Investigating the activation and regulation of the proteasome, an essential proteolytic complex

Patrick Masson

Institutionen för Molekylärbiologi & Funktionsgenomik
Stockholm Universitet

Stockholm 2004
Abstract

The proteasome is a major non-lysosomal proteolytic complex present in eukaryotic cells and has a central role in regulating many protein levels. The complex has been shown to participate in various intracellular pathways including cell cycle regulation or quality control of newly synthesized proteins and many other key pathways. This amazing range of substrates would not be possible without the help of regulators that are able to bind to the 20S proteasome and modulate its activity. Among those, the PA700 or 19S regulator and the PA28 family are the best characterized in higher eukaryotes. The 19S regulatory particle is involved in the recognition of ubiquitinated proteins, targeted for degradation by the proteasome. The PA28 (also termed 11S REG) family is composed of two members the PA28αβ and PA28γ. The function of PA28αβ is related to the adaptive immune response with a proposed contribution in MHC class I peptide presentation whereas the biological role PA28γ remains unknown. The main objectives of the laboratory, and subsequently of this thesis are to use Drosophila melanogaster model system and its advantages to better understand the precise contribution of these different activators in the regulation of the proteasome. Using genomic resources, a unique Drosophila PA28 member was identified, characterized and was shown to be a proteasome regulator with all the properties of PA28γ. Through site-directed mutagenesis we identified a functional nuclear localization signal in the homolog-specific insert region. Study of the promoter region revealed that transcription of Drosophila PA28γ (dPA28γ) gene is under control of DREF, a transcription factor involved in the regulation of genes related to DNA synthesis and cell proliferation. To confirm that dPA28γ has a role in cell cycle progression, the effect of removing dPA28γ from S2 cells was tested using RNA interference. Drosophila cells depleted of dPA28γ showed partial arrest in G1/S cell cycle transition confirming a conserved function between Drosophila and mammalian forms of PA28γ. Finally, characterization of the Dictyostelium regulator, an evolutionarily distant member of the PA28γ, was carried out using fluorogenic degradation assays. We are currently knocking-out the gene in order to determine the biological function of the activator. A second part of my work consisted in the generation of a Drosophila assay used to identify in vivo substrates of the 19S regulator, an assay system that has been originally engineered by Dantuma and coworkers in human cell lines. This was achieved by cloning of GFP behind a series of modified ubiquitins that create substrates degraded through different pathways involving the proteasome pathways. The last project of my thesis concerns the characterization of the mechanism for upregulation of proteasomal gene mRNA after MG132 (proteasome inhibitor) treatment. So far, we found that the 5’-UTR of the genes is responsible for this induction. We are now looking for the precise motif involved in this regulation.

© Patrick Masson
ISBN 91-7265-837-1
# TABLE OF CONTENT

## LIST OF PUBLICATIONS

## ABBREVIATIONS

## INTRODUCTION

### PROTEASES: HISTORICAL ASPECTS

### THE 20S PROTEASOME

### STRUCTURAL PROPERTIES

### PROTEASOME ACTIVITY, FUNCTION AND LOCALIZATION

### THE 26S PROTEASOME AND THE UBIQUITIN PATHWAY

### STRUCTURE OF THE 19S REGULATORY PARTICLE

### THE UBIQUITIN-PROTEASOME PATHWAY

### THE PA28 ACTIVATOR OR 11S REG COMPLEX

### STRUCTURE AND ACTIVATION OF THE PROTEASOME

### PA28, IMMUNOPROTEASOME AND IMMUNITY

### THE PA28GAMMA

### STRUCTURE AND ACTIVATION OF THE PROTEASOME

### PA28γ FUNCTION

### PA28γ INTERACTING PROTEINS

### ADVANTAGES OF DROSOPHILA MODEL IN OUR INVESTIGATION:

### DROSOPHILA MELANOGASTER AS MODEL SYSTEM

### DROSOPHILA EMBRYOGENESIS

### RNAI TECHNIQUES

### DROSOPHILA MODEL FOR TRANSCRIPTIONAL REGULATION

## AIM OF THE WORK

## RESULTS AND DISCUSSION

### STUDY OF THE INVERTEBRATE PA28 HOMOLOG

### THE DROSOPHILA PA28 : A HOMOLOG OF HUMAN PA28GAMMA (PAPER I)

### THE DICTYOSTELIUM PA28 HOMOLOG (PAPER IV)

### INVESTIGATING THE FUNCTION OF PA28GAMMA

### AN IN VIVO ASSAY TO STUDY THE DROSOPHILA 26S PROTEASOME (PAPER III)

### STUDY OF THE PROTEASOMAL REGULATION AFTER MG132 TREATMENT (PAPER V)

## CONCLUSIONS AND PERSPECTIVES

## ACKNOWLEDGMENTS

## REFERENCES
LIST OF PUBLICATIONS

This thesis is based on the following publications:


# ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATP</td>
<td>Adenosine-tri-phosphate</td>
</tr>
<tr>
<td>CSN2</td>
<td>Signalosome subunit 2</td>
</tr>
<tr>
<td>COP9</td>
<td>constitutive photomorphogenesis signalosome</td>
</tr>
<tr>
<td>DPE</td>
<td>Downstream promoter element</td>
</tr>
<tr>
<td>DRE</td>
<td><em>Drosophila</em> replication element</td>
</tr>
<tr>
<td>DREF</td>
<td><em>Drosophila</em> replication element factor</td>
</tr>
<tr>
<td>ER</td>
<td>Endoplasmic reticulum</td>
</tr>
<tr>
<td>ERAD</td>
<td>Endoplasmic reticulum associated degradation</td>
</tr>
<tr>
<td>HCV</td>
<td>Hepatitis c virus</td>
</tr>
<tr>
<td>HECT</td>
<td>homologous to E6-AP C terminus</td>
</tr>
<tr>
<td>MHC</td>
<td>Major histocompatibility complex</td>
</tr>
<tr>
<td>GFP</td>
<td>Green fluorescence protein</td>
</tr>
<tr>
<td>RNAi</td>
<td>RNA interference</td>
</tr>
<tr>
<td>MG 132</td>
<td>Cbz-Leu-Leu-leucinal</td>
</tr>
<tr>
<td>MCA</td>
<td>Methylcouramin</td>
</tr>
<tr>
<td>LLVY</td>
<td>Leu-Leu-Val-Tyr</td>
</tr>
<tr>
<td>LPS</td>
<td>Lipopolysaccharide</td>
</tr>
<tr>
<td>LRR</td>
<td>Leu-Arg-Arg</td>
</tr>
<tr>
<td>ODC</td>
<td>Ornithine decarboxylase</td>
</tr>
<tr>
<td>PA28</td>
<td>Proteasome activator</td>
</tr>
<tr>
<td>PCNA</td>
<td>Proliferating cell nuclear antigen</td>
</tr>
<tr>
<td>11S REG</td>
<td>11S Regulator</td>
</tr>
<tr>
<td>RNP</td>
<td>Ribonucleoprotein</td>
</tr>
<tr>
<td>Rpn</td>
<td>Regulatory particle non-ATPase</td>
</tr>
<tr>
<td>Rpt</td>
<td>Regulatory particle triple-A protein</td>
</tr>
<tr>
<td>SLE</td>
<td>Systemic lupus erythematosus</td>
</tr>
<tr>
<td>SNO</td>
<td>Strawberry notch</td>
</tr>
<tr>
<td>TAP</td>
<td>Transporter associated with antigen processing</td>
</tr>
<tr>
<td>UTR</td>
<td>Untranslated region</td>
</tr>
<tr>
<td>VCP</td>
<td>vasolin-containing protein</td>
</tr>
</tbody>
</table>
INTRODUCTION

Proteases: historical aspects

Proteases or proteolytic enzymes constitute the largest protein family and their function is to catalyze the cleavage of peptide bonds. This reaction is one of the most frequent protein modification occurring in nature. The initial interest of scientists towards understanding digestion was the foundation for our current knowledge on enzymology in general and proteases in particular. During the early 18th century, Reaumur discovered that gastric juices were able to degrade meat chemically and not mechanically which was the current idea held at that time (Reaumur, 1761). Continuing the work of Reaumur, Spallanzani found that the isolated gastric juice was more efficient at body temperature rather than room temperature (Spallanzani, 1780). In 1836, while also investigating the digestive process, the German physiologist Theodor Schwann isolated a substance responsible for digestion in the stomach, and named it pepsin, which is the first enzyme prepared from animal tissue. In 1877, Kühne proposed the general term enzyme (Kühne, 1877). Whereas catalytic properties and specificity of enzymes were discovered shortly after, it took until mid-twenties for the scientific community to accept that enzymes were proteins. In early 1930s Northrop and his collaborators crystallized pepsin, trypsin and chymotrypsin and demonstrated the purity of their obtained crystals (Northrop, 1946). Rapidly improving protein purification techniques as well as other major discoveries during the last century including the DNA double-helix structure in 1953 (Watson and Crick, 1953) and the genetic code (Crick, 1968), allowed researchers to characterize an impressive amount of new proteases. Today several thousand proteolytic enzymes have been described and characterized.

The importance of proteases in living organisms as well as the growing number of publications concerning them directed scientists to classify them into families. In Drosophila for instance 564 out of the 13600 known genes are estimated to be proteases, which represents 4,15% of the total amount (Rawlings et al., 2002). In order to achieve a classification, the protease catalytic residue or the structure of the active site involved in the cleavage has been used to establish the main groups (Barrett and Rawlings, 1991) and so far, six main groups have been categorized according to the specific active site. Within a group, families have been created based upon common evolutionary origins and common
structures. The six groups and examples of proteases within each group are described below:

- aspartic proteases including pepsin A and pepsin B
- metalloproteases including aminopeptidase A
- cysteine proteases including caspase-1 to -14
- serine proteases including trypsin and chymotrypsin
- threonine proteases including the proteasome
- unknown proteases

Recently a detailed classification has been realized and a comprehensive database termed MEROPS has been created (Rawlings and Barrett, 1999). A second nomenclature termed the EC nomenclature of proteases lists a total of 13 subgroups.

Cellular pathways are tightly regulated and significantly depend upon the rate of synthesis and degradation of the involved proteins. Proteases are thus required to precisely regulate the crucial pathways needed by the cell in order to survive, both in a qualitative and quantitative way.

In 1953, Simpson established the surprising requirement for energy (ATP) in protein breakdown. Several lines of research were later undertaken to determine the different functions of ATP served in proteolytic processes. This led in 1980 to the identification of the energy-dependent formation of a covalent linkage between a protein substrate and one or more ATP-dependent proteolysis factor 1 (APF-1) molecules, termed today as ubiquitin (Hershko et al., 1980). Three years later, Tanaka and coworkers identified an additional step requiring ATP for the ubiquitin dependent protein degradation to function properly (Tanaka et al., 1983). This second step led to the identification of a large protease termed the proteasome whose function is to degrade the ubiquitinated substrate in an ATP-dependent manner (Hough et al., 1986).

Now, the functions of the proteasome are well documented and this complex has been shown to degrade a wide range of cellular proteins. Cell cycle progression for example requires the degradation of many short-lived proteins in order to proceed; and the vast majority of them are targeted to the proteasome. The impressive range of participations of the proteasome would not be possible without the help of different activators that greatly increase its performances within the cell. Studying the precise contribution of these different activators will provide much more understanding on the overall proteasome activity and functions.
The 20S proteasome

The proteasome is a large multicatalytic protease involved in many intracellular pathways including degradation of aberrant or damaged proteins and inactivation of proteins that play key roles in regulatory processes. The proteasome is composed of four stacked rings, each one containing seven subunits, which form a barrel-like structure with a $\alpha_1$-$\beta_1$-$\beta_1$-$\beta_1$-$\alpha_1$-$\beta_1$-$\beta_1$ stoichiometry (Puhler et al., 1992). The catalytic sites are located inside a central cavity, which restricts access to specific substrates targeted for degradation. Other proteases such as ClpAP or HslU share this quaternary structure, which is very convenient to avoid undesired and uncontrolled peptide hydrolysis throughout the cell (Kessel et al., 1995), (Sousa et al., 2000). After its discovery in human erythrocytes (Harris, 1968), the 20 S proteasome has been characterized in a wide variety of organisms, including all the eukaryotes and part of prokaryotes. Sequenced genome analysis indicates however, that most eubacteria such as Escherichia coli do not possess proteasomes but instead share a proteasome-like system termed HslUV. The HslUV complex has a diameter of 110 Å and a length of 75Å whereas the diameter and length of the archaeal Thermoplasma acidophilum proteasome are 100Å and 175Å respectively. Structurally, HslUV consists of only two homohexameric rings, explaining the shorter length (fig. 1).

Fig. 1. Crystal structure of E. coli HslUV, and proteasomes from Thermoplasma acidophilum and yeast respectively. Adapted from (Groll and Clausen, 2003).
Structural properties

Typically, two different genes encode the archaeal proteasome subunits alpha and beta (Zwickl et al., 1992). The active sites are located at the N-terminal threonine of the beta subunits and face the central chamber. Two narrow openings of approximately 13 Å reside at both ends of the cylinder (Lowe et al., 1995). The eukaryotic proteasome is made of 14 different subunits, 7 alpha and 7 beta, encoded by 14 distinct genes that confer a more complex organization (Heinemeyer et al., 1994). The catalytic subunits are also located on the beta subunits but only three of them are active (β1, β2 and β5 respectively Y, Z, X in the human nomenclature). Despite this difference, they share their catalytic mechanism with an N-terminal nucleophilic threonine, which places the proteasome in the family of Ntn (N terminal nucleophile) hydrolases (Seemuller et al., 1995), (Heinemeyer et al., 1997).

Upon assembly of the eukaryotic complex, the N-terminal extensions located on the active subunits, also termed propeptides, are autocatalytically removed to allow generation of the N-terminal threonine inside the cylinder (Schmidt and Kloetzel, 1997). In addition to this function, Arendt and coworkers have shown that these propeptides also play a role in protecting the N-terminal threonine against N-alpha acetylation (Arendt and Hochstrasser, 1999). In the inactive beta subunits, this propeptide is never cleaved, leading to a structure where the threonine is masked by the prosequence.

While the Thermoplasma proteasome only possesses one activity as a result of the identical beta subunits, the eukaryotic complex has three distinct proteolytic activities for each active beta subunits: the β1 subunit has a peptidylglutamyl-hydrolyzing activity also termed postacidic or caspase-like activity, meaning that the peptide bond hydrolysis occurs directly after acidic residues. The β2 subunit carries a trypsin-like activity cleaving after basic residues. Finally, a chymotrypsin-like activity has been assigned to the β5 subunit. This increased number of activities is partially due to the differences in the structural architecture of the proteolytic chambers. Studies using calpain I inhibitors together with the yeast proteasome crystal structure identified essential residues for the different activities (Groll et al., 1997). The S1 pocket (specificity) is mainly shaped by five conserved residues with the residue 45 forming the bottom of the pocket and largely contributing to its character. The residue 45 is an arginine in the β1 subunit, a positively charged amino acid well suited for the peptidylglutamyl-hydrolazing activity. Concerning β2, the residue is a glycine forming a spacious S1 pocket, which is convenient for large residues, and glutamine 53 at the bottom of the pocket also contributes to the trypsin-like activity. Finally, β5 possess a methionine at
position 45, an apolar residue ideal for chymotrypsin-like activity. In addition, inhibitor studies using two analogue compounds MB1 and MB2 (fig. 2) differing only in their P3 (position 3) and P4 residues showed that residues other than P1 also influence the degradation (Groll et al., 2002). This result highlights the importance of the whole inner chamber imposing a physical constraint to the substrate rather than the sole S1 pocket.

**Fig 2.** Chemical structures of the two Peptide Vinyl Sulfone Proteasome inhibitors MB1 Ac-PRLN-VS (top) and MB2 Ac-YLLN-VS (bottom) differing on their P3 and P4. Adapted from Groll et al. 2002.

Finally, mutational studies revealed that neighboring subunits interfere with the functions of the catalytic subunits (Groll et al., 1999).

**Proteasome activity, function and localization**

The cleavage fragments that are generated *in vitro* by the mammalian 20S proteasome vary in length between 3 and 22 amino acids with an average of 7-8 amino acids (Kisselev et al., 1999). In other terms, the products leave the catalytic chamber before complete degradation to individual amino acids. Additional proteases have been suggested to function downstream of the 20S proteasome to further process fragments and produce single amino acids, such as the tricorn protease (Tamura et al., 1996). By itself, the 20S proteasome only poorly degrades oligopeptides and unfolded proteins, and the purified complex has negligible activity towards native proteins (Tanaka et al., 1986). The 20S typically requires binding of additional activators to be able to degrade folded proteins or long polypeptides. Among these, the 19S regulatory complex binds to the 20S core and induces protein degradation through
the ubiquitin pathway that will be discussed later. Other types of activators exist in the cell with one member of the PA28 family being the focus of this thesis. The proportions of these different activating complexes is believed to depend on the cell type, the cell cycle phase and metabolic conditions (Brooks et al., 2000). The free 20S core particle however has been reported to be in large excess in the cell (Brooks et al., 2000), possibly suggesting either a specific role in proteolysis in absence of any bound regulatory particles, or the excess of 20S prevents unwanted competition between the different regulators in gaining access to the 20S activity.

In HeLa cells, the 20S proteasome is very abundant and constitutes approximately 0.6% of the total cellular protein (Hendil, 1988). A number of studies indicate that the proteasome is widespread both in the cytoplasm and the nucleus. The precise distribution is very difficult to determine, and probably depends on cell type, growth conditions, and methods used for detection (Wojcik and DeMartino, 2003). Electron microscopy studies have determined that 14% of cytoplasmic proteasomes is associated to the outer surface of the endoplasmic reticulum in rat hepatocytes (Rivett et al., 1992). The importance of this fraction is probably due, at least in part, to the involvement of the proteasome in the ERAD process (Endoplasmic Reticulum-associated degradation) consisting of the degradation of abnormal or misfolded proteins from the ER. Subfractionation of the microsomes shows that the proteasomes are associated with the smooth endoplasmic reticulum and with the cis-golgi but are practically absent from the rough ER (Palmer et al., 1996). Another explanation for the presence of the proteasome in the ER comes from its involvement in antigen presentation. Indeed, it can associate with the ER membrane where it could in theory come in contact with the TAP transporter (transporter associated with antigen processing), an essential complex required for the MHC class I peptide presentation (Rechsteiner et al., 2000).

The proteasome is also present in the nucleus. The crystal structure of the mammalian proteasome identified four putative nuclear localization signals on the alpha subunits 1, 2, 3, and 4 with the consensus sequence K-K(R)-X-K(R) that could in principle interact with the importin-α for nuclear import (Unno et al., 2002). In the nucleus, the proteasome is localized in various regions including the euchromatin and the periphery of the heterochromatin and nucleoli. However, this overall distribution is very dynamic and is very cell specific. Indeed, during cell cycle progression, there is an increase of nuclear proteasome from G1 to G2 phases. During prophase, 20S proteasomes accumulate around the condensing chromosomes whereas in late anaphase they colocalize with α-tubulin of the spindle fibers. In telophase and early interphase of the
daughter cells, an intensive nucleus staining was still observed (Amsterdam et al., 1993). Alternatively, during programmed cell death, proteasomes from rat ovarian granulosa cells are removed from the nucleus and accumulate within the apoptotic blebs at the periphery of the cell (Pitzer et al., 1996).

The great dynamic distribution of proteasomes together with their ability to form various complexes with different regulators greatly increases the complexity of this critical complex and may also increase the number of essential proteolytic roles that the 20S proteasome can fulfill within the cell. In the next part I will focus on the two most studied regulatory particle families, the 19S and the PA28 family and I will describe the reasons why they render the 20S core particle a much more efficient and specific destruction machinery.
The 26S proteasome and the ubiquitin pathway

The 19S regulatory particle is composed of at least 18 subunits and can attach to one or both ends of the 20S core particle in the presence of ATP. Binding on both ends leads to the structure called the 26S proteasome. One problem in the field is the multiple names given to each subunit in different organisms. For instance, one subunit is termed S5a in human, Rpn10 or Mcb-1 in Yeast, and p54 in Drosophila.

Structure of the 19S regulatory particle

Unlike the 20S core, the 19S complex structure is not resolved. Indeed, only a low-resolution structure of the complexes was provided by electro-microscopy whereas precise positioning of individual subunits remains largely unknown (Walz et al., 1998). Purified yeast proteasome mutants (Glickman et al., 1998) instead provided the data concerning the structure of the complex. The 19S complex is composed of a base immediately adjacent to the 20S core and a distal lid (fig. 3).

![Diagram of the 19S proteasome activator](image)

**Fig. 3.** Subunit distribution of the 19S proteasome activator. The 19S represented as 17 subunits subdivided into the base and the lid (derived from Ferrell et al. 2000).

The base of the 19S is composed of 8 subunits including 6 ATPases plus the two largest subunits S1 and S2. The six ATPases (termed S4, S6, S6’, S7, S8, and S10b in human and Rpt1 to Rpt6 in yeast) belong to the family of AAA-ATPases (ATPases associated with a variety of cellular
activities) with a shared conserved region of approximately 400 amino acid residues. Despite the unknown structure, several lines of evidence suggest that these ATPases subunits form a heterohexameric ring. Several complexes such as VCP, CDC48 from the AAA-ATPase family are known to also form hexameric rings and this may be a conserved feature. Sequence alignments show a similar architecture of the six genes with a central nucleotide-binding domain about 60% identical between members, and a 40% identity concerning the last 150 amino acids.

Several regulatory functions have been attributed to the base. Among these, the base is thought to perform two essential steps for protein degradation, unfolding and translocation of the protein substrates to the 20S proteasome. The unfolding occurs at the surface of the ATPase ring and the translocation follows after the ATP-dependent unfolding (Navon and Goldberg, 2001). The base is also able to open the gate of the 20S core in order to allow access to the catalytic chamber by rearrangements of the alpha subunits. Furthermore, some specific functions have been identified for certain subunits of the base. For instance, the subunit S6′ has been proposed to associate with polyubiquitin chains (Lam et al., 2002), which is the signal that is covalently bound to the substrate and targets it to the 26S proteasome for degradation, see ubiquitin section below. In conclusion, the base is composed of a hexameric ring where subunits cooperatively work to unfold and translocate substrates to the 20S cylinder, but also may function individually or in smaller complexes for different purposes.

The S5a subunit located between the base and the lid is believed to stabilize the two subcomplexes (Glickman et al., 1998). Indeed, deletion of S5a in yeast causes dissociation of the lid and the base. The entire 26S proteasome displays a strong and specific activity for polyubiquitin chains, but S5a is one of the few individual subunits identified to be able to bind ubiquitin chains through two binding sites termed PUBs (Young et al., 1998) or UIM for ubiquitin interacting motif (Hofmann and Falquet, 2001).

The lid of the metazoan 19S particle is composed of at least eight subunits: S3, S9, S10a, S11, S12, S13, S14, p55 and shares strong similarities with another multiprotein complex, the COP9 (constitutive photomorphogenesis) signalosome (Henke et al., 1999). Several proteins of both complexes contain a characteristic PCI domain, a structural motif important for complex assembly. The subunit S9 physically interacts with subunit CSN2 of the signalosome suggesting a potential link between both complexes (Lier and Paululat, 2002). Besides this homology, specific functions have been assigned to several of the subunits forming the lid. For instance, deletion of Rpn3 subunit in budding yeast (S3) caused a metaphase arrest. The Rpn3 mutants strongly altered
degradation of several substrates by the 26S proteasome such as G1-phase cyclin Cln2, S-phase cyclin Clb5, and the anaphase inhibitor Pds1 that are essential for cell cycle progression (Bailly and Reed, 1999). Additionally, mutational studies of another lid subunit Rpn9 (S11) also showed defects in cell cycle progression in yeast by delaying substrates degradation by the 26S and increasing multiubiquitinated protein pool at 37°C (Takeuchi and Toh-e, 2001). Finally, a mutation in the Rpn11 (S13) gene in yeast results in a cell cycle arrest, an over replication of nuclear and mitochondrial DNA, as well as an altered mitochondrial morphology (Rinaldi et al., 1998). These studies show a fundamental role for the lid of the 19 regulatory particle, especially during the cell cycle progression. However, additional studies are needed to understand the range of functions or specific interactions that the 19S probably possesses.

The ubiquitin-proteasome pathway

Despite the reported ability of the proteasome to degrade several substrates without ubiquitin, such as oxidized proteins (Grune et al., 1997) or ornithine decarboxylase ODC (Gandre and Kahana, 2002), most of the substrates are degraded by the proteasome/ubiquitin pathway. Ubiquitin is a very conserved protein consisting of 76 amino acids (8 kDa) found to be involved in the destruction of a large number of proteins, by covalently attaching to its substrate. Hershko and coworkers proposed a model where ubiquitin acts as a destruction signal in an ATP-dependent manner (Hershko et al., 1980). Once proposed, this model led to an extensive search for the enzymes or complexes responsible for the actual degradation of these ubiquitin tagged proteins. The 26S proteasome was shown shortly after to be involved in this conserved pathway (Hough et al., 1986). Today, ubiquitin conjugation is recognized as a multifunctional signaling mechanism with regulatory significance comparable to phosphorylation. Most of the substrates are efficiently targeted to the 26S proteasome when they are attached to a polyubiquitin chain consisting of at least 4 ubiquitin molecules (Pickart, 2004). Recent studies show that monoubiquitination might serve others purposes, such as re-localization (Stelter and Ulrich, 2003). The ubiquitination of a substrate has been extensively studied and is well understood. It requires three types of enzymes, named E1s, E2s and E3s. The first ubiquitin is activated on its C-terminal by E1s (Ub activating enzymes). The E2 (Ub conjugating enzymes) transiently carries the activated ubiquitin to its corresponding E3 (Ub ligating enzyme). Finally, the E3 ligases are thought to be most directly involved in substrate recognition. The E3s can be divided in two main families: the HECT domain E3s and the RING E3s. Studies in yeast showed a physical association of certains ubiquitin
ligases with the 26S proteasome suggesting a participation in the delivery of substrates to the proteasome (Xie and Varshavsky, 2000).

Polyubiquitinated proteins are then recognized by the 26S proteasome and degraded. The recognition by the 26S is still not known precisely, but the S5a subunit was originally suspected to play a role in ubiquitin chain recognition since it contains two UIM, ubiquitin interacting motifs, characterized by alternative leucine and alanine residues LALAL (Young et al., 1998). This model is in accordance with deletion study of the Drosophila S5a that showed lethality and multiple mitotic defects (Szlanka et al., 2003). However, knockout of S5a subunit in yeast displays a viable phenotype with minor defects in protein degradation (Wilkinson et al., 2000). In that particular case, it was proposed that several subunits might be involved in the recognition, and thus deletion of S5a is rescued by other cellular components. Indeed, the S5a mutant is lethal in combination with mutations of the other proteasomal component encoded-genes Rpn1, Rpn11 and Rpn12 (Wilkinson et al., 2000). Additionally, chemical cross-linking studies in mammalian suggested that one ATPase subunit, namely S6’, interacts with polyubiquitin chain and may thus be involved in the degradation signal recognition (Lam et al., 2002).

In conclusion, despite the large number of papers published concerning the structure and function of the 26S proteasome, many questions remain unanswered. Particularly, the precise mechanism involved in the polyubiquitin chain recognition is not yet well understood, the exact role of ATP in the hydrolysis is still unclear, and specific functions for each 19S subunit remain to be discovered. A number of questions should be resolved with the determination of the whole 26S proteasome structure.
The PA28 activator or 11S REG complex

In 1992, two groups purified and biochemically characterized a complex able to enhance the proteasomal degradation of fluorogenic peptides in an ATP-independent manner (Dubiel et al., 1992; Ma et al., 1992). However, this activator was unable to enhance proteasomal degradation of native proteins, such as bovine serum albumin or ubiquitin-lysozyme conjugates by the 20S proteasome. Native gel electrophoresis of the human activator showed a complex of approximately 200 kDa, and two-dimensional electrophoresis revealed that this complex is composed of two different subunits. Cloning and sequencing of the corresponding genes revealed that the two subunits share about 50% similarities in their primary sequence, and have a molecular mass of 28 and 29 kDa respectively (Mott et al., 1994). Two different names have been attributed to the complex, De martino and his colleagues designated it proteasome activator PA28, whereas Rechsteiner and colleagues named it 11S REG since the regulator has a sedimentation coefficient around 11S. Electron-microscopy studies have shown that the activator occurs as an oligomeric ring able to bind to both ends of the 20S proteasome (Gray et al., 1994), or at only one end of the 20S proteasome whereas the other end is associated to the 19S regulatory particle, and this combined structure has been termed the hybrid proteasome.

Structure and activation of the proteasome

Having a size of approximately 200 kDa, work was undertaken to determine the exact number of subunits forming the complex. Using densitometric quantification of immunoprecipitates obtained with either anti-PA28α or anti-PA28β antibodies, Ahn and coworkers observed a 1:1 stoichiometry and suggested a hexameric structure with 3 alpha and 3 beta subunits (Ahn et al., 1996). This result was confirmed in another study where the pattern of crosslinked PA28 using bis (sulfosuccinimidyl) substrate indicated a hexamer, probably composed of alternating alpha and beta subunits (Song et al., 1996). However, later investigations differed from these initial results. Indeed, coexpression of wild type PA28α and PA28β in bacterial cells revealed a large majority of alpha3-beta4 heptamers using ESI-TOF mass spectrometry (Zhang et al., 1999). The human recombinant PA28α crystal structure determined at a 2.8 Å resolution also confirms a heptameric structure (fig. 4) (Knowlton et al., 1997). The monomer is predominantly helical with four long α-helices. The sequence between the helices 1 and 2 is made of 39 residues forming a flexible loop not resolved on the structure. This disordered...
region has been termed homolog-specific insert since it constitutes the most divergent region between PA28α and PA28β in term of amino acid sequence. Deletion studies showed that this insert is not required for binding and activating the proteasome by PA28α, whereas removal of the PA28β insert reduces binding of this subunit as well as PA28αβ binding to proteasomes (Zhang et al., 1998b). The PA28α insert might play a role in interacting with others protein complexes since it contains a KEKE motif, a stretch of alternating glutamate and lysine residues suggested to be important in protein-protein interactions. This motif was also found in other proteins such as hsp90, hsp70, calnexin and two subunits of the proteasome (Realini et al., 1994). In addition to this insert, the PA28 complex contains two essential parts allowing interaction and activation of the 20S proteasome. Screening single-site PA28α mutants for altered activity identified a region between Arg-141 and Gly-149 critical for proteasome activation. The last ten residues were also found to be important for binding/activation because Pro240 together with Met247 and the C-terminal residue Tyr249 produced inactive heterooligomers when mutated (Zhang et al., 1998a). This result was confirmed by the crystal structure showing that the two regions were immediately adjacent.

**Fig. 4.** Crystal structure of recombinant human PA28α (Knowlton et al., 1997). A. side view. B. top view.

One challenge in the field is to determine the precise mechanism responsible for proteasome activation by PA28 association. As mentioned earlier, the PA28 is able to selectively enhance short peptide degradation but not native proteins. For instance, LLVY-MCA fluorogenic peptide was cleaved approximately 100 fold, whereas the peptide LLE-MCA cleavage was only enhanced 22 fold. The crystallisation of PA28 together with the 20S proteasome would give a clear view of the activation mechanism, but unfortunately no group has succeeded in this task. Instead, Whitby and coworkers managed to crystallise the PA26 together with yeast 20 S proteasome (Whitby et al., 2000). The proteasome
activator PA26 has been isolated from *Trypanosoma brucei* (Yao et al., 1999), and its structure resembles very much the PA28α structure even if the two sequences only share 14% similarity. The result of this PA26-proteasome structure suggested a conformational change of the outer alpha subunits of the 20S particle without affecting the beta subunits conformations. In contrast, several biochemical studies have suggested a change in the beta subunit conformation after PA28 binding. Among these, a study on peptide cleavage of the 20S together with PA28 showed that addition of PA28 does not change the cleavage pattern but rather increase dual cleavage of 25 mer peptides, instead of single cleavage observed with 20S proteasome alone (Dick et al., 1996). The authors proposed a model where the beta subunits were affected by PA28 allowing dual-cleavage to occur. These results however are not incompatible with the model where PA28 affects the degradation only by selective peptide entrance onto its channel. It may also increase the speed of the cleaved product to exit from the proteasome in order to rapidly empty the catalytic centers for faster degradation.

**PA28, immunoproteasome and immunity**

Shortly after the characterization of PA28, several studies suggested a role in immunity. The first evidence appeared when both mRNA and protein levels were found to be strongly induced upon treatment with the immune cytokine γ-interferon (Ahn et al., 1995), (Groettrup et al., 1995). γ-interferon is well known to induce genes involved in MHC class I presentation such as the transporter associated with antigen processing, TAP. Furthermore, the following findings concerning PA28 reinforce this idea of a functional link to immunity:

- The complex has been identified only in animals possessing an adaptive immune system, from zebra fish to mammals.
- Within the cell, the complex is located both in the cytoplasm and the nucleus but upon γ-interferon treatment, the complex is relocalized to specific structures called PML bodies, these bodies have been shown to participate in the immune response (Fabunmi et al., 2001).

In addition to PA28, organisms with adaptive immune system possess three additional proteasomal beta subunits LMP2, LMP7 and MECL1 (also called LMP10). These subunits are able to replace the three constitutive active subunits beta1, beta5, and beta2 after γ-interferon induction (Groettrup et al., 1996). The newly incorporated subunits change the structural and catalytic properties of the proteasome, leading to a new complex termed the immunoproteasome (Aki et al., 1994). Studies focusing on the immunoproteasome assembly revealed a cooperative incorporation of the subunits: MECL-1 requires LMP2 for
efficient incorporation, and preproteasomes containing LMP2 and MECL-1 require LMP7 for efficient maturation (Griffin et al., 1998), (Groettrup et al., 1996) suggesting an important role of all three subunits. The enzymatic activity is also modified by the three subunits but the precise contribution of each catalytic site and the overall modifications remain controversial. The incorporation of LMP2 and LMP7 subunits has been reported to down-regulate cleavage C-terminal of acidic residues and increase the cleavage C-terminal of hydrophobic residues (Driscoll et al., 1993), (Aki et al., 1994). However, other groups did not observe a significant change of the chymotrypsin-like activity (Ustrell et al., 1995). Ehring and coworkers favor a model where the cleavage site was not determined by the P1 position alone, corresponding to the residue situated directly in front of the cleavage site (Ehring et al., 1996). Finally, some reports mention contradictory results with a decrease in the chymotrypsin-like activity of the 20S after γ-interferon treatment (Boes et al., 1994), (Kuckelkorn et al., 2002). These contradictory results may be explained by the differences in proteasome preparations and the different organisms used during the various studies. However, the use of fluorogenic peptides and synthetic polypeptides clearly indicates a difference in activity between constitutive proteasomes and immunoproteasomes.

The fact that the proteasome releases short peptides led several groups to study the proteasomal antigen generation capacity. Several studies established that purified 20S proteasomes were able to produce antigenic peptides (8 to 9 residues) from polypeptides or intact proteins (Dick et al., 1994), (Niedermann et al., 1995). Furthermore, the use of peptide aldehyde inhibitors of the proteasome coupled with introduction of ovalbumin into lymphoblasts demonstrated that the proteasome is essential for the presentation on MHC class I molecules of an ovalbumin-derived (Rock et al., 1994). Many groups evaluated the immunoproteasome contribution in antigen presentation but contradictory results emerged from these studies. First, work on human lymphoblastoid cell lines LMP2 and LMP7 deficient revealed that these two subunits are not essential for antigen presentation (Arnold et al., 1992). Yewdell and coworkers made the same type of observation and showed that cells lacking LMP2 and LMP7 could still present antigenic peptides derived from viral proteins (Yewdell et al., 1994). Nevertheless, LMP2 depleted mice have reduced levels of CD8+ T lymphocytes and generate fewer influenza nucleoprotein-specific cytotoxic T lymphocytes precursors (Van Kaer et al., 1994). In agreement to this last finding, mice with a deletion of the gene encoding LMP7 have reduced level of MHC class I cell-surface expression and present the endogenous antigen HY inefficiently (Fehling et al., 1994). So, despite the lack of marked effects
in cell lines, the data supports an essential role of the immunoproteasome in MHC class I peptide presentation.

The PA28αβ is also strongly induced after γ-interferon treatment and several groups studied the influence of the activator in antigen processing. In order to test this hypothesis, the proteasomal digestion of a 25-mer from murine cytomegalovirus pp89 was assessed in presence or absence of PA28 and the results showed a change both in quality and in quantity of peptides produced (Groettrup et al., 1995). Furthermore, a detailed study on the mechanism of action revealed that proteasomal generation of MHC class I epitopes was optimized in presence of PA28 by inducing a coordinated double cleavage mechanism (Dick et al., 1996). Another clear example of the role of PA28 in antigen processing was obtained by Sun and coworkers who found that melanoma cells lacking both PA28 and immunoproteasomes did not display a specific epitope to CTLs derived from the melanoma antigen tyrosinase-related protein 2 (TRP2). However, PA28 expression completely rescued the epitope presentation (Sun et al., 2002).

These studies mentioned above demonstrate an important function of immunoproteasomes and PA28 in the antigen presentation pathway, but the exact contribution of both complexes and the link between them during immune response remains difficult to appreciate. For instance, mice lacking both PA28α and PA28β showed normal processing of ovalbumin and influenza A virus derived antigens (Murata et al., 2001), (Sun et al., 2002). However, the knockout mice completely lose the ability to process the TRP2-derived peptide as observed by Sun and coworkers (Murata et al., 2001). In addition, the PA28 has been shown in several occasions to be able to enhance antigen presentation of some CTL epitopes independently of changes in proteasome composition (van Hall et al., 2000), (Schwarz et al., 2000).

In conclusion, despite the increasing understanding of PA28 and immunoproteasomes roles in antigen presentation, several aspects of their biological functions remain to be discovered.
**The PA28γamma**

Systemic Lupus Erythematosus (SLE) is an autoimmune disease affecting a large number of organs, particularly joints, skin, kidneys, heart and lungs. This disease is also characterized by an array of autoantibodies directed against the nucleosome, its DNA and histone components as well as other self-proteins such as PCNA (proliferating cell nuclear antigen) or small nuclear RNPs. One of these proteins called Ki autoantigen was identified in 1981 as a soluble nuclear antigen from a Japanese SLE patient (Tojo et al., 1981). In a clinical and serological study carried out by Riboldi and coworkers on 516 subjects, anti-Ki autoantibodies were found in 12% of patients having SLE. Interestingly, this study also showed a significant correlation between Ki and PCNA antibodies (Riboldi et al., 1987).

The use of anti-Ki antibody as a probe led to isolation of the bovine and human corresponding cDNAs. The two proteins share 95% sequence homology and the gene expression is correlated with proliferative conditions, since the mRNA abundance in mouse fibroblast cells greatly increased 4 hours after serum stimulation whereas in quiescent cells mRNA level was almost non-detectable (Nikaido et al., 1989).

By sequence comparison, the Ki autoantigen was found to share strong homology with both subunits of the PA28 activator described earlier. Ki shares 41% amino acid sequence identity with PA28α and 34% with PA28β (Ahn et al., 1995) and was found to coimmunoprecipitate as a homopolymer together with the 20S proteasome and reversibly associate with the 20S on glycerol gradients (Tanahashi et al., 1997). In accordance with these properties, the Ki autoantigen has been renamed PA28γ or 11 S REGγ.

**Structure and activation of the proteasome**

The PA28γ was first purified by Sakamoto and coworkers from rabbit thymus (Sakamoto et al., 1989). They found a molecular mass of 32 kDa for the monomer and 224 kDa for the native protein, so the complex is likely to be heptameric. These results have been confirmed by purification of a 6-his-PA28γ fusion protein and analysis the resulting fractions on SDS-page. In that case, the molecular mass was calculated as 215 kDa confirming the heptameric structure of the ring (Wilk et al., 2000).
Even if the structure of the complex is unknown, the primary sequence possesses all the features for being a PA28 family member and supposes a similar shape as PA28. Indeed, the activation loop is conserved between human PA28α and PA28γ and the essential proline and tyrosine located in the C-terminal region are also conserved. Concerning the homolog-specific insert region, the KEKE motif is not present but there is, around the middle of the loop, an amino acid sequence homologous to a nuclear localization signal (Takasaki et al., 1996) suggesting a distinct function from the hypothetical protein-protein interaction domain.

The enhancement of the proteasomal activity by PA28γ for small peptides is very different compared to PA28αβ. Unlike PA28, binding of the mammalian recombinant PA28γ strongly stimulates the trypsin-like activity of the proteasome whereas the chymotrypsin-like activity was originally reported to be almost unaffected by the presence of PA28γ compared to the activation pattern observed with PA28αβ (Realini et al., 1997). My recent work with Drosophila indicates a more complicated picture with PA28γ functioning as an actual inhibitor to the chymotrypsin-like activity. Surprisingly, a substitution of Lys188, completely changes the activation pattern of PA28γ that becomes similar to PA28 pattern (Li et al., 2001). The fact that the lysine is located almost on the top of the ring implies that either the channel promotes or inhibits peptide transfer to the 20S proteasome or there are long-range conformational changes along the complex that result in changing the beta subunits structure.

Overall, the mammalian PA28γ is able to activate the proteasome but to a lower extent than the PA28 complex (Realini et al., 1997).

**PA28γ function**

In 1989, a first hint of the biological function came with a study performed by Nikaido and coworkers that showed that the expression of PA28γ (termed the Ki antigen by the authors in relation to the formation of autoantibodies in SLE) was growth regulated with an expression pattern similar to c-myc. The level of mRNA was barely detectable in quiescent cells whereas an increase was observed in the cell by 4 hours after serum stimulation. They suggested a role related to cell growth regulation and proliferation (Nikaido et al., 1989). Once the homology between the Ki antigen and the PA28 established, several groups studied the effect of γ-interferon on the level of PA28γ mRNA or protein, in order to evaluate the putative role of PA28γ in immunity. First investigations showed a very slight increase in mRNA levels (Ahn et al.,
1995), (Jiang and Monaco, 1997), but later results contradicted these findings, and two reports concluded a complete loss of PA28γ protein level after γ-interferon without affecting the mRNA level (Tanahashi et al., 1997) and a reduced protein level in liver of lymphocytic choriomeningitis virus-infected mice respectively (Khan et al., 2001) suggesting a function rather independent of the immune response. The most valuable study in the search for the activator function was probably undertaken by Murata and coworkers that generated PA28γ deficient mice. The lack of the complex did not give any appreciable abnormalities to the newborn mice, but resulted in smaller body size when compared to wild-type mice (Murata et al., 1999). The analysis of the PA28γ depleted cultured embryonic fibroblasts performed in the same study revealed growth retardation when compared to wild type cells. Interestingly, PA28γ depleted cells were slightly larger than wild type cells. To examine the cause of growth retardation, Murata and colleagues studied cell cycle progression by flow cytometry. The results showed an increase of cells in G1 and a decrease in S phase for PA28γ -/- cells, suggesting an important function for the entry into S phase (Murata et al., 1999). This apparent role in cell cycle progression agrees well with the observation that cells in various thyroid neoplasms show an abnormal high expression of PA28γ especially where the cells are rapidly growing (Okamura et al., 2003).

**PA28gamma interacting proteins**

A fair amount of information concerning the *in vivo* function of the complex arose from different studies involving protein-protein interactions. Indeed, two-hybrid screen experiments have in several occasions found PA28γ as major binding protein to the bait used, supposing that PA28γ interacts with many cellular factor other than the proteasome, and all these data might give some hints on the intracellular function of the complex. Here is a brief summary of these findings:

- Interaction with Caspases -3 and -7: screening of a human brain adult cDNA library with a constitutive active mutant of caspase 7 led to the isolation of PA28γ. Subsequent work showed that human PA28γ is cleaved by both caspase –3 and –7 through a DGLD cleavage site located in the homolog-specific insert region of human PA28γ. This study showed that human PA28γ is an endogenous substrate for these caspases (Araya et al., 2002).

- Interaction with Emerin: using a high-stringency yeast two hybrid method to screen a human heart cDNA library, Wilkinson and coworkers identified PA28γ as one of 5 candidates interacting with Emerin. Emerin is a nuclear membrane protein that interacts with lamin A/C at the nuclear
envelope. The function of Emerin is largely unknown (Wilkinson et al., 2003).

Interaction with MEKK3: MEKK3 is a MAPK kinase kinase that acts on the JNK and p38 activation pathways that are induced by cellular stress including UV and γ irradiations. The PA28γ was isolated in a two-hybrid screen using MEKK3 as bait. Once the interaction identified, the authors confirmed the relevance of the binding by GST pull-down assay, showed that MEKK3 was able to increase PA28γ protein expression in cos-7 cells, as well as phosphorylate this one. These results suggest a role for PA28γ in response to various stresses (Hagemann et al., 2003). The phosphorylation is of particular interest because the activation of 20S proteasome by PA28 in rabbit reticulocytes has been proposed to require phosphorylation (Li et al., 1996).

Interaction with the hepatitis virus C core protein: once again, the use of two-hybrid screen identified PA28γ as binding factor, this time of the hepatitis virus C core protein. This interaction was later demonstrated in cell culture as well as in the liver of HCV core transgenic mice. Interestingly, knock-out of PA28γ led to the export of the HCV core protein from the nucleus to the cytoplasm, whereas overexpression of the gene increased degradation of the HCV core protein. These results suggest the indirect role or “undesired” function of PA28γ is to interact with the HCV core protein and direct it to the nucleus where it can exert its pathogenesis (Moriishi et al., 2003).

Additional interactions: recently, a large-scale study was performed on 10,623 predicted transcripts from Drosophila (Giot et al., 2003). Using two-hybrid based protein interaction-map, a total of 4679 proteins were found to share 4780 interactions in a high confidence map. Concerning the Drosophila homolog of PA28γ, a total of ten interactions were identified, mainly with unknown genes, but only one with a high confidence. This last one was the interaction with CG3162 (snRNP U2 component), a gene that encodes a product with pre-mRNA splicing factor activity involved in nuclear mRNA splicing, via the spliceosome.

All these various data obtained by yeast two-hybrid method seem to indicate that the PA28γ plays different functions in addition to the regulation of the G1 to S phase transition and give insights in what could be these functions.
Advantages of Drosophila model in our investigation:

**Drosophila melanogaster as model system**
The *Drosophila* genome size is approximately 180Mb whereas the human genome consists of 3200Mb. Although the two genomes have a vast size difference, the number of genes identified is much closer between the two genomes. Indeed, the number of genes in human is estimated to vary between 30000 and 40000 whereas *Drosophila* may possess 13600 or more genes.

**Drosophila embryogenesis**
The fruit-fly *Drosophila melanogaster* is probably the most popular model system for the study of developmental processes for several reasons. First, *Drosophila* is very convenient and powerful as a model organism because it has short life cycle of two weeks, making it possible to study numerous generations in a relative short time. Also, it is large enough that many attributes can be seen with the naked eye or under low-power magnification. Moreover, it has a very long history in biological research (since the early 1900s) and there are many useful tools to facilitate genetic study, including a detailed literature available describing the various stages of the *Drosophila* development.

**RNAi techniques**
RNA interference (RNAi) has emerged as a powerful tool for the silencing of gene expression in animals. RNAi is mediated by small interfering RNAs (siRNAs) 21 to 25 nucleotides long originally produced from larger double stranded RNAs in vivo through the action of Dicer (Agrawal et al., 2003). In the second step, the siRNAs join an RNase complex, RISC (RNA-induced silencing complex), which acts on the cognate mRNA and degrades it. Because of its specificity and efficiency, RNAi is being considered as a simple and rapid method not only for functional genomics, but also for gene-specific therapeutic activities that target the mRNAs of disease-related genes.

**Drosophila model for transcriptional regulation**
In this section, I will present the opportunities of working in transcription regulation using *Drosophila* as model system, and I will mainly focus on two aspects using examples relevant to my work:
- First, transcription factors are well characterized in this organism, and a lot of available information is unique and has not yet been discovered in others species, or only partially. The well-defined DRE/DREF system is one example,
-second, the pathways are typically simpler than in mammals with fewer genes involved making Drosophila a powerful tool to study transcription. The E2F transcription factor family is well suited as an example of this characteristic.

The DRE/DREF System

In 1993, Hirose and coworkers compared the promoters of two genes involved in DNA replication, PCNA (proliferating cell nuclear antigen) and DNA polymerase α. This sequence analysis led to the identification of a conserved palindromic sequence of 8 nucleotides, TATCGATA, present in both promoters (Hirose et al., 1993). This motif was termed DRE standing for DNA replication related element based on the important roles that PCNA and polymerase α have in DNA synthesis. The efficiency of DRE elements has been evaluated and confirmed both in transgenic flies and cultured cells (Hirose et al., 1999). Gel mobility shift assays detected the presence of a protein factor in nuclear extracts of cultured Drosophila Kc cells able to specifically bind to DRE sequences. The factor, named DREF, is a transcription factor regulating proliferation-related genes in Drosophila by binding to DRE regulatory elements. DREF is a protein of 701 amino acid residues with a molecular weight of 80 kDa (Hirose et al., 1996). Deletion analysis indicated that the DNA binding domain was located in the basic amino acid-rich region between residues 16 and 105. Immunocytochemical analysis demonstrated the presence of DREF polypeptides in the nuclei after the eighth nuclear division cycle, suggesting that a nuclear accumulation of DREF was important for the coordinate zygotic expression of DNA replication-related genes carrying DRE sequences (Hirose et al., 1996).

The E2F pathway

The E2F transcription factors play various roles in controlling entry and progression through the S phase of cell cycle, apoptosis and differentiation. The S phase, where DNA synthesis occurs, requires the activation of a large set of genes. The E2F transcription factor family regulates many of these genes.

Human E2F family:
The human family of E2F transcription factors is composed of at least 6 members, E2F1-6, and each one heterodimerizes with one of the three DP proteins (DP1-3). All combinations of E2F/DP complexes can exist in-vitro (Dyson, 1998). Based on their transcriptional properties as well as their sequence homologies, the E2F family can be divided into three subgroups. E2F-1, E2F-2, and E2F-3 are bound to their target genes in vivo when expressed during late G1 and S-phase, activating them and
driving cellular proliferation (DeGregori, 2002). Overexpression of each protein is sufficient to drive resting cells into cell cycle (DeGregori et al., 1997) and depletion of all three at the same time prevents entry into S-phase. For these properties this subgroup has been classified as activator E2Fs and their expression is maximal as cell approach G1-S phase boundary. On the contrary, E2F4 and E2F5 are uniformly expressed throughout the cell cycle with significant levels during G0, and they preferably bind to their target genes in resting cells and during early G1 when transcription is turned off. These observations suggest that they are mainly involved in repression during early cell cycle progression (Moberg et al., 1996) and for that reason have been termed repressor E2Fs. Finally the E2F6 shares only the core DNA binding domain and dimerization domain but lacks the C-terminal transactivation and the pocket protein binding domains (Morkel et al., 1997). Recent studies revealed that E2F6 is able to recruit multiple factor to form a complex targeting the E2F responsive elements. Among these factors, chromatin-modifying proteins can be recruited primarily in quiescent cells (Ogawa et al., 2002). This result suggests a role in transcription silencing in quiescent cells by modifying the structure of the chromatin.

*Fly E2F family*

The *Drosophila* E2F family is composed of only two members E2F1 and E2F2 and only one gene coding for the DP complex has been identified. The E2F1 transcription factor shows considerable homology to human E2F1 with over 65% identity in the DNA binding region (Ohtani and Nevins, 1994). The factor was shown to bind specifically to the E2F recognition site, and to activate transcription of a reporter gene located downstream of the polymerase α gene promoter that contains E2F recognition sequences. In addition, the E2F1 is required for S phase during *Drosophila* embryogenesis (Duronio et al., 1995), reinforcing the idea of a true homolog of the human counterpart. In contrast, DrosE2F2 represses the transcription of E2F reporters and the loss of DrosE2F2 results in increase of gene expression (Frolov et al., 2001). Therefore, the two types of complexes seem to play antagonistic roles in *Drosophila* and acts on the cell cycle the same way the E2F family in human regulates cell proliferation, but with fewer members.
AIM OF THE WORK

The proteasome field has expanded dramatically the last few years, especially because of its involvement in many intracellular pathways. Furthermore, the use of proteasome inhibitors for cancer patients or other diseases is an approaching treatment, and has already started in several cases. In order to use the proteasome as a target in the treatment of various diseases, the specific functions must be investigated and should be known or at least evaluated to minimize possible side effects. The main objectives of our laboratory are to use the advantages of a well-established model system Drosophila to investigate the functions of the different proteasome regulators identified in these organisms. The main focus of my project was the study of the Drosophila PA28 (or REG) homolog, a unique gene coding for a putative PA28 (REG) member isolated from a Drosophila cDNA library. The secondary objectives involved the generation of a Drosophila in vivo assay to monitor the 26S proteasome degradation, and the genomic search for global proteasomal regulation mechanisms using sequenced promoters available from the FlyBase database of the Drosophila genome projects (2003).
RESULTS AND DISCUSSION

Study of the invertebrate PA28 homolog

The Drosophila PA28: A homolog of human PA28γ (Paper I)

The first part of the project consisted in characterizing the Drosophila homolog of the PA28 gene family identified in mammals. At that time, only one group (Paesen and coworkers) had identified a PA28 homolog in invertebrates, *Rhipicephalus appendiculatus*, sharing 55% identity with the human PA28γ primary structure (Paesen and Nuttall, 1996). No biochemical study had been performed and the work only mentioned the identification of a cDNA. In *Drosophila*, the sequencing of the genome was advancing but no clear sequence shared significant similarity with PA28 genes. However, a cDNA from an adult cDNA library was identified as a PA28 homolog. We subcloned it and started our investigation at that point.

Characterization of Drosophila PA28

The purification of the recombinant protein using DEAE and gel filtration chromatography techniques revealed a rather large complex of around 200 kDa, in accordance with previous reports (Wilk et al., 2000), suggesting that the bacterial recombinant was able to form a polymer. In order to perform biochemical assays, *Drosophila* 20S proteasome was also purified from a large batch of fly embryos. Fluorogenic peptide assays demonstrate that our activator shares similar preference for activating the proteasome trypsin-like activity in a similar manner as the human PA28γ. The proteasomes and PA28 are well conserved between species, especially in their interaction domains. This finding probably motivated Whitby and coworkers to resolve the structure of yeast proteasome together with the *Trypanosoma brucei* PA26 (Whitby et al., 2000). In order to investigate the properties of the *Drosophila* activator, a mammalian proteasome was also tested in our assays, and the trypsin-like activity was stimulated but to a lower extent (5 fold compared to 22 fold for *Drosophila* 20S). Surprisingly, the chymotrypsin-like activity monitored by the degradation of LLVY-MCA fluorogenic peptide was decreased in presence of the *Drosophila* activator. This inhibition was the first example where a PA28 member could function as a positive and negative regulator of proteasomal
activity. This is the reason why we named it dREG for *Drosophila* proteasome regulator. For clarity reasons, the term PA28 will still be used here, since PA28 nomenclature is more common in the scientific community, and dREG will be employed for the *Drosophila* complex. Interestingly, the chymotrypsin-like activity of the mammalian proteasome was also inhibited by dREG. This inhibitory effect was later observed by Li and coworkers with recombinant human PA28γ (Li et al., 2001). Surprisingly, they also observed that substitution of Lys188 to Asp188 or Glu188 abolished this effect, and the mutant showed the exact same activation pattern as PA28. They suggested a change in the 20S proteasome beta-subunit conformation induced by the mutation. One of the main challenges concerning the PA28 family will be to understand precisely the way PA28 activates the proteasome. So far, contradictory reports do not put forward a clear mechanism, but rather suggest a contribution of both the regulator inner channel and conformational changes in the 20S proteasome complex.

**Homolog-specific insert regions and PA28 localisation**

Based on the fluorogenic assays, the dREG activation pattern resembled the PA28γ rather than the PA28, with a major activation of the trypsin-like activity. In order to better characterize the *Drosophila* homolog, we performed immunofluorescence with a polyclonal antibodies raised from the recombinant protein injected in a rabbit. The results clearly showed a major nuclear localization whereas none or very little staining was observed in the cytoplasm. This observation reinforces the homology to vertebrate PA28γ that also localizes in the nucleus whereas PA28αβ is mainly found in the cytoplasm (Soza et al., 1997). Interestingly, a putative monopartite nuclear localization signal (NLS) was found within the insert specific region forming the flexible loop (see introduction) in *Drosophila*. The signal is usually composed of 3 or 4 amino-acids, with lysine or arginine at positions 1, 2, and 4, that are recognized by importin-α. The mutation of the signal from KRQR to SSQS was done by site-directed mutagenesis, and the obtained plasmid was transfected into *Drosophila* mbn-2 cells. Our results showed a strong cytoplasmic staining supporting an important role of the signal in the localization of the complex. Primary sequence alignments showed that the signal is conserved in all organisms possessing a PA28γ at approximately the same position (always within the flexible loop).

From these experiments, it is clear that the homolog-specific insert region plays other functions than contributing to the binding of PA28
complex to the proteasome as originally proposed by Rechsteiner and colleagues (Zhang et al., 1998b). The sequence of PA28α and P28β possess an additional KEKE motif that may serve other purposes. The presence of these nuclear signals, highly conserved between species attest to an important physiological function. The distribution of proteasomes in the cell is very dynamic and varies during different processes, such as the increase of nuclear proteasomes during the cell cycle from G1 to the nuclear envelope breakdown in early mitosis. One might speculate a contribution of the PA28 family in the overall 20S cellular distribution since little is known about the mechanisms responsible for the balance between the nuclear and cytoplasmic pools of proteasomes. In that sense, the PA28 might bind the 20S proteasome directly after being assembled and then re-localised it where the proteasome should play its specific function. The observation of the capase-7 interaction with PA28γ agrees with this hypothesis. Indeed, proteasomes are removed from the nucleus during apoptosis and accumulate within apoptotic blebs. In vitro human PA28γ subunits are cleaved by both caspase-3 and –7, in the specific sequence DGDL, showing that PA28γ is an endogenous substrate of caspases –3 and –7 (Araya et al., 2002). Interestingly, the specific cleavage site is located very close to the NLS, four amino acid before the first lysine of the NLS. The cleavage can thus destroy the ability of the PA28γ to be transported inside the nucleus, directing the proteasome out of this structure. The homolog specific insert region has been shown not to play any role in proteasome binding and activation for PA28α (Zhang et al., 1998b). We did test this hypothesis in Drosophila by looking at dREG protein levels in Drosophila cells undergoing apoptosis after staurosporin treatment, but we could not observe the specific cleavage found in human described above (unpublished data), meaning that the mechanism might not be conserved between species. This hypothesis concerning a role for PA28s in directing the 20S proteasome to specific structures has also been proposed by Zhang and coworkers who proposed that the insert region of PA28 could serve to couple the calnexin TAP-MHC class I complexes in the endoplasmic reticulum, promoting transfer of antigenic peptides to the MHC class I molecules (Zhang et al., 1998b).

Study of the dREG promoter (Paper II)

Identifying gene regulation at the transcriptional level by studying promoters is a powerful tool to understand the function of a gene. The fully sequenced Drosophila genome allowed us to quickly analyze the promoter region located upstream of dREG. The promoter region is
about 926 bp long and the two adjacent genes strawberry notch (sno) and dREG, are arranged in opposite directions with both promoter regions within a short intergenic region. This region includes two typical DRE motifs, an 8 basepair palindromic sequence where the *Drosophila* replication related transcription factor is able to bind (see introduction). In addition, the first intron possesses a typical E2F binding site (fig. 5).

**Fig. 5.** Schematic representation of dREG promoter region. The putative transcription factor-binding sites are represented as well as their positions relative to the longest 5´transcript found.

The first step in the promoter analysis was to determine the transcriptional start of the gene. To our surprise, we identified 4 different transcriptional starts located close together within a 35 bp region. This promiscuous transcriptional initiation may be explained by the absence of a TATA-box or DPEs (downstream promoter elements) associated with TATA-less promoters upstream of dREG gene.

After this initial step, the two DRE elements as well as the E2F binding motif were tested by cloning the upstream region of dREG gene in front of the LacZ reporter gene. Site-directed mutagenesis was carried out on the proposed elements, and β-galactosidase was assayed after transfection of each of the constructs in mbn-2 cells. The results showed a complete abolishment of the β-gal activity when the first DRE sequence was mutated, whereas mutation of the downstream DRE was found not to affect the expression of the reporter. These results demonstrated that the first DRE site is essential for transcription of dREG in mbn-2 cells. We were surprised to observe that the second site had absolutely no effect on transcription. However, similar results have been reported in the literature: Choi and coworkers identified three DRE motifs upstream of the *Drosophila* TBP gene and mutation of one of these motif did not decrease luciferase activity whereas the two other did significantly (100% and 80% of reduced expression respectively) (Choi et al., 2000). The other site might serve other purposes. Indeed, BEAF, a factor involved in chromatin remodeling, has been shown to compete with DREF and bind
to the sequence CGATA which is also part of the DRE motif TATCGATA (Hart et al., 1999). However the BEAF activity, which requires three close CGATA motifs, was impossible to monitor by transient transfection experiments (Hart et al., 1999), so the DRE motif present in dREG promoter whose mutation did not affect the β-gal activity might be used by such factors other than DREF.

Concerning the E2F element, we found a surprising increase of β-gal activity when the E2F site was mutated. The E2F family in Drosophila is composed of two members that play antagonists roles. Drosophila E2F1 is required for S phase progression whereas E2F2 inhibits transcription of gene possessing E2F motif in their promoter (Frolov et al., 2001). The site present in dREG first intron may be occupied in majority by the E2F member that inhibit the transcription of the gene. Thus, mutation of the site would rather show an increase of the β-gal activity. The fact that several members of a family are capable to bind to the same exact site makes the result harder to interpret. For instance, the Drosophila polymerase α gene promoter possesses three E2F binding sites that act as stimulators in Kc cells, but two of them act rather negatively on promoter activity during development (Yamaguchi et al., 1997). The E2F activity depends also highly on the cell cycle phase. E2F sites in mammalian promoters can act negatively in G0 and G1, but release of RB protein from the complex act as a positive effect increasing the transcription. Therefore, the in vivo situation is more complex than what we observed in cell lines. However, we can still conclude that DREF and E2F both act on dREG promoter suggesting a role in G1 to S phase transition and maybe related to DNA replication. Interestingly, two recent publications suggest a similar transcriptional activation in other organisms. Microarray experiments on E2F1 in mice fibroblast cells indicated PA28γ as positive candidate (Ma et al., 2002b) meaning that the E2F activation of PA28γ RNA is conserved between species. Another recent study revealed the presence of a homolog to DREF in the human genome, with a similar putative DRE element (Ohshima et al., 2003). The PA28γ possess such a site in its promoter confirming a conserved activation pathway from Drosophila to human.

In parallel with our attempts to characterize the dREG promoter, we also analyzed the promoters of the Drosophila proteasome genes for DRE motifs. Surprisingly, several but not all of the subunits possess one or two DRE motifs in their proximal promoter regions. One of the few genes that contain multiple DRE sequence is the 20S β-2 gene. In Drosophila, two additional β-2 isoforms have been identified β2R1, and β2R2. Their promoters do not contain any DRE motifs, and northern-blot analysis revealed that the isoforms are male-specific expressed (Ma et al.,
In general, DREF might contribute at least in part, to varying the expression levels of certain proteasomes genes thereby allowing proteasome structural heterogeneity. Alternatively, the DREF-induced subunits may serve a specific function during the G1 to S phase transition independent of the proteasome. Future work should help in understanding overall proteasome regulation.

The Dictyostelium PA28 homolog (Paper IV)

*Dictyostelium* is a powerful system for research in cell and developmental biology. The organism has unique advantages for studying fundamental cellular processes including cytokinesis, motility, phagocytosis, chemotaxis, signal transduction, and different aspects of development (Eichinger and Noegel, 2003). One of the main advantages of this model system resides in the fact that *Dictyostelium* grows as a separate, independent cell but interacts to form multicellular structures when challenged by conditions such as starvation. Furthermore, these approaches have benefited from bioinformatics resources including online databases. In a *Dictyostelium* cDNA developmental library representing all the mRNAs expressed when the organism becomes multicellular, we surprisingly identified a single putative gene that had high similarities to PA28s. Cloning of this gene and overexpression in *E. coli* system led to purification of a large complex suggesting a similar size to native PA28s. The use of fluorogenic peptides showed that the complex has the capacity of activating purified 20S proteasome from *Dictyostelium*, with a major activation of the trypsin-like activity monitored by LLR-MCA fluorogenic peptide. We also performed immunofluorescence experiments with rabbit polyclonal antibodies raised against the protein. The staining was observable almost exclusively in the nucleus. Examination of different developmental stages (before or after starvation, meaning in both monocellular or multicellular forms) also showed a nuclear staining of the *Dictyostelium* complex. These data suggested a true homolog of PA28γ in accordance with what we observed in Drosophila (see results above), the nuclear complex PA28γ probably arising first in the evolution before the PA28 (made of α and β subunits) complex itself only found in organisms with adaptive immune system (Fig. 6).
**Fig. 6.** Schematic representation of the main features of a eukaryote supertree constructed from 100 genes by Bapteste et al. (2002), with the distribution of REG α, β and γ overlaid. The tree demonstrated the clustering of amoeboid lineages (here represented by *Dictyostelium*) as a monophyletic grouping, the Conosa, distinct from plants, animals & fungi. The arrow indicates the root of the eukaryote tree as suggested from the distribution of DHFR-TS fusion genes (Stechmann & Cavalier-Smith 2002). The DHFR-TS fusion suggests the root is located between Opisthokonts and Amoebozoa and, if correct, suggests that REG γ was present in the ancestral eukaryote, and that it has been lost from lineages such as Fungi and Plants. Coloring indicates which eukaryote supergroup the represented lineages/groups belong to.
**Investigating the function of PA28gamma**

We decided to start our investigation by looking at the *Drosophila* complex after induction of the immune response. Since the PA28 present in higher eukaryotes has been shown to be involved in the immune response, we speculated that the ancient form might as well serve a related function in invertebrates. In order to test this hypothesis, we treated mbn-2 cells with lipopolysaccharides (LPS), a compound present in the outer cell membrane of gram negative bacteria that induces an immune response. Our results did not show an increase in dREG expression neither at the RNA or the protein level (Paper I), indicating that dREG is not induced after immune response as PA28 does with γ-interferon.

The second part, done in relation to the promoter analysis described above, was to knock down dREG from *Drosophila* cells and determine the effects on the cells cycle by FACS analysis. This experiment was motivated by the result from Murata and coworkers who observed a G1->S phase inhibition in mouse embryonic fibroblasts lacking PA28γ (Murata et al., 1999). The results obtained from our experiments clearly showed the same change in the G1->S phase transition in *Drosophila* but the effect was more pronounced in the insect system (fig. 7). The two studies agree on the important function of PA28γ on the cell cycle progression.

![Fig. 7](image-url)  
**Fig. 7.** Left panel, comparison of the cell cycle transition defects between PA28γ knock-out in mouse embryos fibroblasts (adapted from Murata et al.) and right panel, knock-down experiment of the *Drosophila* homolog (our study) using RNAi.

Our most recent attempt to gain insight into the biological significance of the complex consisted of immunostaining of Drosophila embryos with polyclonal antibodies raised against dREG in order to follow the expression profile during embryogenesis. Indeed, gene transcriptional expression patterns have been measured at a genomic
scale during a complete time-course of *Drosophila* development and dREG transcriptional activity showed highest levels during embryogenesis. Staining of syncytium early embryos revealed a disperse localization with no apparent concentration around the chromosomes suggesting no direct interaction with DNA (Paper II). The appearance of nuclear envelopes lead to the nuclear localization of dREG. Later, *Drosophila* embryos go through gastrulation, a process involving four major morphogenic events: ventral furrow formation, posterior midgut invagination, cephalic furrow formation and germband extension (Costa et al., 1994). The first three events are driven by cell shape changes. During this phase, the invaginating cells showed increase levels of dREG. Interestingly, a recent screen to identify gastrulation changes identified three proteasome subunits which had altered abundance protein levels in ventralized versus lateralized embryos and RNAi knockdown of these subunits caused ventral furrow defects, confirming the role of these proteins in ventral furrow morphogenesis (Gong et al., 2004). The combination of our past work and this recent finding supports an important role for the proteasome and dREG during gastrulation that would be worth examining in more detail. After this step, dREG was found uniformly in all nuclei except during germ band elongation where a cytoplasmic localization was apparent in discrete domains. These domains were identified to be mitotic domains that consist of specific cell clusters that undergo mitosis in a synchronous manner (Foe, 1989). Thereby, the cytoplasmic staining results of the nuclear envelope breakdown, again demonstrating a function without direct interaction with DNA.

**Concluding remarks**

Overall, PA28γ seems to be the most ancient member of the PA28 family. PA28 arose probably from PA28γ by gene duplication and became an efficient activator recruited by the adaptive immune system. So what is role of the ancient form? It is difficult to draw any picture from the activation results in experiments using fluorogenic peptides, since the compounds used in fluorogenic assays including MCA (methyl-couramin) are relatively large when compared to the small polypeptides made of three or four residues. That of course renders difficult any interpretation. Results involving more "natural peptides" like use of 25-mer by Li et al. showed that PA28γ is helping the proteasome simply by increasing the number of smaller peptides released by the complex. However, we still can describe the main properties of this complex based on our study as well as the literature:
In the cell, PA28γ is nuclear (completely or in a very large percentage) and several experiments showed that:
1- The complex is able to increase proteasomal degradation, with a major effect on the trypsin-like activity, whereas it inhibits the chymotrypsin-like activity.
2- The complex is induced and required for proper G1->S phase transition, from Drosophila to human.
3- The Drosophila possess a DRE site and may possess E2F site important for transcription regulation, and the mechanism is likely to be conserved since human E2F1 activates PA28γ as well.
4- The removal of PA28γ in mouse and in Drosophila induced a growth retardation from G1->S.

So what could be the function of the activator in this transition? This section is based upon a speculative interpretation of the current data. One can speculate a role for PA28γ in the removal of the long fragments generated by the 26S proteasome that could potentially inhibit the cell cycle progression. This involvement of PA28γ in cell cycle regulator degradation has also been proposed by Murata and colleagues that suggested a role in p21 or p27 degradation (Murata et al., 1999). These suggestions need of course to be tested on the bench to be validated.

An explanation could originate from the structure of certain activator or inhibitors of the cell cycle. For instance, p21 is a mammalian cell cycle inhibitor, involved in the inhibition of the G1 phase. Although this protein is not known in Drosophila, the same type of system probably exists (Avedisov et al., 2001). Recent work suggested that p21 possess specific regions that are essential for its function. For instance, PCNA is important for cell cycle progression and p21 can modulate its activity through its C-terminal domain (Zheleva et al., 2002). Furthermore, Chen and coworkers identified a 39 amino acid fragment of p21 that was sufficient to bind to PCNA and partially inhibit DNA replication in vivo (Chen et al., 1996). It is now clear that the proteasome is involved in the degradation of many of these cell cycle regulators including p21 (Bloom et al., 2003). A putative role for PA28γ complex could be to activate the trypsin-like activity of the 20S proteasome in order to further process the degradation products. This mechanism would avoid cell cycle inhibition as a result of release of longer polypeptides by the proteasome that could still inhibit cell cycle progression, such as p21 polypeptide that has been shown to inhibit PCNA activity.
An in vivo assay to study the Drosophila 26S proteasome
(Paper III)

My involvement in this project concerned the cloning of three different constructs from mammalian to a constitutive Drosophila vector pAc5.1 and generating stable Drosophila S2 cell lines. Originally, Dantuma and coworkers developed a convenient reporter system to monitor proteasome activity in living cells by using short-lived substrates (Dantuma et al., 2000). The assay consists of a GFP molecule fused to ubiquitin (fig. 8).

![Ubiquitin-GFP](image)

- X=M for resistant control fusion
- X=R for N-end rule pathway
- X=G76V for UFD pathway

**Fig. 8.** Schematic representation of the fusion genes generated in order to monitor the 26S proteasome activity in vivo. X represents the residue that differs from the three constructs.

GFP was originally derived from a jellyfish protein that has the property to emit autofluorescence when stimulated at the correct wavelength. This stable protein with a half-life over 24 hours became a powerful tool to monitor cellular localization of substrates when fused to this protein. The constructs presented here also possess ubiquitin, a 76 residue conserved protein that is responsible for targeting substrates to degradation by the 26S proteasome. As shown on fig. 8, three different plasmids have been generated according to the residue placed in front of the GFP protein:

1- the **M construct** generates a stable control fusion substrate that does not contain a degradation signal upon ubiquitin cleavage and is stable in cells.

2- the **R construct** generates an N-end rule pathway substrate. Indeed, once the ubiquitin is cleaved, the argine becomes the N-terminal residue and has been shown to be fastly degraded through the N-end rule pathway (approximately 2 minutes when R is in front of β-galactosidase) (Bachmair et al., 1986).
3- The G76V construct generates a mutated uncleavable ubiquitin. In this case the construct is designed to function as a ubiquitin fusion degradation (UFD) substrate. In other words, the deubiquitination of the Ub fusion is completely inhibited if the C-terminal Gly76 is substituted by another residue. Thus, the non-removable ubiquitin will result in a fast multiubiquitination and will lead to rapid degradation of the GFP by the 26S proteasome.

The three constructs generated upstream of a constitutive promoter and stably expressed in Drosophila cell lines have been very useful to monitor the performance of the 26S proteasome when challenged by different treatments. Indeed, the proteasome activity can easily be measured by the described in vivo assay after the use of proteasome inhibitors including MG132 or RNAi interference against various proteasomal subunits, or other cellular component that may compromise the proper function of the proteasome when depleted from the cells.

**Study of the proteasomal regulation after MG132 treatment (Paper V)**

Paper III, study of the RNAi knockdown of S5a and S13 subunits revealed a strong induction of 20S and 19S proteasomal subunits. This interesting result was also observed by DeMartino and colleagues who found the same activation pattern (Wojcik and DeMartino, 2002). In order to investigate the mechanism responsible for this induction, we amplified by PCR and subsequently cloned three 26S proteasome subunit promoters namely alpha2, beta2 (20S proteasome) and S2 (19S proteasome), as well as Ter94 promoter (non proteasomal gene that is also induced after MG132 or S5a depletion; see Paper V) upstream of a LacZ reporter vector in order to perform transient transfection experiments. The β-galactosidase levels monitoring the promoter efficiencies showed an increase after MG132 induction, as expected from the micro-array data. Interestingly, when the 5´UTR of the α2 and S2 promoters were removed, this induction was lost suggesting a role of this region in the induction of the mRNA. The next step consisted in replacing the dREG 5´UTR from the promoter construct generated in the study of dREG promoter, by the S2 5´UTR, to observe if the induction was still present. Preliminary results suggested that the dREG promoter acquired the ability to be induced after MG132 treatment strongly supposing the presence of an essential regulatory motif in the S2 5´UTR. Sequence alignment using the MEME program revealed that the Drosophila genes
coding for the proteasomal subunits may have a conserved motif within their 5´ UTR (fig. 9).

<table>
<thead>
<tr>
<th>NAME</th>
<th>STRAND</th>
<th>START</th>
<th>P-VALUE</th>
<th>SITES</th>
</tr>
</thead>
<tbody>
<tr>
<td>S5a</td>
<td>-</td>
<td>105</td>
<td>3.60e-07</td>
<td>ATTTTCGTAT TTTTTGTTTGTG CACTTTGCCAC</td>
</tr>
<tr>
<td>beta5</td>
<td>+</td>
<td>35</td>
<td>9.67e-07</td>
<td>A0GCAACGAA TTTTTGTTTGTG CATCGGATTT</td>
</tr>
<tr>
<td>beta2</td>
<td>+</td>
<td>9</td>
<td>9.67e-07</td>
<td>TTTTACTAA TTTTTGTTTGTG AGTCGCCCTG</td>
</tr>
<tr>
<td>S3</td>
<td>-</td>
<td>154</td>
<td>2.96e-06</td>
<td>CCGACTTTC TTTTTGTTTGTG TCAAATCCCA</td>
</tr>
<tr>
<td>S10a</td>
<td>-</td>
<td>59</td>
<td>8.31e-06</td>
<td>CAAAAACCA TTTTTGTTTGTG TCAAATCCCA</td>
</tr>
<tr>
<td>S12</td>
<td>+</td>
<td>58</td>
<td>8.95e-06</td>
<td>GTTAAATCG TTTTTGTTTGTG AGTTTGCCCT</td>
</tr>
<tr>
<td>S6prime</td>
<td>-</td>
<td>32</td>
<td>1.03e-05</td>
<td>CTTCTAAATT TTTTTGTTTGTG AGCCCAAGTC</td>
</tr>
<tr>
<td>beta3</td>
<td>-</td>
<td>38</td>
<td>1.73e-05</td>
<td>DAAGATGACT TTTTTGTTTGTG CTCTGCGTGT</td>
</tr>
<tr>
<td>alpha2</td>
<td>+</td>
<td>26</td>
<td>1.73e-05</td>
<td>GCTGTGT CAT TTTTTGTTTGTG TCGAATATAA</td>
</tr>
<tr>
<td>S8</td>
<td>-</td>
<td>14</td>
<td>1.99e-05</td>
<td>ATGACTCTTG TTTTTGTTTGTG GAGATGACAC</td>
</tr>
<tr>
<td>S12</td>
<td>+</td>
<td>1</td>
<td>2.06e-05</td>
<td>TTTTTGTTTGTG TTTAAATAAAA</td>
</tr>
<tr>
<td>beta7b</td>
<td>+</td>
<td>25</td>
<td>2.92e-05</td>
<td>TGACTCTTGA TTTTTGTTTGTG CCCGTCAATT</td>
</tr>
<tr>
<td>beta4</td>
<td>+</td>
<td>3</td>
<td>3.17e-05</td>
<td>TA ATTTTTGTTTGTG CCTGATTCG</td>
</tr>
<tr>
<td>alpha4</td>
<td>-</td>
<td>17</td>
<td>4.37e-05</td>
<td>ATACCT TTTTTGTTTGTG GAAATGCCCA</td>
</tr>
<tr>
<td>S7</td>
<td>-</td>
<td>64</td>
<td>4.91e-05</td>
<td>TTATGCGATG TTTTTGTTTGTG TTATGCGG</td>
</tr>
<tr>
<td>beta6</td>
<td>+</td>
<td>26</td>
<td>5.46e-05</td>
<td>TTTGTTTGTT TTTTTGTTTGTG GAGTTAAATG</td>
</tr>
<tr>
<td>alpha3</td>
<td>+</td>
<td>39</td>
<td>5.77e-05</td>
<td>AAGCAACGAC ATTTTTGTTTGTG ATTCACTGTA</td>
</tr>
<tr>
<td>S12</td>
<td>-</td>
<td>112</td>
<td>6.55e-05</td>
<td>TTCTCAGGAC TCTTTGTTTGTG GCTAAATTAC</td>
</tr>
<tr>
<td>S9</td>
<td>+</td>
<td>42</td>
<td>7.42e-05</td>
<td>TTCTTGCTAT TTTTTGTTTGTG AGTGTTTTT</td>
</tr>
<tr>
<td>S4</td>
<td>+</td>
<td>89</td>
<td>7.42e-05</td>
<td>TTCTTGCTAT TTTTTGTTTGTG AGTTTTTTT</td>
</tr>
<tr>
<td>S3</td>
<td>+</td>
<td>138</td>
<td>9.42e-05</td>
<td>TACCTAGCAT TTTTTGTTTGTG AGTGAAAGAA</td>
</tr>
<tr>
<td>S6</td>
<td>+</td>
<td>48</td>
<td>1.68e-04</td>
<td>CGCAGTCTGCT TTTTTGTTTGTG AGTGATCATT</td>
</tr>
</tbody>
</table>

**Fig. 9.** Result of the MEME motif search on twenty-two proteasomal 5´ UTRs. The search program identified a consensus sequence that might be of importance for the transcriptional up-regulation seen on the micro-array experiments as well as the transfection assays (see Paper V).

In conclusion, the putative motif present in the gene 5´ UTRs may induce transcription of the genes after proteasome inhibitor treatment either on the DNA level or the RNA level. Our group is currently investigating the mechanism responsible for this transcriptional up-regulation.
CONCLUSIONS AND PERSPECTIVES

The proteasome has a wide range of activities within the cell, and this thesis discussed some aspects of them. The identification of a Drosophila homolog of PA28γ showed that this gene is conserved in invertebrates and seems to share the same function related to cell cycle progression. Currently, we are unable to point out the precise biological function of the regulator, but we have instead gained insights into its role. The deletion of the homolog gene in Dictyostelium should be of great help to reveal the biological relevance of the complex since we can easily monitor developmental or chemotaxis defects produced by the removal of a gene of interest in this organism. The study of the dREG promoter showed that the transcriptional control of the gene was governed at least in part by DREF and E2F transcription factors and that the mechanism of activation is probably conserved in humans. Other aspects of the proteasomal regulation have been addressed in this thesis and future work in the laboratory will surely provide insight into the co-regulated transcriptional activation of all the proteasome genes.
ACKNOWLEDGMENTS

Thanks to all the persons who supported me directly or indirectly during this thesis.

I would like to especially thank:

-Patrick Young, my supervisor for always offering me help and guidance when necessary, and for also letting me independent at other times. Thanks also for the great help during the preparation of this thesis.

-Josefin Lundgren, my coworker during most of the five years, thanks for your great enthusiasm and our long discussions. I hope that I will be as sharp as you one day (and thanks for the organization lessons of course...).

I also would like to express my gratitude to the persons not from our lab that helped me during this period:
- Ylva and her group, especially Gunnel for all the Drosophila work, thanks for teaching me very useful Drosophila techniques.
- Marie and her group, thanks for the help with the RNA work.
- a special thanks to Britta for preparing all the buffers, you save us a lot of time!!

Special thanks for all the nice people I met at the Department, especially:
- Alex, thanks for never refusing me a coffee at any time of the day, and for all these discussions where we never agreed, it would not have been so fun otherwise!
- Eduard, it’s good that you brought some Mediterranean touch in the lab (or Catalan touch I should say…)
- Rula, thanks for all the nice time we spent inside as well as outside the lab (merci à Eyad aussi donc…)
- Anthony thanks for your help and advices, and good luck with your complicated theories,
- Susannah and Haleh, as Alex said, thanks for the very nice discussions at lunches we had together (sometimes he can be right…).
- thanks to all the other very nice persons I met during this time: David, Linus, Poppy, Ignasi, Mats, and Johan among others.

More personally, I would like to thank my parents and my brother Christophe for their constant support.
I will always remember the persons that were always there for me here in Sweden:
- Mounia parents, Amine, Nadia and Samira, thanks a lot for introducing me to both Algerian and Swedish cultures, and thanks all of you for your great generosity.
- Hanna and Marek as well as Loïc, thanks for sharing a lot of excellent moments.

Finally, I dedicate this thesis for the two princesses of my life, my wife Mounia and my daughter Ines. Thanks Mounia for your never-ending support, and for your love.
REFERENCES


49


