDP-2-acetamido-2-deoxy-3-oxo-D-glucuronate aminotransferase	112	O antigen biosynthetic process (UniProtKB-UniPathway)
Q9I687	PA0430	*metF*	Methylenetetrahydrofolate reductase	111	Cellular amino acid metabolic process
Q9HT66	PA5507		Uncharacterized protein	107	na
P33883	PA1432	*lasI*	Acyl-homoserine-lactone synthase	107	Quorum sensing (UniProtKB-KW)
O69754	PA1904	*phzF1*	Trans-2,3-dihydro-3-hydroxyanthranilate isomerase	105	Antibiotic biosynthetic process (UniProtKB-KW)
Q9I3B2	PA1609	*fabB*	Beta-ketoacyl-ACP synthase I	105	Cellular lipid metabolic process
Q9I1Z9	PA2116		UPF0317 protein PA2116	105	na
Q9I4D7	PA1201		Probable transcriptional regulator	103	Transcription, DNA-templated (UniProtKB-KW)
Q9I297	PA2014	*iluB*	Methylcrotonyl-CoA carboxylase beta subunit	102	Cellular catabolic process

Abbreviations: DB: databases; na: none available/unknown

A high number of proteins were identified as potential interactors of PaCCP. The great majority, around 74%, can be group in the general category of metabolism-related proteins (biosynthesis, catabolism and generation of precursors metabolites). Other functional categories (motility, cytokinesis, secretion, protein folding, quorum sensing, and transcriptional regulators) are underrepresented.

Immunoprecipitation assays were also performed for identification of interactors of PaCCP. These types of experiments have the advantage of capturing the endogenous protein avoiding artifacts of the recombinant production and purification. The proteins identified with this approach are listed in Table 3.3. In this case, fewer proteins were identified, most of them with protein scores below 100. This is probable due to the lower concentration of the endogenous PaCCP in the lysate. From a statistical point of view, the majority of proteins identified in the immunoprecipitations experiments are not considered significant results. Nevertheless, it is useful preliminary information that must be validate in future experiments using higher levels of the endogenous protein. There was very little overlap in the results obtained by the two methods. A possible explanation is that the higher concentration of PaCCP in pull-down experiments, resulted in higher levels of nonspecific proteins that interfered the binding of true interactors, specially the low abundant. The binding of nonspecific proteins, in particular from ribosomes, was a common

characteristic to both types of experiments although higher in pull-down assays. The detection of protein PilA in the immunoprecipitation assays is particularly relevant for the potential implication of PaCCP in *pilA* expression. Further experiments should be done to elucidate whether or not PilA interacts with PaCCP.

Table 3.3. List of proteins identified in immunoprecipitation experiments with anti-PaCCP antibodies.* Protein score is -10*Log(P), where P is the probability that the observed match is a random event. ‡ indicates proteins also identified in pull-down experiments (Table 3.2).

UniProt ID	Locus ID	Gene	Protein name	Protein Score*	PseudoCAP function or (annotation from other DB)
P72151	PA1092	*fliC*	B-type flagellin‡	173	Ciliary or bacterial-type flagellar motility
P30718	PA4385	*groL*	60 kDa chaperonin	86	Protein folding
Q9HUC3	PA5060	*phaF*	Polyhydroxyalkanoate synthesis protein PhaF	86	Metabolic process
P04739	PA4525	*pilA*	Fimbrial protein	72	Cellular component movement
Q9HYT5	PA3309		Universal stress protein‡	47	Response to stress (InterPro)
Q9HW86	PA4315	*mvaT*	Transcriptional regulator MvaT, P16 subunit	47	Regulation of transcription, DNA-templated
Q9HW49	PA4352		Uncharacterized protein	46	Response to stress (InterPro)
Q9HTL0	PA5348	*hupA*	DNA-binding protein HU-alpha	45	Chromosome condensation (UniProtKB-KW)
Q9I6Z1	PA0141		Uncharacterized protein	38	Phosphorylation (GOC)
Q9I1M0	PA2249	*bkdB*	Branched-chain alpha-keto acid dehydrogenase subunit E2	31	Cellular amino acid metabolic process

Abbreviations: DB: databases; na: none available/unknown

According to the phenotypes observed in the mutant strain, more proteins from motility machineries and/or regulators would be expected. Most proteins of motility systems are poorly represented in the soluble fraction because they are inserted or associated with the cytoplasmic and the outer membranes. The results obtained may be biased by the cellular fraction used and future interatomic studies should also include the membrane fraction.

Draw conclusions from experiments involving MS-based protein identifications are difficult because are so sensitive that a large number of identified proteins could be contaminants. Additionally, the finding of substrates for PaCCP with this interatomic approach, it is not probable. Interactions enzyme-substrate are usually transients and could be dissociated by the end of the protocol. Anyway, this approach provides a first glance of

the potential interactors of PaCCP and some candidates can be suggested for future validation of the interaction.

Proteins that were found by the two methods (Tables 3.2 and 3.3), B-type flagellin and universal stress protein (PA3309) could be consider first. PA3309, a protein with the signature domain of universal stress proteins (Usp), is important for pyruvate fermentation and the survival of *P. aeruginosa* during anaerobic conditions (Schreiber et al., 2006). This protein is abundant in *P. aeruginosa* biofilms according to proteomic studies and selective green fluorescent protein labeling (Williamson et al., 2012)

Considering the phenotypes showed by the Δ*paccp* strain, it is not surprising to find proteins related with motility machineries like B-type flagellin, the subunit that polymerizes to form the filaments of bacterial flagella (Jarrell and McBride, 2008). In Δ*paccp* strain, swarming motility was greatly increased, although this phenotype cannot be attributed to different levels of B-type flagellin or defects in the biogenesis of flagellum (Figures 3.16 and 3.17 top). Regarding the possibility that flagellin is a substrate, the hydrolysis of its C-terminal has not been reported and this portion is important for polymerization of the flagellum (Mimori-Kiyosue et al., 1997).

Other potential interactor worth analyzing is the enzyme polyphosphate kinase (PPK), responsible for the synthesis of inorganic polyphosphate. PPK is essential in *P. aeruginosa* not only for swimming, swarming and twitching motilities (Rashid and Kornberg, 2000; Rashid et al., 2000a) but also for the development of biofilms, production of the virulence factors elastase and rhamnolipid, and for virulence in the burned-mouse pathogenesis model. All of these effects are likely exerted through a defect in QS and responses (Rashid et al., 2000b). As PaCCP has effects over a wide array of phenotypes, a direct or indirect functional relation with QS could also exist. This hypothesis is further support by the fact that the enzyme acyl-homoserine-lactone synthase (LasI) appears as a potential interactor of PaCCP (Table 3.2). LasI produces N-3-oxo-dodecanoyl-homoserine lactone and forms part of the *las* signaling system together with the signal receptor LasR, which binds its cognate signal and activates transcription of certain target genes (Schuster and Greenberg, 2006).

The secondary metabolites hydrogen cyanide and phenazines has been described as virulence factors in the "fast-killing" *C. elegans* model (Cezairliyan et al., 2013; Gallagher and Manoil, 2001). Two components of enzyme complexes responsible for the synthesis of these

metabolites, the hydrogen cyanide synthase (HcnB) and phenazine biosynthesis protein PhzC, were identified as potential interactors of PaCCP (Table 3.2). HcnB is a subunit of the membrane-bound flavoenzyme encoded by *hcnABC* and transcription of the *hcn* genes appears to depend directly on the QS regulators LasR and RhlR (Gallagher and Manoil, 2001). PhzC is encoded in one of the two seven-gene phenazine operons responsible for the synthesis of phenazine-1-carboxylic acid. The regulation of phenazine synthesis in *P. aeruginosa* is mediated by N-acyl-homoserine lactone and involves at least two different QS circuits, las and rhl, organized in a hierarchical cascade (Mavrodi et al., 2001).

Another identified potential interactor of PaCCP is the actin homolog in rod-shaped bacteria, MreB (Table 3.2). This cytoskeletal protein forms discrete "patches" that move perpendicularly to the cell axis and is supposed to act as a scaffold for transporting proteins to different locations. MreB is essential for maintenance of cell shape, chromosome segregation, and polar localization of several bacterial proteins like PilT, the pilin depolymerase of *P. aeruginosa* T4P (Divakaruni et al., 2007; Gitai et al., 2004). It is difficult to predict the nature of the interaction between MreB and PaCCP but the relation with cytoskeleton proteins would be a shared featured with eukaryotic CCPs, which have tubulins as substrate (Rogowski et al., 2010).

Finally, the transcriptional regulators identified in the interactomic assays could also be taking into account for further analysis. Some of them are uncharacterized proteins, opening the opportunity for discovery of novel regulatory mechanisms of virulence factors in *P. aeruginosa*.

The field of bacterial CCPs has been practically unexplored until now. These proteins were identified a few years ago and only found in a reduced number of gram-negative bacteria, mostly proteobacteria (Rodriguez de la Vega et al., 2007). PaCCP impacts several virulence factors in *P. aeruginosa* MPAO1, and the absence of the protein results in a less virulence strains in a *C. elegans* slow-killing model. A similar result was obtained for the orthologous of *paccp* in the pathogenic *P. aeruginosa* strain LESB58 that resulted one of the 47 genes essential for lung infection in the chronic rat lung infection model (Winstanley et al., 2009). Several questions remains to be answered for bacterial CCPs, for instance, its limited distribution in a small group of bacteria. It is possible that CCPs, predicted to be in the last eukaryotic common ancestor (Rodriguez de la Vega Otazo et al., 2012), were

retained in some bacteria because provide adaptive advantages in both their natural and host environments, or belong to an exclusive pathway not describe so far. Bacteria lacking CCPs could develop alternative functional mechanisms, more effective in its environments, where CCPs function is exert by other proteins. A similar selective mechanism may have occurred in eukaryotes where CCPs are found in the major groups but only in the basal lineages of plants and fungi. The functions of eukaryotic CCPs have been primarily related with cilia and basal bodies, which were lost in superior fungi and plants (Rodriguez de la Vega Otazo et al., 2012).

3.3. Chapter 3. CCP6, the shortest variant of human cytosolic carboxypeptidases

The CCP6, among the six human orthologous, is the closest to bacterial CCPs, in terms of sequence length and predicted 3D structure. Human CCP6 (hCCP6) is mainly formed by the N-terminal and CP-domain (Rodriguez de la Vega et al., 2007), which in the practice makes it as a minimal version of human CCPs. Hence, it was decided to over-express and characterize some properties of such hCCP6 because it was expected to be one of the human forms experimentally easier to handle.

3.3.1. Production and evaluation of the enzymatic activity of hCCP6 in HEK 293F cells

The cDNA encoding for hCCP6 (Uniprot Q5VU57) was cloned into vector bacterial/mammalian expression plasmid pTriEx6, introducing a N-terminal *Strep*-tag II (Schmidt and Skerra, 2007) and a hemagglutinin epitope (HA) in the protein. In addition, a catalytic inactive gene mutant hCCP6 E425Q (equivalent to E270 in bCPA) (Kim and Lipscomb, 1990; Wu et al., 2012) was generated and cloned into the same vector. These vectors were transiently transfected in HEK 293F cells and the expression was analyzed by Western blot using the anti-HA antibody. hCCP6 and mutant E425Q proteins were successfully overexpressed, with the expected molecular weight (i.e. 58 kDa) (Figure 3.19A).

In parallel, aiming to get large quantities of the recombinant proteins for crystallization experiments, attempts were made to produce it in different *E.coli*-based expression system. Nevertheless, bacterial system was discarded because in all attempts, the recombinant protein produced was not active and the protein yield was low due to a variety of truncated forms. This latter could result from proteolytic degradation and/or premature termination of transcription and/or translation. Probably, mammalian CCPs are difficult to produce in such systems, due to their size (up to 120 kDa for some member) and proteolytic susceptibility, among others. Previous studies successfully used the baculovirus system for CCP1 and CCP5 expression and purification (Berezniuk et al., 2013a; Berezniuk et al., 2012; Wu et al., 2012). Our group already has produced some active mammalian CCPs in the HEK

293F expression system (Tanco et al., 2014, submitted; Tort et al., 2014, submitted) and the results obtained here corroborate that this is a suitable system for studying CCPs.

The enzymatic activity of produced hCCP6 against tubulins was analyzed in protein extracts of HEK 293F cells, overexpressing active or enzymatically inactive hCCP6. The potential protein substrates in the cleared lysates were separated by SDS-PAGE followed by Western blot using antibodies that recognized different posttranslational modifications (PTM) of tubulin (Figure 3.19A).
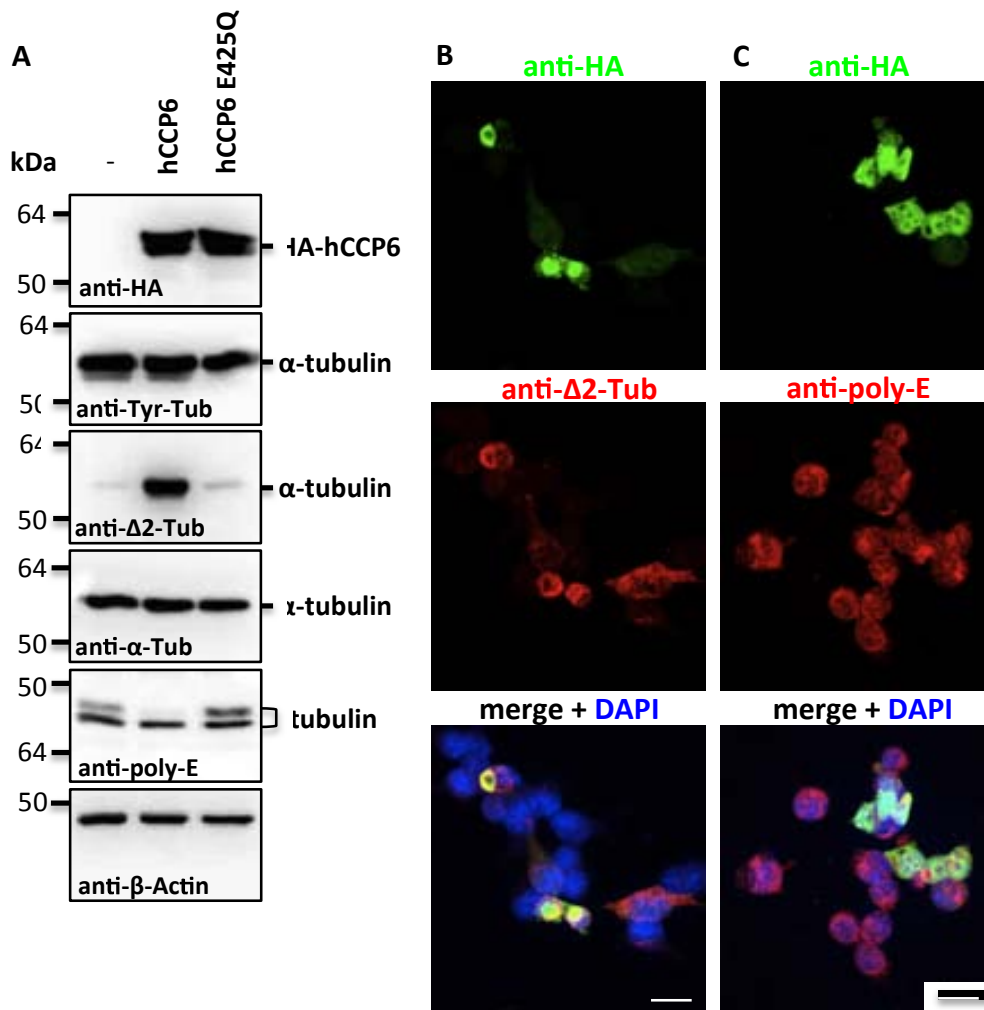
Figure 3.19. Enzymatic characterization of hCCP6 on tubulins. A. Immunoblot of protein extracts from HEK 293F cells expressing active (HA-hCCP6) or inactive (HA-hCCP6 (E425Q)) CCP6. B, C. Immunofluorescence of HEK 293T cells expressing HA-hCCP6. B. Anti-Δ2-tubulin antibody specifically labels the cell expressing HA-hCCP6. C. PolyE antibody signal is reduced in the cell expressing HA-hCCP6. Scale bar is 20 μm.

From this assays it can be concluded that overexpressed hCCP6 did not alter the levels of tyrosinated tubulin, and therefore was not the enzyme responsible for C-terminal tubulin detyrosination. In contrast, there was a remarkable increase in the levels of Δ2 tubulin,

which indicates that hCCP6 was able to hydrolyze the C-terminal Glu of Glu-tubulin. This effect was dependent of the enzymatic activity of hCCP6 because no variation in Δ2 tubulin levels was detected when the catalytic inactive mutant was over-expressed. This is in agreement with previous results obtained in mouse CCP6 (Rogowski et al., 2010) and with its ortholog in *C.elegans* (Kimura et al., 2010).

On the other hand, the polyglutamate (polyE) side-chains attached to glutamates in the C-terminus of α- and β-tubulins, were labeled with an antibody that recognizes side-chains of at least three consecutive glutamate residues (Shang et al., 2002). In the presence of active hCCP6, the signal of the PolyE antibody disappeared only from a band that displayed the typical slightly lower mobility in SDS-PAGE, characteristic of the α-tubulin family (Figure 3.19A). Further experiments carried out in our laboratory, with a modified SDS-PAGE methodology to improve the separation of α- and β-tubulin families followed by Western blot with anti-polyE antibody confirmed that hCCP6 acts specifically on α-tubulin. This is the first report in which hCCP6 activity is specifically linked to one family of tubulins.

The activity displayed by hCCP6 was confirmed by immunofluorescence staining of HEK 293T cells over-expressing hCCP6, using either anti-Δ2 tubulin or anti-polyE antibodies (Figures 2B and C). The increasing in Δ2 tubulin levels was clearly seen as a strong increased signal, specifically in the cells that over-expressed hCCP6 (Figure 3.19B). On the other hand, the signal of the polyE antibody decreased only moderately in the cells over-expressing hCCP6. This result is consistent with the observation in Western blot that only the polyE chains of α-tubulins are deglutamilated by hCCP6 (Figure 3.19C).

The activity of a CCP on a specific family of tubulins has not been reported before. The study of the activity of CCP1 on tubulins by mass spectrometry showed that this enzyme can shorten glutamates side chains from both, α- and β-tubulins (Berezniuk et al., 2012). The specificity of hCCP6 could be a regulatory mechanism for enzymes sharing substrate within the same cellular environment. The precise intervention of each CCP could be influenced by the relative expression of the CCPs genes, the subcellular localization, the longitude of polyE chains and the specific activity toward a family of tubulins, among others.

These regulatory factors, in turn, have important implications in the regulation of microtubule functions. Recent advances have evidenced the importance of tubulin PTMs in the regulation of the interactions and function of microtubule-associated proteins (MAPs)

and molecular motors. Polyglutamylation has been shown to enhance the microtubule-severing activity of spastin and katanin (Lacroix et al., 2010) and MAP binding (Bonnet et al., 2001). A recent publication (Sirajuddin et al., 2014) shows that tubulin isotypes are also important for the interactions of MAPs and motors with microtubules. For example, kinesin-1 and dynein motilities are regulated primarily by the C-terminal tail of β-tubulin (TUBB2A isotype), whereas kinesin-2 and kinesin-13 show more complex regulation involving C-terminal tail of both α- and β-tubulins (TUBA1A/TUBB2A isotypes).

3.3.2. Enzymatic activity of hCCP6 on TRAFD1 and HMGB3

CCPs not only display deglutamylating activity toward tubulins but also are able to hydrolyze the C-terminal encoded glutamates from other proteins. Rogowski et al, (2010) identified two substrates of mouse CCP1, myosin light chain kinase 1 (MLCK1) and telokin, a protein corresponding to the C-terminal fragment of MLCK1. Mouse CCP4 and CCP6 have also been report to remove all but one of the C-terminal glutamates from MLCK1 and telokin (Rogowski et al., 2010). A recent study identified seven potential substrates of human CCP1 from C-terminal Combined FRActional Diagonal Chromatography (C-terminal COFRADIC) analysis and one from shotgun experiments (Tanco et al., 2014, submitted). Two of these substrates, protein High Mobility Group Protein B3 (HMGB3) and TRAF-type zinc finger domain-containing protein 1 (TRAFD1) were further validated by affinity purification and Western blot analysis. CCP1 was able to removed the six Glu in the C-terminus of TRAFD1 (GAGDA↓EEEEEE-COOH) and 15 acidic residues from the C-terminus of HMGB3 (EEEDE↓EEEEEEEEEEEEDE-COOH) (Tanco et al., 2014, submitted).

CCPs may have other common substrates besides tubulins, therefore the ability of hCCP6 to hydrolyze the acidic C-terminus of TRAFD1 and HMGB3 was evaluated. The genes encoding proteins TRAFD1 and HMGB3 were cloned into vector pTriEx6 for production of N-terminal Strep-tag and HA-tag fusion proteins. The expression vector encoding for TRAFD1 was transfected alone or co-transfected with plasmid encoding for hCCP6 in HEK 293F cells. The same strategy was followed for protein HMGB3 and transfections were done in this case in HEK 293T cells. After 72 h of transfection, cells were lysed, and the cleared protein extracts were analyzed by SDS-PAGE followed by Western blot with anti-HA antibody.
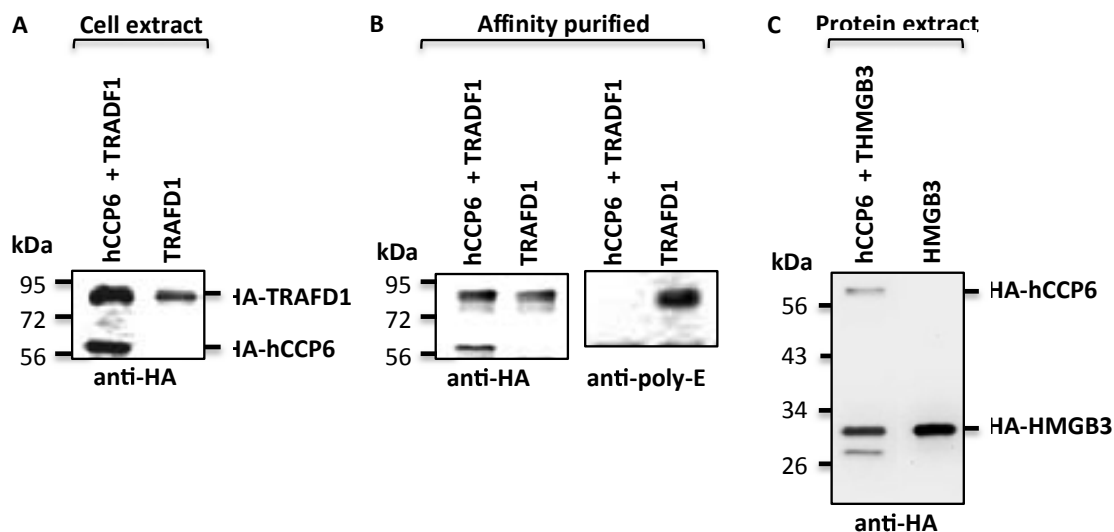
Figure 3.20. Co-expression of hCCP6 with TRAFD1 or HMGB3. A. HEK 293F cells were transfected with a plasmid expressing TRAFD1 alone or co-transfected with a plasmid expressing hCCP6 and equal amounts of each cell extract were analyzed by Western blot using anti-HA antibody B. Analysis of affinity-purified TRAFD1. Purification of TRAFD1 from cell extract of HEK 293F transfected with a plasmid expressing TRAFD1 alone or co-expressed with a plasmid expressing hCCP6. Equal amounts of eluted fractions were separated by SDS-PAGE followed by Western blot with anti-HA and anti-polyE antibodies. C. HEK 293T cells were transfected with a plasmid expressing HMGB3 alone or co-transfected with a plasmid expressing hCCP6 and equal amounts of each cell extract were analyzed by Western blot using anti-HA antibody.

Proteins TRAFD1 and hCCP6 are clearly detected in the Western blot of HEK 293F cells over-expressing both HA-tag fusion proteins (Figure 3.20A). Nevertheless, no difference in the electrophoretic mobility of TRAFD1, indicative of an enzymatic processing, was observed with anti-HA antibody. Furthermore, the anti-polyE antibody was not sensitive enough to detect TRAFD1 thus it was decided to obtain a sample enriched in the protein by partial purification via its N-terminal Strep-tag. HEK 293F cell extracts were loaded into a Strep-Tactin column and eluted fractions containing the proteins were analyzed by Western blot with anti-HA and anti-polyE antibodies (Figure 3.20B). Proteins hCCP6 and TRAFD1 effectively co-purified as was evidenced by the corresponding bands in the Western blot with the anti-HA antibody. The anti-polyE antibody recognizes the C-terminal acidic tail of TRAFD1, when overexpressed alone but not in the presence of active hCCP6. This means that hCCP6 is able to hydrolyze the glutamates from the C-terminus of TRAFD1 (Figure 3.20B). However, it is unknown whether TRAFD1 processing takes place in the cellular environment or during the purification process due to co-purification of the proteins.

The production of HA-tagged hCCP6 and HMBG3 was also detected in HEK 293T cells (Figure 3.20C). An additional band, with lower electrophoretic mobility, was observed which

is similar to the Western blot pattern obtained in CCP1 substrate validation experiments (Tanco et al., 2014, submitted). In this study the authors demonstrates that CCP1 processes the C-terminus of HMGB3 generating a fragment of 23,377 ± 11.67 Da. Apparently, hCCP6 is able to perform a similar activity on HMGB3.

The activity of hCCP6 towards TRAFD1 and HMGB3 was further evaluated *in vitro* with the proteins independently produced in HEK 293F cells and purified by affinity chromatography in Strep-Tactin column. hCCP6 was incubated overnight, at 37 °C with purified TRAFD1 or HMGB3 and an aliquot of the reaction mixture was analyzed by SDS-PAGE followed by Western blot. Control reactions, in the presence of the CP inhibitor ortho-phenanthroline, were performed in parallel (Figures 3.21A, B).

The *in vitro* assay confirmed the activity of hCCP6 towards the glutamic acid patch in the C-terminus of TRAFD1. The reaction was inhibited by ortho-phenanthroline, demonstrating that the processing is directly linked to the deglutamylating activity of hCCP6 (Figure 3.21A).

The result is less clear in the case of HMGB3 (Figure 3.21B). There is a basal degradation of HMGB3 reflected in an additional band in the reaction containing the CP inhibitor. The same band can be seen when hCCP6 and HMGB3 were incubated without inhibitor making it difficult to confirm the specific activity of hCCP6 toward HMGB3.

The C-terminal of HMGB3 contains two residues of Asp and until now the activity of hCCP6 has been described for proteins with Glu at the C-terminus (Rogowski et al., 2010). It is possible that the activity *in vitro* cannot be detected because hCCP6 has low or no specificity for aspartates and acts in coordination with another protease to trim the C-terminal of HMGB3 in the cell. CCP1 can hydrolyze aspartates from the C-termini of proteins although less efficiently than glutamates (Tanco et al., 2014, submitted).
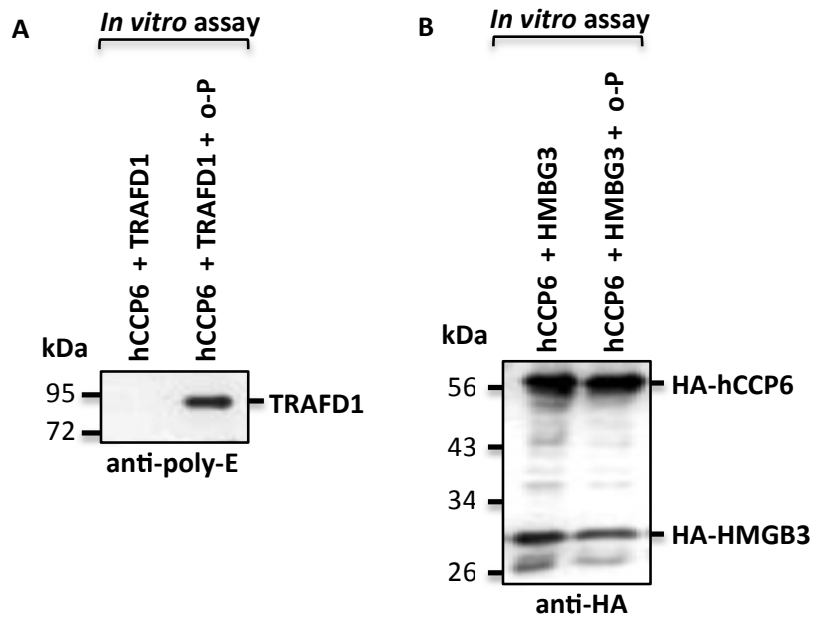
Figure 3.21. *In vitro* **assays with affinity purified Strep-tagged hCCP6, TRAFD1 and HMGB3 proteins. hCCP6 was incubated overnight, at 37 °C with TRAFD1 (A) or HMGB3 (B) in presence or absence of the CP inhibitor o-phenanthroline (o-P). Aliquots of the reaction were analyzed by SDS-PAGE followed by Western blot using anti-polyE (A) or anti-HA (B) antibodies.**

All together, hCCP6 was able to trim the acidic C-termini of TRAFD1 and HMGB3 in HEK 293 cells under over-expression conditions. The *in vitro* assays confirmed the deglutamylating activity of hCCP6 on TRAFD1. Further experiments should be performed to demonstrate whether hCCP6 alone can hydrolyze the Glu and Asp in the C-terminal of HMGB3 and to determine whether TRAFD1 and HMGB3 are physiological substrates of hCCP6.

3.3.3. Inhibition of hCCP6 by known carboxypeptidase inhibitors

The interaction of CCPs with known CP inhibitors of the M14A and B subfamilies has been poorly described. CCPs are zinc-dependent enzymes, and thereby general divalent cation chelators like EDTA and ortho-phenantroline have been used to inhibit CCPs *in vitro* (Berezniuk et al., 2012; Rodriguez de la Vega et al., 2007; Rogowski et al., 2010). Additionally, it has been proposed that the Retinoic Acid Receptor RESponder 1 (RARRES1), a mammalian endogenous CP inhibitor, might inhibit human CCP2 (Sahab et al., 2011). However, the latter result has not been confirmed independently.

In the present work, the interaction of hCCP6 with a more comprehensive group of CP inhibitors, including chelating agents and protein inhibitors was analyzed. For this purpose,

in vitro assays containing Strep-tag affinity-purified hCCP6, purified porcine brain tubulin as substrate and the corresponding inhibitor were performed. The following inhibitors were tested: a) chelating agents, EDTA and ortho-phenantroline; b) exogenous recombinant CP inhibitors, TCI (Tick Carboxypeptidase Inhibitor), ACI (*Ascaris* Carboxypeptidase Inhibitor), LCI (Leech Carboxypeptidase Inhibitor), PCI (Potato Carboxypeptidase Inhibitor) and NvCI (*Nerita versicolor* Carboxypeptidase Inhibitor); and c) the mammalian endogenous recombinant inhibitor latexin (Arolas et al., 2005a; Covaleda et al., 2012; Pallares et al., 2005; Reverter et al., 2000; Sanglas et al., 2009; Villanueva et al., 1998). Reaction mixtures were incubated 5 h, at 37 °C and proteins were separated by SDS-PAGE. The hydrolysis of polyglutamate side-chains of tubulin was analyzed by Western blot with anti-polyE antibody. Control reactions without inhibitor were performed to establish maximal (+hCCP6) or no activity (-hCCP6) of the enzyme. The results are shown in Figure 3.22.
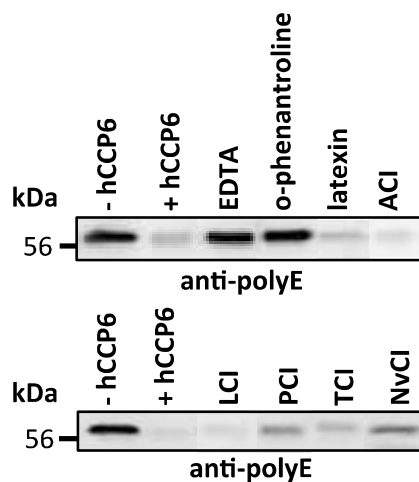
Figure 3.22. Interaction of hCCP6 with carboxypeptidase (CP) inhibitors. Protein hCCP6 (5 µl of fraction eluted from Strep-tag affinity chromatography) was incubated 5 h, at 37 °C with 0,625 µg of purified porcine brain tubulin and CP inhibitors in activity buffer. Identical reactions without inhibitor were done with (+hCCP6) or without (-hCCP6) hCCP6. Inhibitors used were: 10 mM EDTA, 10 mM o-phenantroline, 0.01 mM latexin, 0.01 mM ACI, 0.033 mM LCI, 0.1 mM PCI, 0.1 mM TCI and 0.1 mM NvCI. Aliquots of the reactions were analyzed by SDS-PAGE 10% followed by Western blot using anti-polyE tubulin antibody. Representative image of one replica of the reactions.

The chelating agents, EDTA and ortho-phenantroline completely inhibited hCCP6 as the antibody signal is equivalent to that obtained in the absence of the enzyme. This result is expected for MCPs and is in agreement with previous studies with CCPs (Berezniuk et al., 2012; Rodriguez de la Vega et al., 2007; Rogowski et al., 2010).

The endogenous inhibitor latexin does not inhibit hCCP6 at a concentration of 0.01 mM (Figure 3.22). Previously, we have observed that the bacterial ortholog PaCCP was not able

to bind latexin in non-denaturing PAGE experiments. Latexin is a potent inhibitor of M14A MCPs with equilibrium dissociation constant (Ki) in the range of 1-4 nM (Pallares et al., 2005) which is 10,000 times lower than the concentration used in this study. This protein inhibitor shares sequence/structural similarities with RARRES1 who have been proposed to inhibit human CCP2 (Sahab et al., 2011). The authors found CCP2 in complex with RARRES1 and proposed a functional relationship based on the effects generated by over-expression or silencing of these genes on the levels of tyrosinated α-tubulin in HEK 293 cells (Sahab et al., 2011). Latexin, and probably also RARRES1, follows a mechanism of inhibition based on interaction through a large surface with the enzymes and the blockage of the active-site by an inhibitory loop (Pallares et al., 2005). Interestingly, except hCCP6, the other human CCPs contain extra domains, in addition to the conserved N- terminal and CP domains (Rodriguez de la Vega et al., 2007), which might be involved in interactions with other proteins. These differences in the domain architectures could account for the differences in the interaction with endogenous CP inhibitors between CCP2 and CCP6.

In a close agreement, exogenous proteinaceous inhibitors ACI (from Ascaris) and LCI (from Leech), at concentrations of 0.01 mM and 0.033 mM, respectively, did not inhibit the activity of hCCP6. Other proteinaceus of MCPs like PCI (from potato), TCI (from Tick) and NvCI (from *Nerita versicolor*), at higher concentration (0.1 mM), seem to slightly inhibit hCCP6, reflected in a small increased signal of the anti-polyE antibody (Figure 3.22). These inhibitors competitively inhibit M14A MCPs with nanomolar affinities and NvCI displays a stronger inhibition in the picomolar range (Arolas et al., 2007; Covaleda et al., 2012). It is clear that they are not good inhibitors for CCPs as only a mild effect can be seen at much higher concentrations than that needed for inhibition of M14A MCPs. It is worth remembering that bacterial PaCCP did not form stable complexes with inhibitors PCI, NvCI and TCI. The mechanism of inhibition of exogenous inhibitors relies on the interaction of the C-terminal tail with the active site groove of the CP mimicking substrate binding (Arolas et al., 2005a; Reverter et al., 2000). A possible explanation for the lack of inhibition or a poor one, could be that these inhibitors are not efficiently fixed in the active-site of CCPs (extended enough) as it was described for non-proteinaceous inhibitors GEMSA and BZS in co-crystallization experiments with PaCCP (section 3.1.6). Besides the fixation in the active-site, additional interactions that confer stability to the complex could be missing.

3.3.4. Interactomic studies

The ability of hCCP6 to trim proteins other than tubulins suggests that it could be involved in cellular processes yet unexplored. Hence, an interactomic approach was used to identify potential functional partners of hCCP6 combining immunoprecipitation coupled to mass spectrometry. Initially, an attempt to immunoprecipitate the endogenous enzyme with anti-hCCP6 antibody in HEK 293F cells was made, but the protein was not detected probably due to low or no gene expression in this cell line. Instead, the HA-tagged hCCP6, over-expressed in HEK 293F cells, was captured with anti-HA antibody and the interacting partners were identified by nano-ESI-MS/MS.

Magnetic beads coated with protein G (protein G-beads) were used for immunoprecipations assays. First it was used the direct method (Figure 2.2A) in which anti-HA antibodies were cross-linked to beads and then incubated with the soluble fraction of HEK 293F cells, over-expressing hCCP6 or transfected with the empty vector (background control). After several washes the proteins were eluted from the beads and identified by mass spectrometry. The list of potential interactors of hCCP6, after removal of proteins identified in the background control is present in Table 3.4.

Half of the proteins identified are related to the microtubule cytoskeleton like molecular motors and modifying/organizing proteins. This was expected since tubulins are substrates of hCCP6. Proteins of actin cytoskeleton, like myosins, were also found. Other proteins identified are involved in the regulation of the cell cycle and cell division that also related with the microtubules cytoskeleton. The only exception is the transcriptional regulator Zinc finger protein 768 that probably resides in the nucleus (according to general annotations in UNIPROT database). hCCP6 should not interact *in vivo* with nuclear proteins since it was localized in the cytoplasm of HELA cells (Rodriguez de la Vega Otazo et al., 2012).

Immunoprecipitation experiments were also performed by the indirect method (Figure 2.2B) in which the anti-HA antibody is first incubated with HEK 293F cell lysates and the antibody-antigen complex is then captured by adding protein G-beads to the sample. HEK 293F cells, transfected with the empty vector were used as background control. The results are listed in Table 3.5.

Ribosomal proteins were abundant in controls and samples in both methods and, therefore, were removed from the final list of results. 40S ribosomal protein S9 was identified as one of the natural substrates of human CCP1 (Tanco et al., 2014, submitted) raising the possibility to a functional relationship between CCPs and ribosomal proteins. In the case that hCCP6 truly interacts with ribosomal proteins cannot be identified by the methods used here.

Noteworthy, the indirect method allowed the identification of a larger number of potential interacting partners of hCCP6 than the direct method. Once again, cytoskeletal proteins including actin, myosin, dynein 1 and MAP7 were identified (Table 3.5).

Tabla 3.4. Proteins identified in immunoprecipitations experiments by the direct method.* Protein score of one replica. Values greater than 100 are significant ($p<0.05$).

UniProt	Gene	Protein name	Protein Score*	Function (UniProt database)
P35579	MYH9	Myosin-9	307	Appears to play a role in cytokinesis, cell shape, secretion and capping. During cell spreading, plays an important role in cytoskeleton reorganization, focal contacts formation and lamellipodial retraction.
P14649	MYL6B	Myosin light chain 6B	123	Regulatory light chain of myosin. Does not bind calcium.
Q13409	DYNC1I2	Cytoplasmic dynein 1 intermediate chain 2	107	Non-catalytic accessory components of the cytoplasmic dynein 1 complex that are thought to be involved in linking dynein to cargos and to adapter proteins that regulate dynein function.
Q9Y6G9	DYNC1LI1	Cytoplasmic dynein 1 light intermediate chain 1	103	Non-catalytic accessory components of the cytoplasmic dynein 1 complex that are thought to be involved in linking dynein to cargos and to adapter proteins that regulate dynein function.
Q3KQU3	MAP7D1	MAP7 domain-containing protein 1	101	Microtubule cytoskeleton organization.
O75122	CLASP2	CLIP-associating protein 2	204	Microtubule plus-end tracking protein that promotes the stabilization of dynamic microtubules. Stabilizing function at the kinetochore which is essential for the bipolar alignment of chromosomes on the mitotic spindle
Q96N16	JAKMIP1	Janus kinase and microtubule-interacting protein 1	126	May play a role in the microtubule-dependent transport of the GABA-B receptor. May play a role in JAK1 signaling and regulate microtubule cytoskeleton rearrangements.
P51610	HCFC1	Host cell factor 1	113	Involved in control of the cell cycle. Involved in the activation and repression of transcription of several genes. May be involved in acetylation of nucleosomal histone H4 on several lysine residues.
Q14980	NUMA1	Nuclear mitotic apparatus protein 1	188	Non-mitotic structural role in the nuclear matrix. Required for maintenance and establishment of the mitotic spindle poles. May be involved in

UniProt	Gene	Protein name	Protein Score	Function
				coordination of the alignment of the mitotic spindle to the cellular polarity axis.
O43175	PHGDH	D-3-phosphoglycerate dehydrogenase	131	G1 to G0 transition. Cellular amino acid biosynthetic process. Brain development. Metabolic processes. Glial cell development. Neural tube development. Neuron projection development. Regulation of gene expression. Spinal cord development.
P48634	PRRC2A	Protein PRRC2A	114	May play a role in the regulation of pre-mRNA splicing.
Q9H5H4	ZNF768	Zinc finger protein 768	104	May be involved in transcriptional regulation.

Tabla 3.5. Proteins identified in immunoprecipitations experiments with indirect method.* Protein score of one replica. Values greater than 100 are significant (*p*<0.05). ‡ Proteins also identified by the direct method (Table 3.4).

UniProt	Gene	Protein name	Protein Score*	Function (UniProt database)
P68032	ACTC1	Actin, alpha cardiac muscle 1	961	Actins are involved in various types of cell motility.
P21333	FLNA	Filamin-A	180	Promotes orthogonal branching of actin filaments and links actin filaments to membrane glycoproteins. Anchors various transmembrane proteins to the actin cytoskeleton and serves as a scaffold for a wide range of cytoplasmic signaling proteins. Tethers cell surface-localized furin, modulates its rate of internalization and directs its intracellular trafficking.
P35579	MYH9	Myosin-9‡	2000	Appears to play a role in cytokinesis, cell shape, secretion and capping. During cell spreading, plays an important role in cytoskeleton reorganization, focal contacts formation and lamellipodial retraction.
P60660	MYL6	Myosin light polypeptide 6	291	Regulatory light chain of myosin. Does not bind calcium.
P35580	MYH10	Myosin-10	213	Appears to play a role in cytokinesis, cell shape, and secretion and capping. Involved in the stabilization of type I collagen mRNAs for CO1A1 and CO1A2. During cell spreading, plays an important role in cytoskeleton reorganization, focal contacts formation, and lamellipodial extension; this function is mechanically antagonized by MYH9.
Q14204	DYNC1H1	Cytoplasmic dynein 1 heavy chain 1	1089	Cytoplasmic dynein 1 acts as a motor for the intracellular retrograde motility of vesicles and organelles along microtubules.
Q13409	DYNC1I2	Cytoplasmic dynein 1 intermediate chain 2‡	163	The intermediate chains mediate the binding of dynein to dynactin via its 150 kDa component DCNT1. Involved in membrane-transport, such as Golgi apparatus, late endosomes and lysosomes.
Q3KQU3	MAP7D1	MAP7 domain-containing protein 1‡	105	Microtubule cytoskeleton organization.

P49368	CCT3	T-complex protein 1 subunit gamma	857	Molecular chaperone; assists the folding of proteins upon ATP hydrolysis. As part of the BBS/CCT complex may play a role in the assembly of BBSome, a complex involved in ciliogenesis regulating transports vesicles to the cilia. Known to play a role, in vitro, in the folding of actin and tubulin.
P17987	TCP1	T-complex protein 1 subunit alpha	778	
Q99832	CCT7	T-complex protein 1 subunit eta	810	
P40227	CCT6A	T-complex protein 1 subunit zeta	551	
P48643	CCT5	T-complex protein 1 subunit epsilon	274	
P50991	CCT4	T-complex protein 1 subunit delta	294	
P50990	CCT8	T-complex protein 1 subunit theta	107	
P50454	SERPINH1	Serpin H1	201	Could be involved as a chaperone in the biosynthetic pathway of collagen.
O95816	BAG2	BAG family molecular chaperone regulator 2	102	Inhibits the chaperone activity of HSP70/HSC70 by promoting substrate release.
P35998	PSMC2	26S protease regulatory subunit 7	195	ATP-dependent degradation of ubiquitinated proteins.
Q13200	PSMD2	26S proteasome non-ATPase regulatory subunit 2	224	Acts as a regulatory subunit of the 26 proteasome.
P62195	PSMC5	26S protease regulatory subunit 8	191	The regulatory complex confers ATP dependency and substrate specificity to the 26S complex.
P62191	PSMC1	26S protease regulatory subunit 4	121	
Q13618	CUL3	Cullin-3	652	Core component of multiple cullin-RING-based BCR E3 ubiquitin-protein ligase complexes
Q86VP6	CAND1	Cullin-associated NEDD8-dissociated protein 1	432	Key assembly factor of SCF E3 ubiquitin ligase complexes. Acts as a F-box protein exchange factor.
Q93008	USP9X	Probable ubiquitin carboxyl-terminal hydrolase FAF-X	210	Processing of ubiquitin precursors and of ubiquitinated proteins. Regulatory role of protein turnover. Regulates chromosome alignment and segregation in mitosis.
Q92616	GCN1L1	Translational activator GCN1	353	Translation activator and perform an EF3-related function on the ribosome.
Q9Y265	RUVBL1	RuvB-like 1	101	Possesses single-stranded DNA-stimulated ATPase and ATP-dependent DNA helicase activity. Component of the NuA4 histone acetyltransferase complex which is involved in transcriptional activation of select genes. Proposed core component of the chromatin remodeling INO80 complex which is involved in transcriptional regulation, DNA replication and probably DNA repair
Q9Y230	RUVBL2	RuvB-like 2	155	
Q8IZL8	PELP1	Proline-, glutamic acid- and leucine-rich protein 1	124	Coactivator of estrogen receptor-mediated transcription and a corepressor of other nuclear hormone receptors and sequence-specific transcription factors. Can facilitate ER non-genomic signaling via SRC and PI3K interaction in the cytosol. Plays a role in E2-mediated cell cycle progression by

				interacting with RB1. May have important functional implications in ER/growth factor cross-talk. Interacts with several growth factor signaling components. Involved in nuclear receptor signaling via its interaction with AR and NR3C1. May promote tumorigenesis via its interaction with and modulation of several oncogenes including SRC, PI3K, STAT3 and EGFR. Plays a role in cancer cell metastasis via its ability to modulate E2-mediated cytoskeleton changes and cell migration via its interaction with SRC and PI3K. Functions as the key stabilizing component of the Five Friends of Methylated CHTOP complex.
O94992	HEXIM1	Protein HEXIM1	209	RNA polymerase II transcription inhibitor. May also regulate NF-kappa-B, ESR1, NR3C1 and CIITA-dependent transcriptional activity.
P78527	PRKDC	DNA-dependent protein kinase catalytic subunit	791	Serine/threonine-protein kinase that acts as a molecular sensor for DNA damage. Also involved in modulation of transcription.
P33991	MCM4	DNA replication licensing factor MCM4	313	Acts as component of the MCM2-7 complex (MCM complex) which is the putative replicative helicase essential for 'once per cell cycle' DNA replication initiation and elongation in eukaryotic cells.
P33993	MCM7	DNA replication licensing factor MCM7	628	
P25205	MCM3	DNA replication licensing factor MCM3	132	
Q15366	PCBP2	Poly(rC)-binding protein 2	135	Major cellular poly(rC)-binding protein. Binds also poly(rU).
Q15365	PCBP1	Poly(rC)-binding protein 1	214	Single-stranded nucleic acid binding protein that binds preferentially to oligo dC.
P52701	MSH6	DNA mismatch repair protein Msh6	177	Component of the post-replicative DNA mismatch repair system.
Q14974	KPNB1	Importin subunit beta-1	387	Functions in nuclear protein import.
Q92973	TNPO1	Transportin-1	213	Functions in nuclear protein import as nuclear transport receptor
Q8TEX9	IPO4	Importin-4	150	Functions in nuclear protein import as nuclear transport receptor.
P55060	CSE1L	Exportin-2	481	Export receptor for importin-alpha.
O95373	IPO7	Importin-7	111	Functions in nuclear protein import.
Q13428	TCOF1	Treacle protein	142	May be involved in nucleolar-cytoplasmic transport. May play role in early embryonic development. May participate in certain stages of ribosome biogenesis.
P46060	RANGAP1	Ran GTPase-activating protein 1	105	GTPase activator for the nuclear Ras-related regulatory protein Ran.
P06748	NPM1	Nucleophosmin	159	Involved in diverse cellular processes such as ribosome biogenesis, centrosome duplication, protein chaperoning, histone assembly, cell proliferation, and regulation of tumor suppressors p53/TP53 and ARF.
O14654	IRS4	Insulin receptor substrate 4	150	Acts as an interface between multiple growth factor receptors and a complex network of intracellular signaling molecules containing SH2 domains. Plays a role in

				growth, reproduction and glucose homeostasis.
P55786	NPEPPS	Puromycin-sensitive aminopeptidase	190	Involved in proteolytic events essential for cell growth and viability. Plays a role in the antigen-processing pathway for MHC class I molecules.
Q9Y678	COPG1	Coatomer subunit gamma-1	166	The coatomer is a cytosolic protein complex that binds to dilysine motifs and reversibly associates with Golgi non-clathrin-coated vesicles, which further mediate biosynthetic protein transport from the ER, via the Golgi up to the trans Golgi network. In mammals, the coatomer influences the Golgi structural integrity, as well as the processing, activity, and endocytic recycling of LDL receptors. Required for limiting lipid storage in lipid droplets.
P53618	COPB1	Coatomer subunit beta	164	
P53621	COPA	Coatomer subunit alpha	238	
Q9UBF2	COPG2	Coatomer subunit gamma-2	103	
P08670	VIM	Vimentin	221	Vimentins are class-III intermediate filaments attached to the nucleus, ER, and mitochondria, either laterally or terminally.
Q07812	BAX	Apoptosis regulator BAX	112	Accelerates programmed cell death by binding to, and antagonizing the apoptosis repressor BCL2. Promotes activation of CASP3, and thereby apoptosis.
P17812	CTPS1	CTP synthase 1	310	Pyrimidine metabolism. CTP biosynthesis via de novo pathway.
Q9BSD7	NTPCR	Cancer-related nucleoside-triphosphatase	128	Hydrolyzes nucleoside diphosphates with lower efficiency.
P11586	MTHFD1	C-1-tetrahydrofolate synthase, cytoplasmic	112	One-carbon metabolism; tetrahydrofolate interconversion.
P07195	LDHB	L-lactate dehydrogenase B chain	227	Fermentation; pyruvate fermentation to lactate.
P49327	FASN	Fatty acid synthase	1019	Fatty acid synthetase catalyzes the formation of long-chain fatty acids.
P27708	CAD	CAD protein	624	"Fusion" protein encoding four enzymatic activities of the pyrimidine pathway.
P17858	PFKL	6-phosphofructokinase liver type	367	Catalyzes the third step of glycolysis.
P06737	PYGL	Glycogen phosphorylase, liver form	235	Important allosteric enzyme in carbohydrate metabolism.

Abbreviations. Endoplasmic Reticulum (ER); Bardet-Biedl Syndrome (BBS)

Several CCT (also known as TRiC) family of group II chaperonins were identified. These proteins are essential for BBSome (BBS: Bardet-Biedl Syndrome) assembly and normal ciliary function (Seo et al., 2010). The potential interaction of hCCP6 with CCT chaperonins is in agreement with the localization of hCCP6 in the basal bodies of ciliated NIH-3T3 cells and the hypothesis that might play a role in the *de novo* assembly of cilia and basal body (Rodriguez de la Vega Otazo et al., 2012). This potential interaction would be worth to be validated.

Proteins of the 26S proteasome complex and ubiquitin ligases were also identified as interactors of hCCP6. Adult Purkinje cell degeneration (pcd) mice brain showed altered levels of intracellular peptides (Berezniuk et al., 2010). It has been proposed that the proteasome activity and peptide levels could be direct affected by the defect in CCP1 activity or by the altered ciliary function caused by CCP1 defect (Berezniuk et al., 2013b). It possible that CCPs could be related with proteasome activity and this interaction must be further explored.

Subunits of coatomer complex, related with transport of vesicles through the Golgi network and Golgi structural integrity (Wang et al., 2010) were identified. Protein hCCP6 was localized in the Golgi apparatus in HELA cells (Rodriguez de la Vega Otazo et al., 2012) indicating a potential role in this membrane-bound compartment and supporting the study of the possible interaction with the coatomer complex.

The regulation of the expression of CCPs is a key aspect to better understand its functional range inside the cell. In this sense, some of the proteins identified in this study that modulate transcription of genes, could be regulators of hCCP6 expression and the validation of the interaction would be the first step to elucidate new regulatory pathways.

4. General Discussion

Our laboratory of Protein Engineering and Proteomics, at the Institute of Biotechnology and Biomedicine has extensively studied metallocarboxypeptidases (MCPs) from M14 family, in particular M14A and M14B subfamilies. In recent years, the laboratory has contributed crucially to the discovery and initial characterization of a new group of MCPs that have been classified into the M14D subfamily or cytosolic carboxypeptidases (CCPs; because of their main localization at the cytosol). CCPs conserved the catalytic and zinc-binding residues characteristic of the family (Rodriguez de la Vega et al., 2007) but also have several qualitative differences that have represented a challenge for its characterization (larger size and multidomain nature). The present thesis was aimed to gain insight into the structure, enzymatic properties and function of CCPs by using the CCP from *Pseudomonas aeruginosa* and the human CCP6 as models.

Chapter 1 was devoted to describe the structure of the CCP from *P. aeruginosa* (PaCCP). Structural studies of MCPs have been very important to identify key catalytic residues, to describe the interaction with substrates and inhibitors, to explore regulatory mechanisms and to design probes for the development of therapeutic strategies. The structure here described is the first report for a member of the M14D subfamily and certainly contributes to a better understanding of the properties of these MCPs. Structural studies were carried out firstly with a bacterial CCP because they have smaller size and bacterial proteins are typically easier to analyze and handle. PaCCP was produced and purified as a recombinant protein in *E. coli* and behaves as a monomer in solution. In contrast, CCPs from *Burkholderia* species are forming a tetramer in solution according to our results and others (Rimsa et al., 2014) evidencing a variability within bacterial CCPs.

A tridimensional crystal structure of PaCCP was derived at high resolution and comprises a novel β-sandwich N-terminal domain followed by the classical carboxypeptidase (CP) α/β-hydrolase domain (Vendrell et al., 2000). The structure and sequence of the N-terminal domain is unique in the whole M14 family and it is arranged in such way that partially blocks the entrance to the active site. This suggests a regulatory mechanism mediated by protein reorientation or restriction to allow the entrance of substrates. Cleavage by proteases, between the N- and CP-domains of PaCCP, has not been

observed suggesting that is not involved in a proteolytic activation as it happens in by enzymatic activation as M14A MCPs (Arolas et al., 2007; Vendrell et al., 2000).

The CP domain of PaCCP is smaller and more compact in comparison to other members of the M14 family (Vendrell et al., 2000) and conserves the topology of secondary structure elements. The zinc binding and catalytic residues are conserved and appear to be in optimal structural arrangement to carry out the expected CP activity. The exception is Tyr306, which is displaced with respect to the its equivalent canonical Tyr248 in the M14A and M14B families and would need structural rearrangement, probably upon binding to its substrate and/or an allosteric molecule, to be able to participate in the standard CP reaction.

The proposed mechanism for regulation of PaCCP activity is further support by the apparent absence of enzymatic activity against a variety of substrates, including standard chromogenic substrates used for M14A and M14B MCPs, peptides of different length and sequence and tubulins, the common substrates of mammalian CCPs (Rogowski et al., 2010; Tort et al., 2014, submitted). In the case of mammalian CCPs, it has been possible to detect the activity specifically towards natural protein substrates or, in the case of CCP1, also toward peptides mimicking the C-terminal of *in vivo* substrates (Rogowski et al., 2010; Tanco et al., 2014, submitted; Wu et al., 2012). The specific activity of CCPs, highly dependent of the nature of the substrate, appear to be a general feature of the subfamily and could function as a regulatory mechanism that prevent unspecific protein cleavage in the intracellular environment.

The substrate specificity of mammalian CCPs, as in M14A MCPs, probably is defined by the residue in position S1´ (Wu et al., 2012), which use to be an arginine, a fact that is compatible with the observed specificity for acidic residues. In PaCCP, the presence of a free cysteine in S1´ position is very unusual in the MCP family. PaCCP mutant in which cysteine in S1´ was replaced by arginine did not conferred specificity for glutamates chains of tubulins. Future research should be done to found the adequate substrates for this protein and determine the potential involvement of the free cysteine in S1´ position in the substrate recognition.

The determination of the structure of PaCCP, co-crystalized with organic CP inhibitors of M14A/B MCPs, showed that these molecules are not completely fixed in the specificity pocket of the active-site and probably would not be retained in competing conditions. This

is another element that supports the requirement of molecular rearrangement of PaCCP to display the enzimatic activity. This could also explain the observed lack of binding of PaCCP to proteinaceus inhibitors of M14A MCPs.

Finally, the crystal structure of PaCCP also was useful for the three-dimensional modeling of the N- and CP-domains of human CCPs. These domains conserved the global fold and its relative spatial position between distant evolutionary related organisms.

Chapter 2 describes microbiological phenotypes of *Pseudomonas aeruginosa* strain deficient in protein PaCCP (Δ*paccp*) in comparison with the wild-type strain (MPAO1). The mutant strain showed an increased protease secretion and swarming motility and decreased biofilm formation while swimming motility and resistance to some antibiotics were similar to MPAO1 strain. The most dramatic change was a total lack of twitching motility apparently due to low intracellular levels of PilA, insufficient to promote polymerization of the pili. The intracellular concentration of PaCCP is a critical parameter to exert its function in the modulation of PilA levels and therefore of twitching motility. PaCCP could be involved in the regulation of the expression of *pilA*, and probably of other genes that account for the multiples changes observed in the mutant strain.

The deficiency in well recognized virulence factors like motility and biofilm formation (Balasubramanian et al., 2013) in the mutant Δ*paccp*, resulted in a less virulence strain in a *C. elegans* slow-killing model. Further studies should be carry out in mammalian models to evaluate the importance of this protein in the human pathogenesis.

The study of potential interacting partners of PaCCP led to identification of several proteins among which some candidates related with motility, regulation systems, cell shape and polarity and pathogenicity are proposed for further validation. This analysis could be useful to fully understand the biological processes in which PaCCP is involved and would help to find a suitable substrate to characterize its enzymatic activity. The *in silico* evaluation of candidates, *ie* by molecular docking and dinamics could be a complementary and/or complementary approach to find substrates for PaCCP.

Chapter 3 describes some functional properties of CCP6, the human ortholog with the greatest similarity in domains architecture to bacterial CCPs (Rodriguez de la Vega et al., 2007). Human CCP6 (hCCP6) and a catalytic inactive mutant hCCP6 E425Q (equivalent to E270, the main catalytic residue in bovine CPA) were successfully produced in the HEK 293

mammalian cells system. The active protein, as was first demonstrated for the mouse orthologous one (Rogowski et al., 2010), hydrolyzes the encoded glutamate in the C-terminus of tubulin to generate Δ2 tubulin. Additionally, hCCP6 hydrolyzes from the C-terminal region of α-tubulins post-translationally added side chain glutamates specifically. This particular reaction preference (α-vs-β tubulin), has been described for enzymes responsible for the polyglutamylation of tubulin (inverse reaction) (Magiera and Janke, 2014) but never for CCPs. This finding unveils a new element in the mechanism for patterns generation of tubulin post-translational modifications (PTMs) known as the "tubulin code"(Magiera and Janke, 2014). In turn, this specificity could contribute to the regulation mechanisms of CCPs functions inside the cell. These enzymes posses overlapping enzymatic activity and its intervention must be highly coordinate. For instance CCPs have different intracellular localization depending on the cell cycle phase and the presence of cilia (Rodriguez de la Vega Otazo et al., 2012).

Cellular experiments with overexpressed proteins and *in vitro* assays using two partially purified cellular substrates of human CCP1 (Tanco et al., 2014, submitted) showed that hCCP6 can hydrolyzes glutamates from the C-terminal of other proteins besides tubulin. The identification of natural substrates of hCCP6 should be performed to know its whole cellular functional spectrum. The C-terminomics screening by mass spectrometry (Tanco et al., 2014, submitted) have proved to be a useful tool for this purpose and, additionally, it would allow to determine if hCCP6 hydrolyzes C-terminal aspartates, since our results are inconclusive in this sense.

On the other hand, inhibition test with chelating agents, exogenous carboxypeptidases inhibitors and the endogenous inhibitor latexin showed that hCCP6 is only strongly inhibit by chelating agents. Besides, interacting partners of hCCP6 were identified by immunoprecipitation of protein complexes followed by mass spectrometry identification. The most promising partners, suggesting links between hCCP6 and proposed functional mechanisms, are recommended for further validation. In agreement with its function in the modification of tubulin PTMs, proteins associated with the microtubules cytoskeleton were identified. Additionally identification of chaperonins involved in the BBsome assembly and normal ciliary function further supports the proposed role of hCCP6 in the assembly of cilia and basal bodies (Rodriguez de la Vega Otazo et al., 2012). Other identified proteins also

suggest that hCCP6 could be related to proteasome activity and the transport of vesicles through the Golgi network, the latter further hypothesis reinforced by its localization in the Golgi apparatus (Rodriguez de la Vega Otazo et al., 2012).

5. Conclusions

Chapter 1

1. The cytosolic carboxypeptidase from *Pseudomonas aeruginosa* MPAO1 (PaCCP) was purified to homogeneity behaving as monomer in solution.

2. The structure of PaCCP was solved by X-ray crystallography at high resolution (1.6 Å). PaCCP is folded in a novel β-sandwich N-terminal domain followed by the classical carboxypeptidase (CP) catalytic-like domain at C-terminus. The detailed structural disposition of the active-site residues is optimal to carry out the expected CP activity.

3. The partial blocking of the entrance to the active-site of PaCCP, the displacement of the residue Tyr306 (equivalent to the canonic Tyr248 in the M14A subfamily), known to be important for the catalysis in M14A metallocarboxypeptidases (MCPs) as well as the incomplete binding of M14A and B MCPs inhibitors in the active-site of PaCCP, suggest that a rearrangement is required for the enzyme to be active. This could explain the lack of a suitable substrate found to measure the CP activity of PaCCP. The structural rearrangement could be induced upon binding a very specific substrate(s) or a regulator. Alternatively, the produced protease might constitute an inactive precursor.

4. Known organic inhibitors of M14A and B MCPs, although co-crystallized with PaCCP, only bind partially to its potential active-site whilst known proteinaceus M14A MCPs inhibitors do not form complex with the enzyme. This highlights significant differences between MCPs of M14 subfamily.

5. The overall fold of PaCCP can be used as a general structural model for the M14D subfamily. The structural modeling of all human members of such family evidences a conserved global fold for the amino and carboxypeptidase domains and conserved positions of key residues among evolutionary distant members of M14D subfamily.

Chapter 2

1. Several microbiological phenotypes including protease secretion, biofilm formation, motilities and virulence in *Caenorhabditis elegans* model are altered in the *Pseudomonas aeruginosa* MPAO1 mutant strain deficient in PaCCP (Δ*paccp*), suggesting the potentiality of PaCCP as target for possible therapeutic and pharmaceutical goals.

2. The complete lack of twitching motility in such Δ*paccp* strain seems to be related to a defect in *pilA* gene expression, a gene that codifies the structural protein of bacterial pili. According to this, PaCCP could be involved in the regulation of gene expression in *P. aeruginosa* MPAO1 considering the wide range of modified phenotypes in Δ*paccp* strain. It is a relevant issue since twitching motility is a recognized virulence factor in *P. aeruginosa.*

3. A large number of potential interacting partners of PaCCP were identified by proteomics-like interactomic approach. Some of them, related with motility, regulation and virulence factors are proposed for further validation studies to demonstrate whether they are *bona fide* interacting partners. These studies could provide information on the functional mechanisms involving PaCCP in the bacterium.

Chapter 3

1. Human cytosolic carboxypeptidase 6 (hCCP6) has been selected as a distant homolog of PaCCP to establish comparisons. The HEK 293 mammalian cells are a suitable system for production and partial purification of such CCP and of its catalytic inactive form (E425Q mutant in a residue equivalent to E270, essential in M14A MCPs).

2. hCCP6, in HEK 293 cells overexpressing it, hydrolyzes the C-terminal encoded glutamate and the poly-glutamates chains added posttranslationally on the α-tubulin family.

3. Protein hCCP6 can hydrolyze C-terminal glutamates from TRAFD1 with semipurified components. It remains to determine if this trimming activity has a biological significance. Further experiments should be performed to evaluate whether hCCP6 can also hydrolyze C-terminal aspartates given that assays using the protein HMGB3,

which contains C-terminal aspartates and glutamates were not conclusive in this respect.

4. The activity of hCCP6 is completely inhibited by chelating agents but not by M14A MCPs proteinaceous inhibitors. These establish a functional difference between M14A and M14D subfamilies of MCPs.

5. A large number of potential interacting partners of hCCP6 were identified by proteomics-like interactomic approach. The validation of the interaction detected with proteins related with protein folding, protein degradation, vesicles trafficking and transcriptional regulation would provide information on novel and proposed mechanism involving hCCP6.

6. References

Aagaard, A., Listwan, P., Cowieson, N., Huber, T., Ravasi, T., Wells, C.A., Flanagan, J.U., Kellie, S., Hume, D.A., Kobe, B*., et al.* (2005). An inflammatory role for the mammalian carboxypeptidase inhibitor latexin: relationship to cystatins and the tumor suppressor TIG1. Structure *13*, 309-317.

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W*., et al.* (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr *66*, 213-221.

Alarcon, I., Evans, D.J., and Fleiszig, S.M. (2009). The role of twitching motility in Pseudomonas aeruginosa exit from and translocation of corneal epithelial cells. Invest Ophthalmol Vis Sci *50*, 2237-2244.

Allesen-Holm, M., Barken, K.B., Yang, L., Klausen, M., Webb, J.S., Kjelleberg, S., Molin, S., Givskov, M., and Tolker-Nielsen, T. (2006). A characterization of DNA release in Pseudomonas aeruginosa cultures and biofilms. Mol Microbiol *59*, 1114-1128.

Alonso-del-Rivero, M., Trejo, S.A., Reytor, M.L., Rodriguez-de-la-Vega, M., Delfin, J., Diaz, J., Gonzalez-Gonzalez, Y., Canals, F., Chavez, M.A., and Aviles, F.X. (2012). Tri-domain bifunctional inhibitor of metallocarboxypeptidases A and serine proteases isolated from marine annelid Sabellastarte magnifica. J Biol Chem *287*, 15427-15438.

Aloy, P., Companys, V., Vendrell, J., Aviles, F.X., Fricker, L.D., Coll, M., and Gomis-Ruth, F.X. (2001). The crystal structure of the inhibitor-complexed carboxypeptidase D domain II and the modeling of regulatory carboxypeptidases. J Biol Chem *276*, 16177-16184.

Anderson, R.D., Roddam, L.F., Bettiol, S., Sanderson, K., and Reid, D.W. (2010). Biosignificance of bacterial cyanogenesis in the CF lung. J Cyst Fibros *9*, 158-164.

Arolas, J.L., Lorenzo, J., Rovira, A., Castella, J., Aviles, F.X., and Sommerhoff, C.P. (2005a). A carboxypeptidase inhibitor from the tick Rhipicephalus bursa: isolation, cDNA cloning, recombinant expression, and characterization. J Biol Chem *280*, 3441-3448.

Arolas, J.L., Lorenzo, J., Rovira, A., Vendrell, J., Aviles, F.X., and Ventura, S. (2004). Secondary binding site of the potato carboxypeptidase inhibitor. Contribution to its structure, folding, and biological properties. Biochemistry *43*, 7973-7982.

Arolas, J.L., Popowicz, G.M., Lorenzo, J., Sommerhoff, C.P., Huber, R., Aviles, F.X., and Holak, T.A. (2005b). The three-dimensional structures of tick carboxypeptidase inhibitor in complex with A/B carboxypeptidases reveal a novel double-headed binding mode. J Mol Biol *350*, 489-498.

Arolas, J.L., Vendrell, J., Aviles, F.X., and Fricker, L.D. (2007). Metallocarboxypeptidases: emerging drug targets in biomedicine. Curr Pharm Des *13*, 349-366.

Asikyan, M.L., Kus, J.V., and Burrows, L.L. (2008). Novel proteins that modulate type IV pilus retraction dynamics in Pseudomonas aeruginosa. J Bacteriol *190*, 7022-7034.

Aviles, F.X., Vendrell, J., Guasch, A., Coll, M., and Huber, R. (1993). Advances in metallo-procarboxypeptidases. Emerging details on the inhibition mechanism and on the activation process. Eur J Biochem *211*, 381-389.

Bai, F., Li, Y., Xu, H., Xia, H., Yin, T., Yao, H., Zhang, L., Zhang, X., Bai, Y., Jin, S*., et al.* (2007). Identification and functional characterization of pfm, a novel gene involved in swimming motility of Pseudomonas aeruginosa. Gene *401*, 19-27.

Balasubramanian, D., Schneper, L., Kumari, H., and Mathee, K. (2013). A dynamic and intricate regulatory network determines Pseudomonas aeruginosa virulence. Nucleic Acids Res *41*, 1-20.

Barken, K.B., Pamp, S.J., Yang, L., Gjermansen, M., Bertrand, J.J., Klausen, M., Givskov, M., Whitchurch, C.B., Engel, J.N., and Tolker-Nielsen, T. (2008). Roles of type IV pili, flagellum-mediated motility and extracellular DNA in the formation of mature multicellular structures in Pseudomonas aeruginosa biofilms. Environ Microbiol *10*, 2331-2343.

Barraud, N., Schleheck, D., Klebensberger, J., Webb, J.S., Hassett, D.J., Rice, S.A., and Kjelleberg, S. (2009). Nitric oxide signaling in Pseudomonas aeruginosa biofilms mediates phosphodiesterase activity, decreased cyclic di-GMP levels, and enhanced dispersal. J Bacteriol *191*, 7333-7342.

Beatson, S.A., Whitchurch, C.B., Sargent, J.L., Levesque, R.C., and Mattick, J.S. (2002). Differential regulation of twitching motility and elastase production by Vfr in Pseudomonas aeruginosa. J Bacteriol *184*, 3605-3613.

Berezniuk, I., Lyons, P.J., Sironi, J.J., Xiao, H., Setou, M., Angeletti, R.H., Ikegami, K., and Fricker, L.D. (2013a). Cytosolic carboxypeptidase 5 removes alpha- and gamma-linked glutamates from tubulin. J Biol Chem *288*, 30445-30453.

Berezniuk, I., Sironi, J., Callaway, M.B., Castro, L.M., Hirata, I.Y., Ferro, E.S., and Fricker, L.D. (2010). CCP1/Nna1 functions in protein turnover in mouse brain: Implications for cell death in Purkinje cell degeneration mice. FASEB J *24*, 1813-1823.

Berezniuk, I., Sironi, J.J., Wardman, J., Pasek, R.C., Berbari, N.F., Yoder, B.K., and Fricker, L.D. (2013b). Quantitative peptidomics of Purkinje cell degeneration mice. PLoS One *8*, e60981.

Berezniuk, I., Vu, H.T., Lyons, P.J., Sironi, J.J., Xiao, H., Burd, B., Setou, M., Angeletti, R.H., Ikegami, K., and Fricker, L.D. (2012). Cytosolic carboxypeptidase 1 is involved in processing alpha- and beta-tubulin. J Biol Chem *287*, 6503-6517.

Bergmeyer, H.U., Gawehn, K. & Grassl, M. (1974). Methods of Enzymatic Analysis, Vol 1, 2nd edn (New York, Academic Press).

Bertrand, J.J., West, J.T., and Engel, J.N. (2010). Genetic analysis of the regulation of type IV pilus function by the Chp chemosensory system of Pseudomonas aeruginosa. J Bacteriol *192*, 994-1010.

Bitto, E., Bingman, C.A., Wesenberg, G.E., McCoy, J.G., and Phillips, G.N., Jr. (2007). Structure of aspartoacylase, the brain enzyme impaired in Canavan disease. Proc Natl Acad Sci U S A *104*, 456-461.

Bonnet, C., Boucher, D., Lazereg, S., Pedrotti, B., Islam, K., Denoulet, P., and Larcher, J.C. (2001). Differential binding regulation of microtubule-associated proteins MAP1A, MAP1B, and MAP2 by tubulin polyglutamylation. J Biol Chem *276*, 12839-12848.

Boyd, J.M., and Lory, S. (1996). Dual function of PilS during transcriptional activation of the Pseudomonas aeruginosa pilin subunit gene. J Bacteriol *178*, 831-839.

Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem *72*, 248-254.

Bradley, D.E. (1980). A function of Pseudomonas aeruginosa PAO polar pili: twitching motility. Can J Microbiol *26*, 146-154.

Breidenstein, E.B., Bains, M., and Hancock, R.E. (2012a). Involvement of the lon protease in the SOS response triggered by ciprofloxacin in Pseudomonas aeruginosa PAO1. Antimicrob Agents Chemother *56*, 2879-2887.

Breidenstein, E.B., and Hancock, R.E. (2013). Armand-Frappier outstanding student award -- role of ATP-dependent proteases in antibiotic resistance and virulence. Can J Microbiol *59*, 1-8.

Breidenstein, E.B., Janot, L., Strehmel, J., Fernandez, L., Taylor, P.K., Kukavica-Ibrulj, I., Gellatly, S.L., Levesque, R.C., Overhage, J., and Hancock, R.E. (2012b). The Lon protease is essential for full virulence in Pseudomonas aeruginosa. PLoS One *7*, e49123.

Brunger, A.T. (2007). Version 1.2 of the Crystallography and NMR system. Nat Protoc *2*, 2728-2733.

Burrows, L.L. (2012). Pseudomonas aeruginosa twitching motility: type IV pili in action. Annu Rev Microbiol *66*, 493-520.

Buske, P.J., and Levin, P.A. (2012). Extreme C terminus of bacterial cytoskeletal protein FtsZ plays fundamental role in assembly independent of modulatory proteins. J Biol Chem *287*, 10945-10957.

Cain, J.A., Solis, N., and Cordwell, S.J. (2014). Beyond gene expression: The impact of protein post-translational modifications in bacteria. J Proteomics *97*, 265-286.

Cezairliyan, B., Vinayavekhin, N., Grenfell-Lee, D., Yuen, G.J., Saghatelian, A., and Ausubel, F.M. (2013). Identification of Pseudomonas aeruginosa phenazines that kill Caenorhabditis elegans. PLoS Pathog *9*, e1003101.

Chet, I., and Mitchell, R. (1976). Ecological aspects of microbial chemotactic behavior. Annu Rev Microbiol *30*, 221-239.

Chevance, F.F., and Hughes, K.T. (2008). Coordinating assembly of a bacterial macromolecular machine. Nat Rev Microbiol *6*, 455-465.

Chiang, P., Habash, M., and Burrows, L.L. (2005). Disparate subcellular localization patterns of Pseudomonas aeruginosa Type IV pilus ATPases involved in twitching motility. J Bacteriol *187*, 829-839.

Choi, K.H., Kumar, A., and Schweizer, H.P. (2006). A 10-min method for preparation of highly electrocompetent Pseudomonas aeruginosa cells: application for DNA fragment transfer between chromosomes and plasmid transformation. J Microbiol Methods *64*, 391-397.

CLSI (2012). Performance standards for antimicrobial susceptibility testing; 22th informational supplement (Wayne, P).

Coll, M., Guasch, A., Aviles, F.X., and Huber, R. (1991). Three-dimensional structure of porcine procarboxypeptidase B: a structural basis of its inactivity. EMBO J *10*, 1-9.

Comolli, J.C., Hauser, A.R., Waite, L., Whitchurch, C.B., Mattick, J.S., and Engel, J.N. (1999). Pseudomonas aeruginosa gene products PilT and PilU are required for cytotoxicity in vitro and virulence in a mouse model of acute pneumonia. Infect Immun *67*, 3625-3630.

Cornforth, D.M., Popat, R., McNally, L., Gurney, J., Scott-Phillips, T.C., Ivens, A., Diggle, S.P., and Brown, S.P. (2014). Combinatorial quorum sensing allows bacteria to resolve their social and physical environment. Proc Natl Acad Sci U S A.

Covaleda, G., del Rivero, M.A., Chavez, M.A., Aviles, F.X., and Reverter, D. (2012). Crystal structure of novel metallocarboxypeptidase inhibitor from marine mollusk Nerita versicolor in complex with human carboxypeptidase A4. J Biol Chem *287*, 9250-9258.

Craig, L., Pique, M.E., and Tainer, J.A. (2004). Type IV pilus structure and bacterial pathogenicity. Nat Rev Microbiol *2*, 363-378.

Craig, L., Volkmann, N., Arvai, A.S., Pique, M.E., Yeager, M., Egelman, E.H., and Tainer, J.A. (2006). Type IV pilus structure by cryo-electron microscopy and crystallography: implications for pilus assembly and functions. Mol Cell *23*, 651-662.

de Bentzmann, S., and Plesiat, P. (2011). The Pseudomonas aeruginosa opportunistic pathogen and human infections. Environ Microbiol *13*, 1655-1665.

de Castro, E., Sigrist, C.J., Gattiker, A., Bulliard, V., Langendijk-Genevaux, P.S., Gasteiger, E., Bairoch, A., and Hulo, N. (2006). ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. Nucleic Acids Res *34*, W362-365.

Deiteren, K., Surpateanu, G., Gilany, K., Willemse, J.L., Hendriks, D.F., Augustyns, K., Laroche, Y., Scharpe, S., and Lambeir, A.M. (2007). The role of the S1 binding site of carboxypeptidase M in substrate specificity and turn-over. Biochim Biophys Acta *1774*, 267-277.

Deziel, E., Lepine, F., Milot, S., and Villemur, R. (2003). rhlA is required for the production of a novel biosurfactant promoting swarming motility in Pseudomonas aeruginosa: 3-(3-hydroxyalkanoyloxy)alkanoic acids (HAAs), the precursors of rhamnolipids. Microbiology *149*, 2005-2013.

Divakaruni, A.V., Baida, C., White, C.L., and Gober, J.W. (2007). The cell shape proteins MreB and MreC control cell morphogenesis by positioning cell wall synthetic complexes. Mol Microbiol *66*, 174-188.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr D Biol Crystallogr *66*, 486-501.

Estebanez-Perpina, E., Bayes, A., Vendrell, J., Jongsma, M.A., Bown, D.P., Gatehouse, J.A., Huber, R., Bode, W., Aviles, F.X., and Reverter, D. (2001). Crystal structure of a novel mid-gut procarboxypeptidase from the cotton pest Helicoverpa armigera. J Mol Biol *313*, 629-638.

Fazli, M., Almblad, H., Rybtke, M.L., Givskov, M., Eberl, L., and Tolker-Nielsen, T. (2014). Regulation of biofilm formation in Pseudomonas and Burkholderia species. Environ Microbiol.

Feinbaum, R.L., Urbach, J.M., Liberati, N.T., Djonovic, S., Adonizio, A., Carvunis, A.R., and Ausubel, F.M. (2012). Genome-wide identification of Pseudomonas aeruginosa virulence-related genes using a Caenorhabditis elegans infection model. PLoS Pathog *8*, e1002813.

Felber, J.P., Coombs, T.L., and Vallee, B.L. (1962). The mechanism of inhibition of carboxypeptidase A by 1,10-phenanthroline. Biochemistry *1*, 231-238.

Fernandez, D., Boix, E., Pallares, I., Aviles, F.X., and Vendrell, J. (2010). Analysis of a new crystal form of procarboxypeptidase B: further insights into the catalytic mechanism. Biopolymers *93*, 178-185.

Fernandez, L., Breidenstein, E.B., Song, D., and Hancock, R.E. (2012). Role of intracellular proteases in the antibiotic resistance, motility, and biofilm formation of Pseudomonas aeruginosa. Antimicrob Agents Chemother *56*, 1128-1132.

Fernandez-Gonzalez, A., La Spada, A.R., Treadaway, J., Higdon, J.C., Harris, B.S., Sidman, R.L., Morgan, J.I., and Zuo, J. (2002). Purkinje cell degeneration (pcd) phenotypes caused by mutations in the axotomy-induced gene, Nna1. Science *295*, 1904-1906.

Foley, J.H., Kim, P.Y., Mutch, N.J., and Gils, A. (2013). Insights into thrombin activatable fibrinolysis inhibitor function and regulation. J Thromb Haemost *11 Suppl 1*, 306-315.

Fraser, G.M., and Hughes, C. (1999). Swarming motility. Curr Opin Microbiol *2*, 630-635.

Friedman, L., and Kolter, R. (2004). Two genetic loci produce distinct carbohydrate-rich structural components of the Pseudomonas aeruginosa biofilm matrix. J Bacteriol *186*, 4457-4465.

Fulcher, N.B., Holliday, P.M., Klem, E., Cann, M.J., and Wolfgang, M.C. (2010). The Pseudomonas aeruginosa Chp chemosensory system regulates intracellular cAMP levels by modulating adenylate cyclase activity. Mol Microbiol *76*, 889-904.

Gallagher, L.A., and Manoil, C. (2001). Pseudomonas aeruginosa PAO1 kills Caenorhabditis elegans by cyanide poisoning. J Bacteriol *183*, 6207-6214.

Garcia-Castellanos, R., Bonet-Figueredo, R., Pallares, I., Ventura, S., Aviles, F.X., Vendrell, J., and Gomis-Rutha, F.X. (2005). Detailed molecular comparison between the inhibition mode of A/B-type carboxypeptidases in the zymogen state and by the endogenous inhibitor latexin. Cell Mol Life Sci *62*, 1996-2014.

Garcia-Saez, I., Reverter, D., Vendrell, J., Aviles, F.X., and Coll, M. (1997). The three-dimensional structure of human procarboxypeptidase A2. Deciphering the basis of the inhibition, activation and intrinsic activity of the zymogen. EMBO J *16*, 6906-6913.

Ghosh-Roy, A., Goncharov, A., Jin, Y., and Chisholm, A.D. (2012). Kinesin-13 and tubulin posttranslational modifications regulate microtubule growth in axon regeneration. Dev Cell *23*, 716-728.

Giltner, C.L., Nguyen, Y., and Burrows, L.L. (2012). Type IV pilin proteins: versatile molecular modules. Microbiol Mol Biol Rev *76*, 740-772.

Gitai, Z., Dye, N., and Shapiro, L. (2004). An actin-like gene can determine cell polarity in bacteria. Proc Natl Acad Sci U S A *101*, 8643-8648.

Glick, R., Gilmour, C., Tremblay, J., Satanower, S., Avidan, O., Deziel, E., Greenberg, E.P., Poole, K., and Banin, E. (2010). Increase in rhamnolipid synthesis under iron-limiting conditions influences surface motility and biofilm formation in Pseudomonas aeruginosa. J Bacteriol *192*, 2973-2980.

Gomis-Ruth, F.X. (2008). Structure and mechanism of metallocarboxypeptidases. Crit Rev Biochem Mol Biol *43*, 319-345.

Gong, H., Zhou, J., Liao, M., Hatta, T., Harnnoi, T., Umemiya, R., Inoue, N., Xuan, X., and Fujisaki, K. (2007). Characterization of a carboxypeptidase inhibitor from the tick Haemaphysalis longicornis. J Insect Physiol *53*, 1079-1087.

Gooderham, W.J., and Hancock, R.E. (2009). Regulation of virulence and antibiotic resistance by two-component regulatory systems in Pseudomonas aeruginosa. FEMS Microbiol Rev *33*, 279-294.

Grishin, A.M., Akparov, V., and Chestukhina, G.G. (2008). Structural principles of the broad substrate specificity of Thermoactinomyces vulgaris carboxypeptidase T--role of amino acid residues at positions 260 and 262. Protein Eng Des Sel *21*, 545-551.

Guasch, A., Coll, M., Aviles, F.X., and Huber, R. (1992). Three-dimensional structure of porcine pancreatic procarboxypeptidase A. A comparison of the A and B zymogens and their determinants for inhibition and activation. J Mol Biol *224*, 141-157.

Haiko, J., and Westerlund-Wikstrom, B. (2013). The role of the bacterial flagellum in adhesion and virulence. Biology (Basel) *2*, 1242-1267.

Harmsen, M., Yang, L., Pamp, S.J., and Tolker-Nielsen, T. (2010). An update on Pseudomonas aeruginosa biofilm formation, tolerance, and dispersal. FEMS Immunol Med Microbiol *59*, 253-268.

Harris, A., Morgan, J.I., Pecot, M., Soumare, A., Osborne, A., and Soares, H.D. (2000). Regenerating motor neurons express Nna1, a novel ATP/GTP-binding protein related to zinc carboxypeptidases. Mol Cell Neurosci *16*, 578-596.

Harvey, H., Habash, M., Aidoo, F., and Burrows, L.L. (2009). Single-residue changes in the C-terminal disulfide-bonded loop of the Pseudomonas aeruginosa type IV pilin influence pilus assembly and twitching motility. J Bacteriol *191*, 6513-6524.

Hass, G.M., Nau, H., Biemann, K., Grahn, D.T., Ericsson, L.H., and Neurath, H. (1975). The amino acid sequence of a carboxypeptidase inhibitor from potatoes. Biochemistry *14*, 1334-1342.

Hatch, R.A., and Schiller, N.L. (1998). Alginate lyase promotes diffusion of aminoglycosides through the extracellular polysaccharide of mucoid Pseudomonas aeruginosa. Antimicrob Agents Chemother *42*, 974-977.

Hemsley, A., Arnheim, N., Toney, M.D., Cortopassi, G., and Galas, D.J. (1989). A simple method for site-directed mutagenesis using the polymerase chain reaction. Nucleic Acids Res *17*, 6545-6551.

Henrichsen, J. (1972). Bacterial surface translocation: a survey and a classification. Bacteriol Rev *36*, 478-503.

Hinz, A., Lee, S., Jacoby, K., and Manoil, C. (2011). Membrane proteases and aminoglycoside antibiotic resistance. J Bacteriol *193*, 4790-4797.

Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K., and Pease, L.R. (1989). Site-directed mutagenesis by overlap extension using the polymerase chain reaction. Gene *77*, 51-59.

Hobbs, M., Collie, E.S., Free, P.D., Livingston, S.P., and Mattick, J.S. (1993). PilS and PilR, a two-component transcriptional regulatory system controlling expression of type 4 fimbriae in Pseudomonas aeruginosa. Mol Microbiol *7*, 669-682.

Hoge, R., Pelzer, A., Rosenau F., Wilhelm, S. (2010). Weapons of a pathogen: Proteases and their role in virulence of Pseudomonas aeruginosa, Vol 1 (FORMATEX).

Hoiby, N., Ciofu, O., Johansen, H.K., Song, Z.J., Moser, C., Jensen, P.O., Molin, S., Givskov, M., Tolker-Nielsen, T., and Bjarnsholt, T. (2011). The clinical impact of bacterial biofilms. Int J Oral Sci *3*, 55-65.

Hooper, N.M. (1994). Families of zinc metalloproteases. FEBS Lett *354*, 1-6.

Hourdou, M.L., Guinand, M., Vacheron, M.J., Michel, G., Denoroy, L., Duez, C., Englebert, S., Joris, B., Weber, G., and Ghuysen, J.M. (1993). Characterization of the sporulation-related gamma-D-glutamyl-(L)meso-diaminopimelic-acid-hydrolysing peptidase I of Bacillus sphaericus NCTC 9602 as a member of the metallo(zinc) carboxypeptidase A family. Modular design of the protein. Biochem J *292 (Pt 2)*, 563-570.

Irazoqui, J.E., Urbach, J.M., and Ausubel, F.M. (2010). Evolution of host innate defence: insights from Caenorhabditis elegans and primitive invertebrates. Nat Rev Immunol *10*, 47-58.

IUBMB (1992). Enzyme Nomenclature (San Diego, California,, Academic Press).

Jacobs, M.A., Alwood, A., Thaipisuttikul, I., Spencer, D., Haugen, E., Ernst, S., Will, O., Kaul, R., Raymond, C., Levy, R.*, et al.* (2003). Comprehensive transposon mutant library of Pseudomonas aeruginosa. Proc Natl Acad Sci U S A *100*, 14339-14344.

Janke, C., and Bulinski, J.C. (2011). Post-translational regulation of the microtubule cytoskeleton: mechanisms and functions. Nat Rev Mol Cell Biol *12*, 773-786.

Jarrell, K.F., and McBride, M.J. (2008). The surprisingly diverse ways that prokaryotes move. Nat Rev Microbiol *6*, 466-476.

Jimenez, P.N., Koch, G., Thompson, J.A., Xavier, K.B., Cool, R.H., and Quax, W.J. (2012). The multiple signaling systems regulating virulence in Pseudomonas aeruginosa. Microbiol Mol Biol Rev *76*, 46-65.

Jing, C., El-Ghany, M.A., Beesley, C., Foster, C.S., Rudland, P.S., Smith, P., and Ke, Y. (2002). Tazarotene-induced gene 1 (TIG1) expression in prostate carcinomas and its relationship to tumorigenicity. J Natl Cancer Inst *94*, 482-490.

Kabsch, W. (2010). Xds. Acta Crystallogr D Biol Crystallogr *66*, 125-132.

Kalinina, E., Biswas, R., Berezniuk, I., Hermoso, A., Aviles, F.X., and Fricker, L.D. (2007). A novel subfamily of mouse cytosolic carboxypeptidases. FASEB J *21*, 836-850.

Karuppiah, V., and Derrick, J.P. (2011). Structure of the PilM-PilN inner membrane type IV pilus biogenesis complex from Thermus thermophilus. J Biol Chem *286*, 24434-24442.

Keil, C., Maskos, K., Than, M., Hoopes, J.T., Huber, R., Tan, F., Deddish, P.A., Erdos, E.G., Skidgel, R.A., and Bode, W. (2007). Crystal structure of the human carboxypeptidase N (kininase I) catalytic domain. J Mol Biol *366*, 504-516.

Kim, H., and Lipscomb, W.N. (1990). Crystal structure of the complex of carboxypeptidase A with a strongly bound phosphonate in a new crystalline form: comparison with structures of other complexes. Biochemistry *29*, 5546-5555.

Kimura, Y., Kurabe, N., Ikegami, K., Tsutsumi, K., Konishi, Y., Kaplan, O.I., Kunitomo, H., Iino, Y., Blacque, O.E., and Setou, M. (2010). Identification of tubulin deglutamylase among Caenorhabditis elegans and mammalian cytosolic carboxypeptidases (CCPs). J Biol Chem *285*, 22936-22941.

Kindrachuk, K.N., Fernandez, L., Bains, M., and Hancock, R.E. (2011). Involvement of an ATP-dependent protease, PA0779/AsrA, in inducing heat shock in response to tobramycin in Pseudomonas aeruginosa. Antimicrob Agents Chemother *55*, 1874-1882.

Kirienko, N.V., Kirienko, D.R., Larkins-Ford, J., Wahlby, C., Ruvkun, G., and Ausubel, F.M. (2013). Pseudomonas aeruginosa disrupts Caenorhabditis elegans iron homeostasis, causing a hypoxic response and death. Cell Host Microbe *13*, 406-416.

Klausen, M., Heydorn, A., Ragas, P., Lambertsen, L., Aaes-Jorgensen, A., Molin, S., and Tolker-Nielsen, T. (2003). Biofilm formation by Pseudomonas aeruginosa wild type, flagella and type IV pili mutants. Mol Microbiol *48*, 1511-1524.

Kohler, T., Curty, L.K., Barja, F., van Delden, C., and Pechere, J.C. (2000). Swarming of Pseudomonas aeruginosa is dependent on cell-to-cell signaling and requires flagella and pili. J Bacteriol *182*, 5990-5996.

Koo, J., Tang, T., Harvey, H., Tammam, S., Sampaleanu, L., Burrows, L.L., and Howell, P.L. (2013). Functional mapping of PilF and PilQ in the Pseudomonas aeruginosa type IV pilus system. Biochemistry *52*, 2914-2923.

Kovach, M.E., Elzer, P.H., Hill, D.S., Robertson, G.T., Farris, M.A., Roop, R.M., 2nd, and Peterson, K.M. (1995). Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. Gene *166*, 175-176.

Krushkal, J., Juarez, K., Barbe, J.F., Qu, Y., Andrade, A., Puljic, M., Adkins, R.M., Lovley, D.R., and Ueki, T. (2010). Genome-wide survey for PilR recognition sites of the metal-reducing prokaryote Geobacter sulfurreducens. Gene *469*, 31-44.

Lacroix, B., van Dijk, J., Gold, N.D., Guizetti, J., Aldrian-Herrada, G., Rogowski, K., Gerlich, D.W., and Janke, C. (2010). Tubulin polyglutamylation stimulates spastin-mediated microtubule severing. J Cell Biol *189*, 945-954.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature *227*, 680-685.

Lee, K.B., and Thomas, J.O. (2000). The effect of the acidic tail on the DNA-binding properties of the HMG1,2 class of proteins: insights from tail switching and tail removal. J Mol Biol *304*, 135-149.

Lee, M., Jin, Y., and Kim, D.H. (1999). 2-Benzyl-2-methylsuccinic acid as inhibitor for carboxypeptidase A. synthesis and evaluation. Bioorg Med Chem *7*, 1755-1760.

Legendre, C., Mooij, M.J., Adams, C., and O'Gara, F. (2011). Impaired expression of hypoxia-inducible factor-1alpha in cystic fibrosis airway epithelial cells - a role for HIF-1 in the pathophysiology of CF? J Cyst Fibros *10*, 286-290.

Lipscomb, W.N., and Strater, N. (1996). Recent Advances in Zinc Enzymology. Chem Rev *96*, 2375-2434.

Lowe, J., and Amos, L.A. (2009). Evolution of cytomotive filaments: the cytoskeleton from prokaryotes to eukaryotes. Int J Biochem Cell Biol *41*, 323-329.

Lyons, P.J., Callaway, M.B., and Fricker, L.D. (2008). Characterization of carboxypeptidase A6, an extracellular matrix peptidase. J Biol Chem *283*, 7054-7063.

Lyons, P.J., and Fricker, L.D. (2011). Carboxypeptidase O is a glycosylphosphatidylinositol-anchored intestinal peptidase with acidic amino acid specificity. J Biol Chem *286*, 39023-39032.

Ma, L., Jackson, K.D., Landry, R.M., Parsek, M.R., and Wozniak, D.J. (2006). Analysis of Pseudomonas aeruginosa conditional psl variants reveals roles for the psl polysaccharide in adhesion and maintaining biofilm structure postattachment. J Bacteriol *188*, 8213-8221.

Madhurantakam, C., Varadamsetty, G., Grutter, M.G., Pluckthun, A., and Mittl, P.R. (2012). Structure-based optimization of designed Armadillo-repeat proteins. Protein Sci *21*, 1015-1028.

Magiera, M.M., and Janke, C. (2014). Post-translational modifications of tubulin. Curr Biol *24*, R351-354.

Mahenthiralingam, E., Campbell, M.E., and Speert, D.P. (1994). Nonmotility and phagocytic resistance of Pseudomonas aeruginosa isolates from chronically colonized patients with cystic fibrosis. Infect Immun *62*, 596-605.

Maniatis, T., E. F. Fritsch, and J. Sambrook (1982). Molecular cloning, a laboratory manual (New York, Cold Spring Harbor Laboratory).

Mavrodi, D.V., Bonsall, R.F., Delaney, S.M., Soule, M.J., Phillips, G., and Thomashow, L.S. (2001). Functional analysis of genes for biosynthesis of pyocyanin and phenazine-1-carboxamide from Pseudomonas aeruginosa PAO1. J Bacteriol *183*, 6454-6465.

McGuffin, L.J., Bryson, K., and Jones, D.T. (2000). The PSIPRED protein structure prediction server. Bioinformatics *16*, 404-405.

Mesaros, N., Nordmann, P., Plesiat, P., Roussel-Delvallez, M., Van Eldere, J., Glupczynski, Y., Van Laethem, Y., Jacobs, F., Lebecque, P., Malfroot, A.*, et al.* (2007). Pseudomonas aeruginosa: resistance and therapeutic options at the turn of the new millennium. Clin Microbiol Infect *13*, 560-578.

Mimori-Kiyosue, Y., Vonderviszt, F., and Namba, K. (1997). Locations of terminal segments of flagellin in the filament structure and their roles in polymerization and polymorphism. J Mol Biol *270*, 222-237.

Mock, W.L., Liu, Y., and Stanford, D.J. (1996). Arazoformyl peptide surrogates as spectrophotometric kinetic assay substrates for carboxypeptidase A. Anal Biochem *239*, 218-222.

Mock, W.L., and Stanford, D.J. (2002). Anisylazoformylarginine: a superior assay substrate for carboxypeptidase B type enzymes. Bioorg Med Chem Lett *12*, 1193-1194.

Mock, W.L., and Zhang, J.Z. (1991). Mechanistically significant diastereoselection in the sulfoximine inhibition of carboxypeptidase A. J Biol Chem *266*, 6393-6400.

Morita, Y., Tomida, J., and Kawamura, Y. (2014). Responses of to antimicrobials. Front Microbiol *4*, 422.

Murray, P.R., E.J. Baron, F.C. Pfaller, M.A. Tenover and R.H. Yolken (1999). Manual of Clinical Microbiology, 7th edn (Washington, DC., American Society for Microbiology).

Nagpal, S., Patel, S., Asano, A.T., Johnson, A.T., Duvic, M., and Chandraratna, R.A. (1996). Tazarotene-induced gene 1 (TIG1), a novel retinoic acid receptor-responsive gene in skin. J Invest Dermatol *106*, 269-274.

Notredame, C., Higgins, D.G., and Heringa, J. (2000). T-Coffee: A novel method for fast and accurate multiple sequence alignment. J Mol Biol *302*, 205-217.

O'Toole, G.A. (2011). Microtiter dish biofilm formation assay. J Vis Exp.

O'Toole, G.A., and Kolter, R. (1998). Flagellar and twitching motility are necessary for Pseudomonas aeruginosa biofilm development. Mol Microbiol *30*, 295-304.

Ollis, D.L., Cheah, E., Cygler, M., Dijkstra, B., Frolow, F., Franken, S.M., Harel, M., Remington, S.J., Silman, I., Schrag, J*., et al.* (1992). The alpha/beta hydrolase fold. Protein Eng *5*, 197-211.

Otero, A., Rodriguez de la Vega, M., Tanco, S., Lorenzo, J., Aviles, F.X., and Reverter, D. (2012). The novel structure of a cytosolic M14 metallocarboxypeptidase (CCP) from Pseudomonas aeruginosa: a model for mammalian CCPs. FASEB J *26*, 3754-3764.

Overhage, J., Bains, M., Brazas, M.D., and Hancock, R.E. (2008). Swarming of Pseudomonas aeruginosa is a complex adaptation leading to increased production of virulence factors and antibiotic resistance. J Bacteriol *190*, 2671-2679.

Pallares, I., Bonet, R., Garcia-Castellanos, R., Ventura, S., Aviles, F.X., Vendrell, J., and Gomis-Ruth, F.X. (2005). Structure of human carboxypeptidase A4 with its endogenous protein inhibitor, latexin. Proc Natl Acad Sci U S A *102*, 3978-3983.

Papaioannou, E., Wahjudi, M., Nadal-Jimenez, P., Koch, G., Setroikromo, R., and Quax, W.J. (2009). Quorum-quenching acylase reduces the virulence of Pseudomonas aeruginosa in a Caenorhabditis elegans infection model. Antimicrob Agents Chemother *53*, 4891-4897.

Paranchych, W., Sastry, P.A., Frost, L.S., Carpenter, M., Armstrong, G.D., and Watts, T.H. (1979). Biochemical studies on pili isolated from Pseudomonas aeruginosa strain PAO. Can J Microbiol *25*, 1175-1181.

Park, A.J., Murphy, K., Krieger, J.R., Brewer, D., Taylor, P., Habash, M., and Khursigara, C.M. (2014). A Temporal Examination of the Planktonic and Biofilm Proteome of Whole Cell Pseudomonas aeruginosa PAO1 using Quantitative Mass Spectrometry. Mol Cell Proteomics.

Pejler, G., Knight, S.D., Henningsson, F., and Wernersson, S. (2009). Novel insights into the biological function of mast cell carboxypeptidase A. Trends Immunol *30*, 401-408.

Pelicic, V. (2008). Type IV pili: e pluribus unum? Mol Microbiol *68*, 827-837.

Pla, V., Paco, S., Ghezali, G., Ciria, V., Pozas, E., Ferrer, I., and Aguado, F. (2013). Secretory sorting receptors carboxypeptidase E and secretogranin III in amyloid beta-associated neural degeneration in Alzheimer's disease. Brain Pathol *23*, 274-284.

Plummer, T.H., Jr., and Kimmel, M.T. (1980). An improved spectrophotometric assay for human plasma carboxypeptidase N1. Anal Biochem *108*, 348-353.

Poole, K. (2011). Pseudomonas aeruginosa: resistance to the max. Front Microbiol *2*, 65.

Ramos, J. (2004). Pseudomonas (New York, Kluwer Academic/Plenum Publishers).

Rashid, M.H., and Kornberg, A. (2000). Inorganic polyphosphate is needed for swimming, swarming, and twitching motilities of Pseudomonas aeruginosa. Proc Natl Acad Sci U S A *97*, 4885-4890.

Rashid, M.H., Rao, N.N., and Kornberg, A. (2000a). Inorganic polyphosphate is required for motility of bacterial pathogens. J Bacteriol *182*, 225-227.

Rashid, M.H., Rumbaugh, K., Passador, L., Davies, D.G., Hamood, A.N., Iglewski, B.H., and Kornberg, A. (2000b). Polyphosphate kinase is essential for biofilm development, quorum sensing, and virulence of Pseudomonas aeruginosa. Proc Natl Acad Sci U S A *97*, 9636-9641.

Rawlings, N.D., and Barrett, A.J. (1993). Evolutionary families of peptidases. Biochem J *290 (Pt 1)*, 205-218.

Rawlings, N.D., Waller, M., Barrett, A.J., and Bateman, A. (2014). MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. Nucleic Acids Res *42*, D503-509.

Rees, D.C., and Lipscomb, W.N. (1982). Refined crystal structure of the potato inhibitor complex of carboxypeptidase A at 2.5 A resolution. J Mol Biol *160*, 475-498.

Reverter, D., Fernandez-Catalan, C., Baumgartner, R., Pfander, R., Huber, R., Bode, W., Vendrell, J., Holak, T.A., and Aviles, F.X. (2000). Structure of a novel leech carboxypeptidase inhibitor determined free in solution and in complex with human carboxypeptidase A2. Nat Struct Biol *7*, 322-328.

Reznik, S.E., and Fricker, L.D. (2001). Carboxypeptidases from A to z: implications in embryonic development and Wnt binding. Cell Mol Life Sci *58*, 1790-1804.

Rimsa, V., Eadsforth, T.C., Joosten, R.P., and Hunter, W.N. (2014). High-resolution structure of the M14-type cytosolic carboxypeptidase from Burkholderia cenocepacia refined exploiting PDB_REDO strategies. Acta Crystallographica Section D *70*, 279-289.

Rodriguez de la Vega, M., Sevilla, R.G., Hermoso, A., Lorenzo, J., Tanco, S., Diez, A., Fricker, L.D., Bautista, J.M., and Aviles, F.X. (2007). Nna1-like proteins are active metallocarboxypeptidases of a new and diverse M14 subfamily. FASEB J *21*, 851-865.

Rodriguez de la Vega Otazo, M., Lorenzo, J., Tort, O., Aviles, F.X., and Bautista, J.M. (2012). Functional segregation and emerging role of cilia-related cytosolic carboxypeptidases (CCPs). FASEB J *27*, 424-431.

Rogowski, K., van Dijk, J., Magiera, M.M., Bosc, C., Deloulme, J.C., Bosson, A., Peris, L., Gold, N.D., Lacroix, B., Bosch Grau, M.*, et al.* (2010). A family of protein-deglutamylating enzymes associated with neurodegeneration. Cell *143*, 564-578.

Rotanova, T.V., Botos, I., Melnikov, E.E., Rasulova, F., Gustchina, A., Maurizi, M.R., and Wlodawer, A. (2006). Slicing a protease: structural features of the ATP-dependent Lon proteases gleaned from investigations of isolated domains. Protein Sci *15*, 1815-1828.

Roy, A., Kucukural, A., and Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. Nat Protoc *5*, 725-738.

Sahab, Z.J., Hall, M.D., Me Sung, Y., Dakshanamurthy, S., Ji, Y., Kumar, D., and Byers, S.W. (2011). Tumor suppressor RARRES1 interacts with cytoplasmic carboxypeptidase AGBL2 to regulate the alpha-tubulin tyrosination cycle. Cancer Res *71*, 1219-1228.

Samuel, A.D., Pitta, T.P., Ryu, W.S., Danese, P.N., Leung, E.C., and Berg, H.C. (1999). Flagellar determinants of bacterial sensitivity to chi-phage. Proc Natl Acad Sci U S A *96*, 9863-9866.

Sanglas, L., Arolas, J.L., Valnickova, Z., Aviles, F.X., Enghild, J.J., and Gomis-Ruth, F.X. (2010). Insights into the molecular inactivation mechanism of human activated thrombin-activatable fibrinolysis inhibitor. J Thromb Haemost *8*, 1056-1065.

Sanglas, L., Aviles, F.X., Huber, R., Gomis-Ruth, F.X., and Arolas, J.L. (2009). Mammalian metallopeptidase inhibition at the defense barrier of Ascaris parasite. Proc Natl Acad Sci U S A *106*, 1743-1747.

Sauer, K., Cullen, M.C., Rickard, A.H., Zeef, L.A., Davies, D.G., and Gilbert, P. (2004). Characterization of nutrient-induced dispersion in Pseudomonas aeruginosa PAO1 biofilm. J Bacteriol *186*, 7312-7326.

Sauer, R.T., and Baker, T.A. (2011). AAA+ proteases: ATP-fueled machines of protein destruction. Annu Rev Biochem *80*, 587-612.

Schechter, I., and Berger, A. (2012). On the size of the active site in proteases. I. Papain. 1967. Biochem Biophys Res Commun *425*, 497-502.

Schmidt, T.G., and Skerra, A. (2007). The Strep-tag system for one-step purification and high-affinity detection or capturing of proteins. Nat Protoc *2*, 1528-1535.

Schreiber, K., Boes, N., Eschbach, M., Jaensch, L., Wehland, J., Bjarnsholt, T., Givskov, M., Hentzer, M., and Schobert, M. (2006). Anaerobic survival of Pseudomonas aeruginosa by pyruvate fermentation requires an Usp-type stress protein. J Bacteriol *188*, 659-668.

Schrodinger, LLC (2010). The PyMOL Molecular Graphics System, Version 1.3r1.

Schuster, M., and Greenberg, E.P. (2006). A network of networks: quorum-sensing gene regulation in Pseudomonas aeruginosa. Int J Med Microbiol *296*, 73-81.

Semmler, A.B., Whitchurch, C.B., Leech, A.J., and Mattick, J.S. (2000). Identification of a novel gene, fimV, involved in twitching motility in Pseudomonas aeruginosa. Microbiology *146 (Pt 6)*, 1321-1332.

Semmler, A.B., Whitchurch, C.B., and Mattick, J.S. (1999). A re-examination of twitching motility in Pseudomonas aeruginosa. Microbiology *145 (Pt 10)*, 2863-2873.

Seo, S., Baye, L.M., Schulz, N.P., Beck, J.S., Zhang, Q., Slusarski, D.C., and Sheffield, V.C. (2010). BBS6, BBS10, and BBS12 form a complex with CCT/TRiC family chaperonins and mediate BBSome assembly. Proc Natl Acad Sci U S A *107*, 1488-1493.

Shang, Y., Li, B., and Gorovsky, M.A. (2002). Tetrahymena thermophila contains a conventional gamma-tubulin that is differentially required for the maintenance of different microtubule-organizing centers. J Cell Biol *158*, 1195-1206.

Shaya, D., Zhao, W., Garron, M.L., Xiao, Z., Cui, Q., Zhang, Z., Sulea, T., Linhardt, R.J., and Cygler, M. (2010). Catalytic mechanism of heparinase II investigated by site-directed mutagenesis and the crystal structure with its substrate. J Biol Chem *285*, 20051-20061.

Shrout, J.D., Chopp, D.L., Just, C.L., Hentzer, M., Givskov, M., and Parsek, M.R. (2006). The impact of quorum sensing and swarming motility on Pseudomonas aeruginosa biofilm formation is nutritionally conditional. Mol Microbiol *62*, 1264-1277.

Sidyelyeva, G., Wegener, C., Schoenfeld, B.P., Bell, A.J., Baker, N.E., McBride, S.M., and Fricker, L.D. (2010). Individual carboxypeptidase D domains have both redundant and unique functions in Drosophila development and behavior. Cell Mol Life Sci *67*, 2991-3004.

Sirajuddin, M., Rice, L.M., and Vale, R.D. (2014). Regulation of microtubule motors by tubulin isotypes and post-translational modifications. Nat Cell Biol *16*, 335-344.

Skidgel, R.A., and Erdos, E.G. (2007). Structure and function of human plasma carboxypeptidase N, the anaphylatoxin inactivator. Int Immunopharmacol *7*, 1888-1899.

Song, W.S., and Yoon, S.I. (2014). Crystal structure of FliC flagellin from Pseudomonas aeruginosa and its implication in TLR5 binding and formation of the flagellar filament. Biochem Biophys Res Commun *444*, 109-115.

Sourjik, V., and Wingreen, N.S. (2012). Responding to chemical gradients: bacterial chemotaxis. Curr Opin Cell Biol *24*, 262-268.

Strom, M.S., Nunn, D.N., and Lory, S. (1993). A single bifunctional enzyme, PilD, catalyzes cleavage and N-methylation of proteins belonging to the type IV pilin family. Proc Natl Acad Sci U S A *90*, 2404-2408.

Tammam, S., Sampaleanu, L.M., Koo, J., Sundaram, P., Ayers, M., Chong, P.A., Forman-Kay, J.D., Burrows, L.L., and Howell, P.L. (2011). Characterization of the PilN, PilO and PilP type IVa pilus subcomplex. Mol Microbiol *82*, 1496-1514.

Tan, M.W., Mahajan-Miklos, S., and Ausubel, F.M. (1999). Killing of Caenorhabditis elegans by Pseudomonas aeruginosa used to model mammalian bacterial pathogenesis. Proc Natl Acad Sci U S A *96*, 715-720.

Tanco, S., Tort, O., Demol, H., Aviles, F.X., Gevaert, K., Van Damme, P., and Lorenzo, J. (2014, submitted). C‑terminomics screen for natural substrates of cytosolic carboxypeptidase 1 reveals processing of acidic protein C‑termini. Mollecular and Cellular Proteomics.

Tanco, S., Zhang, X., Morano, C., Aviles, F.X., Lorenzo, J., and Fricker, L.D. (2010). Characterization of the substrate specificity of human carboxypeptidase A4 and implications for a role in extracellular peptide processing. J Biol Chem *285*, 18385-18396.

Taylor, T.B., and Buckling, A. (2010). Competition and dispersal in Pseudomonas aeruginosa. Am Nat *176*, 83-89.

Tort, O., Tanco, S., Rocha, C., Bièche, I., Seixas, C., Bosc, C., Andrieux, A., Moutin, M.J., Avilés, F.X., Lorenzo, J.*, et al.* (2014, submitted). The cytosolic carboxypeptidase family is dedicated to posttranslational removal of acidic amino acids. Molecular Biology of the Cell.

Toutain, C.M., Zegans, M.E., and O'Toole, G.A. (2005). Evidence for two flagellar stators and their role in the motility of Pseudomonas aeruginosa. J Bacteriol *187*, 771-777.

Utari, P.D., and Quax, W.J. (2013). Caenorhabditis elegans reveals novel Pseudomonas aeruginosa virulence mechanism. Trends Microbiol *21*, 315-316.

Vallet, I., Olson, J.W., Lory, S., Lazdunski, A., and Filloux, A. (2001). The chaperone/usher pathways of Pseudomonas aeruginosa: identification of fimbrial gene clusters (cup) and their involvement in biofilm formation. Proc Natl Acad Sci U S A *98*, 6911-6916.

Vater, S.M., Weisse, S., Maleschlijski, S., Lotz, C., Koschitzki, F., Schwartz, T., Obst, U., and Rosenhahn, A. (2014). Swimming behavior of Pseudomonas aeruginosa studied by holographic 3D tracking. PLoS One *9*, e87765.

Vendrell, J., Querol, E., and Aviles, F.X. (2000). Metallocarboxypeptidases and their protein inhibitors. Structure, function and biomedical properties. Biochim Biophys Acta *1477*, 284-298.

Vendrell, J.a.A., F. X. (1999). Carboxypeptidases. In Proteases: New Perspectives, T. V., ed. (Switzerland, Birkyhäuser Verlag Basel), pp. 1-12.

Villanueva, J., Canals, F., Prat, S., Ludevid, D., Querol, E., and Aviles, F.X. (1998). Characterization of the wound-induced metallocarboxypeptidase inhibitor from potato. cDNA sequence, induction of gene expression, subcellular immunolocalization and potential roles of the C-terminal propeptide. FEBS Lett *440*, 175-182.

Wang, L., Shao, Y.Y., and Ballock, R.T. (2009). Carboxypeptidase Z (CPZ) links thyroid hormone and Wnt signaling pathways in growth plate chondrocytes. J Bone Miner Res *24*, 265-273.

Wang, T., Parris, J., Li, L., and Morgan, J.I. (2006). The carboxypeptidase-like substrate-binding site in Nna1 is essential for the rescue of the Purkinje cell degeneration (pcd) phenotype. Mol Cell Neurosci *33*, 200-213.

Wang, Y.N., Wang, H., Yamaguchi, H., Lee, H.J., Lee, H.H., and Hung, M.C. (2010). COPI-mediated retrograde trafficking from the Golgi to the ER regulates EGFR nuclear transport. Biochem Biophys Res Commun *399*, 498-504.

Wehbi, H., Portillo, E., Harvey, H., Shimkoff, A.E., Scheurwater, E.M., Howell, P.L., and Burrows, L.L. (2011). The peptidoglycan-binding protein FimV promotes assembly of the Pseudomonas aeruginosa type IV pilus secretin. J Bacteriol *193*, 540-550.

Wehenkel, A., and Janke, C. (2014). Towards elucidating the tubulin code. Nat Cell Biol *16*, 303-305.

Wei, S., Feng, Y., Kalinina, E., and Fricker, L.D. (2003). Neuropeptide-processing carboxypeptidases. Life Sci *73*, 655-662.

Wei, Y., Wang, X., Liu, J., Nememan, I., Singh, A.H., Weiss, H., and Levin, B.R. (2011). The population dynamics of bacteria in physically structured habitats and the adaptive virtue of random motility. Proc Natl Acad Sci U S A *108*, 4047-4052.

Whitchurch, C.B., Tolker-Nielsen, T., Ragas, P.C., and Mattick, J.S. (2002). Extracellular DNA required for bacterial biofilm formation. Science *295*, 1487.

Williamson, K.S., Richards, L.A., Perez-Osorio, A.C., Pitts, B., McInnerney, K., Stewart, P.S., and Franklin, M.J. (2012). Heterogeneity in Pseudomonas aeruginosa biofilms includes expression of ribosome hibernation factors in the antibiotic-tolerant subpopulation and hypoxia-induced stress response in the metabolically active population. J Bacteriol *194*, 2062-2073.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G., McCoy, A.*, et al.* (2011). Overview of the CCP4 suite and current developments. Acta Crystallogr D Biol Crystallogr *67*, 235-242.

Winstanley, C., Langille, M.G., Fothergill, J.L., Kukavica-Ibrulj, I., Paradis-Bleau, C., Sanschagrin, F., Thomson, N.R., Winsor, G.L., Quail, M.A., Lennard, N.*, et al.* (2009). Newly introduced genomic prophage islands are critical determinants of in vivo competitiveness in the Liverpool Epidemic Strain of Pseudomonas aeruginosa. Genome Res *19*, 12-23.

Wu, H.Y., Wang, T., Li, L., Correia, K., and Morgan, J.I. (2012). A structural and functional analysis of Nna1 in Purkinje cell degeneration (pcd) mice. FASEB J *26*, 4468-4480.

Wu, S., Zhang, C., Xu, D., and Guo, H. (2010). Catalysis of carboxypeptidase A: promoted-water versus nucleophilic pathways. J Phys Chem B *114*, 9259-9267.

Xavier, J.B., Kim, W., and Foster, K.R. (2011). A molecular mechanism that stabilizes cooperative secretions in Pseudomonas aeruginosa. Mol Microbiol *79*, 166-179.

Zaborin, A., Romanowski, K., Gerdes, S., Holbrook, C., Lepine, F., Long, J., Poroyko, V., Diggle, S.P., Wilke, A., Righetti, K.*, et al.* (2009). Red death in Caenorhabditis elegans caused by Pseudomonas aeruginosa PAO1. Proc Natl Acad Sci U S A *106*, 6327-6332.

Zolfaghar, I., Evans, D.J., and Fleiszig, S.M. (2003). Twitching motility contributes to the role of pili in corneal infection caused by Pseudomonas aeruginosa. Infect Immun *71*, 5389-5393.