FUNCTIONAL STUDIES OF THE UBIQUITIN-PROTEASOME SYSTEM USING GFP-BASED REPORTERS

Kristina Lindsten

Stockholm 2002
Front cover: Primary fibroblasts from a Ub^{576V}-GFP transgenic mouse treated with proteasome inhibitor \textit{in vitro}. (Photograph: K. Lindsten).

All previously published papers were reproduced by permission from the publisher

Published and printed by Karolinska University Press
Box 200, SE-171 77 Stockholm, Sweden
© Kristina Lindsten, 2002
ISBN 91-7349-398-8
To my family and Nico
FUNCTIONAL STUDIES OF THE UBIQUITIN-PROTEASOME SYSTEM USING GFP-BASED REPORTERS
Kristina Lindsten
Doctoral thesis from the Microbiology and Tumor Biology Center
Karolinska Institutet, Stockholm, Sweden

Ubiquitin-dependent proteasomal degradation is of paramount importance for cellular processes such as cell cycle progression, transcriptional regulation, apoptosis and disposal of misfolded and aberrant proteins. Moreover the ubiquitin-proteasome system is the main producer of peptides used by major histocompatibility complex class I for antigen presentation. Considering its critical involvement in these cellular processes, it is not surprising that aberrant ubiquitin-proteasome-dependent degradation is implicated in human disorders. The aim of this thesis has been to develop reporter systems for quantitative and functional analysis of this system in cells and to use these reporters to gain insight into the interactions between viral and cellular proteins and the ubiquitin-proteasome system.

The green fluorescent protein (GFP) was chosen as a reporter for degradation and was in the initial studies modified by the addition of degradation signals that transformed the GFP into a substrate of the proteasome. Upon inhibition of the system the GFP reporters accumulate and the level of obstruction can be monitored by the fluorescence. The GFP reporters were initially characterised and used for studies in cellular systems but more recently we developed a transgenic mouse model constitutively expressing one of the reporters.

We used the GFP reporters and new site-specific inhibitors to elucidate the individual contribution of the different proteasomal active sites. Our study demonstrated that the catalytic activities are of unequal importance for degradation. Furthermore this study indicated the presence of a non-catalytic modifier site that regulates the activity of the proteolytic sites through binding of peptides.

The role of a cellular and a viral repetitive sequence on proteolysis was also investigated using the GFP reporters. The Epstein-Barr virus derived Gly-Ala repeat functions as a stabilising domain. We concluded that it protects a protein from degradation in a length-dependent manner. Another repetitive sequence with a length-dependent effect is the poly-Gln repeat, which is expanded in several proteins involved in neurodegeneration. We showed that in contrast to the Gly-Ala repeat this repeat stabilises proteins through the formation of insoluble aggregates.

Another protein associated with neurodegeneration, and that also colocalises with different aggregates involved in such diseases, is the transcript mutant of ubiquitin, UBB\textsuperscript{15}. With the use of the GFP reporters we demonstrated that UBB\textsuperscript{15} is a substrate of the proteasome and is ubiquitinated in a specific manner and this causes an inhibition of the ubiquitin-proteasome system. It remains to be resolved if this inhibitory activity contributes to neuropathogenesis.

Together these studies shed some light on the roles of the ubiquitin-proteasome system in normal and pathologic conditions. The cellular and transgenic animal models presented in this thesis will be important tools for further studies on this intriguing proteolytic system.
LIST OF PUBLICATIONS

This thesis is based on the following papers and manuscript that will be referred to in the text by their Roman numerals.


II Lindsten K*, Menéndez-Benito V*, Masucci MG, Dantuma NP. A Transgenic Mouse Model for the Ubiquitin/Proteasome System. *Submitted*


*These authors contributed equally to the work
**ABBREVIATIONS**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>APC</td>
<td>anaphase promoting complex</td>
</tr>
<tr>
<td>ATP</td>
<td>adenosine triphosphate</td>
</tr>
<tr>
<td>CMV</td>
<td>cytomegalovirus</td>
</tr>
<tr>
<td>CP</td>
<td>core particle</td>
</tr>
<tr>
<td>CTL</td>
<td>cytotoxic T lymphocyte</td>
</tr>
<tr>
<td>E1</td>
<td>ubiquitin activating enzyme</td>
</tr>
<tr>
<td>E2</td>
<td>ubiquitin conjugating enzyme</td>
</tr>
<tr>
<td>E3</td>
<td>ubiquitin ligase</td>
</tr>
<tr>
<td>EBNA-1</td>
<td>EBV nuclear antigen 1</td>
</tr>
<tr>
<td>EBV</td>
<td>Epstein-Barr virus</td>
</tr>
<tr>
<td>ER</td>
<td>endoplasmatic reticulum</td>
</tr>
<tr>
<td>ERAD</td>
<td>ER-associated degradation</td>
</tr>
<tr>
<td>GFP</td>
<td>green fluorescent protein</td>
</tr>
<tr>
<td>GRR</td>
<td>glycine rich region</td>
</tr>
<tr>
<td>HBX</td>
<td>Hepatitis B virus X protein</td>
</tr>
<tr>
<td>HIF</td>
<td>hypoxia-inducible factor</td>
</tr>
<tr>
<td>HIV</td>
<td>human immunodeficiency virus</td>
</tr>
<tr>
<td>HPV</td>
<td>human papilloma virus</td>
</tr>
<tr>
<td>IFN</td>
<td>interferon</td>
</tr>
<tr>
<td>IxBα</td>
<td>inhibitor of NFκB</td>
</tr>
<tr>
<td>IKK</td>
<td>IκB kinase</td>
</tr>
<tr>
<td>KSHV</td>
<td>Kaposi sarcoma-associated herpes virus</td>
</tr>
<tr>
<td>MHC</td>
<td>major histocompatibility complex</td>
</tr>
<tr>
<td>NF-κB</td>
<td>nuclear factor-κB</td>
</tr>
<tr>
<td>RP</td>
<td>regulatory particle</td>
</tr>
<tr>
<td>SCA</td>
<td>spinocerebellar ataxia</td>
</tr>
<tr>
<td>SCF</td>
<td>Skp1/Cul1/F box protein</td>
</tr>
<tr>
<td>Ub</td>
<td>ubiquitin</td>
</tr>
<tr>
<td>UBA</td>
<td>ubiquitin-associated domain</td>
</tr>
<tr>
<td>UBB⁺¹</td>
<td>ubiquitin B gene with +1 frame shift</td>
</tr>
<tr>
<td>UBL</td>
<td>ubiquitin-like protein</td>
</tr>
<tr>
<td>UBP</td>
<td>ubiquitin-specific processing proteases</td>
</tr>
<tr>
<td>UCH</td>
<td>ubiquitin C-terminal hydrolase</td>
</tr>
<tr>
<td>UDP</td>
<td>ubiquitin-like domain protein</td>
</tr>
<tr>
<td>UFD</td>
<td>ubiquitin fusion degradation</td>
</tr>
<tr>
<td>UIM</td>
<td>ubiquitin interacting motif</td>
</tr>
<tr>
<td>VHL</td>
<td>von Hippel-Lindau</td>
</tr>
</tbody>
</table>
# LIST OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>1</td>
</tr>
<tr>
<td>LIST OF PUBLICATIONS</td>
<td></td>
</tr>
<tr>
<td>ABBREVIATIONS</td>
<td></td>
</tr>
<tr>
<td>1. AIMS OF THE STUDY</td>
<td>2</td>
</tr>
<tr>
<td>2. GENERAL INTRODUCTION TO THE UBIQUITIN-PROTEASOME SYSTEM</td>
<td></td>
</tr>
<tr>
<td>3. UBIQUITINATION: TARGETING PROTEINS FOR DEGRADATION</td>
<td>6</td>
</tr>
<tr>
<td>3.1 Ubiquitin</td>
<td>6</td>
</tr>
<tr>
<td>3.2 Ubiquitination</td>
<td>7</td>
</tr>
<tr>
<td>3.3 Functions of ubiquitination</td>
<td>9</td>
</tr>
<tr>
<td>3.4 Deubiquitination</td>
<td>11</td>
</tr>
<tr>
<td>3.5 Degradation signals</td>
<td>12</td>
</tr>
<tr>
<td>3.6 Stabilisation signals</td>
<td>15</td>
</tr>
<tr>
<td>4. PROTEASOMAL DEGRADATION: PROTEIN DESTRUCTION BY A COMPLEX MACHINERY</td>
<td>17</td>
</tr>
<tr>
<td>4.1 The 26S proteasome</td>
<td>17</td>
</tr>
<tr>
<td>4.2 The 19S regulatory particle</td>
<td>19</td>
</tr>
<tr>
<td>4.3 Other proteasomal complexes</td>
<td>20</td>
</tr>
<tr>
<td>4.4 Proteasome inhibitors</td>
<td>22</td>
</tr>
<tr>
<td>5. GFP-BASED REPORTERS FOR THE UBIQUITIN-PROTEASOME SYSTEM</td>
<td>25</td>
</tr>
<tr>
<td>5.1 Generation of GFP-based proteasome substrates</td>
<td>25</td>
</tr>
<tr>
<td>5.2 A mouse model for the ubiquitin-proteasome system</td>
<td>27</td>
</tr>
<tr>
<td>5.3 Importance of individual proteolytic sites</td>
<td>28</td>
</tr>
<tr>
<td>5.4 A modifier site regulates the proteasome</td>
<td>29</td>
</tr>
<tr>
<td>6. THE UBIQUITIN-PROTEASOME SYSTEM AND HUMAN DISORDERS</td>
<td>31</td>
</tr>
<tr>
<td>6.1 General overview</td>
<td>31</td>
</tr>
<tr>
<td>6.2 Viral immune escape and the ubiquitin-proteasome system</td>
<td>34</td>
</tr>
<tr>
<td>6.2.1 EBV and EBNA-1</td>
<td>35</td>
</tr>
<tr>
<td>6.2.2 The Gly-Ala repeat of EBNA-1: a stabilising signal</td>
<td>35</td>
</tr>
<tr>
<td>6.3 The ubiquitin-proteasome system in conformational diseases</td>
<td>39</td>
</tr>
<tr>
<td>6.3.1 Characteristics of poly-Gln proteins</td>
<td>40</td>
</tr>
<tr>
<td>6.3.2 Poly-Gln repeat expansions and proteasomal degradation</td>
<td>43</td>
</tr>
<tr>
<td>6.3.3 Aberrant ubiquitin, UBB$^+$, and conformational diseases</td>
<td>44</td>
</tr>
<tr>
<td>6.3.4 Effect of UBB$^+$ on proteasomal degradation</td>
<td>46</td>
</tr>
<tr>
<td>7. CONCLUSIONS AND FUTURE PERSPECTIVES</td>
<td>49</td>
</tr>
<tr>
<td>8. ACKNOWLEDGEMENTS</td>
<td>52</td>
</tr>
<tr>
<td>9. REFERENCES</td>
<td>54</td>
</tr>
<tr>
<td>APPENDIX (Papers I-VI)</td>
<td></td>
</tr>
</tbody>
</table>
I. AIMS OF THE STUDY

The general aim with the studies presented in this thesis was to generate tools for studying the activity and functionality of the ubiquitin-proteasome system *in vivo* and to address questions regarding protein stability, proteasomal function and aberrant ubiquitin-proteasome-dependent degradation in human disorders.

The specific aims were to:

- Generate green fluorescent protein (GFP)-based proteasomal substrates that would allow monitoring of the activity of the ubiquitin-proteasome system in cells as well as in a transgenic mouse model.

- Study the regulation of the proteolytic sites of the proteasome using new site-specific proteasome inhibitors.

- Analyse the effect of naturally occurring viral and cellular repetitive protein sequences on ubiquitin-proteasome-dependent proteolysis.

- Study the effect of an aberrant ubiquitin found in conformational diseases on the ubiquitin-proteasome system.
2. GENERAL INTRODUCTION TO THE UBIQUITIN-PROTEASOME SYSTEM

The ubiquitin-proteasome system was discovered more than 20 years ago but only recently has the field rapidly expanded and gained the scientific interest that reflects its paramount role in cell biology. The major task of the system is the regulated degradation of cytosolic and nuclear proteins that occurs within minutes or hours, a process that is vital for cell survival. In eukaryotes the ubiquitination system most often serves as a triggering signal for degradation. It accounts for targeting selected proteins for degradation by tagging them with chains of a small heat stable protein, the adenosine triphosphate (ATP)-dependent proteolytic factor 1 (APF-1), which was later identified to be identical to ubiquitin, a ubiquitously expressed protein with previously unknown function (275). This system is widely conserved through evolution and even though prokaryotes do not contain a ubiquitin homologue or orthologue some prokaryotes do contain proteasomes, however in a less complex form (53).

Hershko and Ciechanover found in the period from 1970 to 1980 that ubiquitin could be conjugated to Lys residues of proteins by an energy-dependent process requiring ATP. The ubiquitin modification of proteins was followed by rapid degradation by an energy requiring protease (97). Three groups of enzymes were required for accomplishing the ubiquitin conjugation: a ubiquitin activating enzyme (E1), the ubiquitin-conjugating enzymes (E2) and the ubiquitin ligases (E3). In a sequential manner these enzymes collaborate to covalently conjugate ubiquitin first to the substrates and next to one of the Lys within the ubiquitin until a chain of conjugated ubiquitins is formed, which is recognised by the proteasome resulting in degradation of the ubiquitinated protein (98) (Fig. 1).

In analogy with the lysosome, which degrades mostly extracellular proteins in a less selective way, the proteasome controls its proteolytic activity through self-compartmentalisation of the responsible proteases. The proteolytically active sites are facing the inside of a gated barrel-shaped ~2.5 MDa multisubunit complex, keeping cells from uncontrolled destruction of their protein content. The proteasome was first observed 1968 by electron microscopy but it took until 1988 before Goldberg and colleagues renamed the prosome, as it was originally known (217), to the proteasome, highlighting its pivotal role intracellular degradation (9).
The ubiquitin-proteasome system plays a key role in many different cellular events. A primary function of the system is the spatially and temporally controlled inactivation of key regulators of cellular processes through degradation. The progression of the cell cycle is an example of a process that requires strictly regulated turnover of proteins like p53 and cyclins. Under pathological conditions the turnover of such proteins can be dysregulated and lead to either uncontrolled proliferation or apoptosis. As a consequence of its role in cell cycle regulation the ubiquitin-proteasome system is directly involved in human disorders linked to uncontrolled growth such as cancer.

Another crucial function of proteasomal degradation is the clearance of misfolded proteins from the cellular environment. Especially in conditions that result in cellular stress the ubiquitin-proteasome system protects cells from toxic accumulation of aberrant proteins that are prone to form potentially harmful aggregates. Indeed impaired proteasome function has been implicated in several neurodegenerative disorders associated with accumulation of deposits of aggregated proteins. It is however not clear how, or if, the aggregated proteins contribute to the cellular pathology.

Progressive degradation of proteins by the proteasome results in the production of peptides of various lengths. These peptides can be further degraded by cytosolic peptidases generating amino acids that can be reused in protein synthesis. In vertebrates the peptides can also be translocated into the endoplasmatic reticulum (ER) and loaded on major histocompatibility complex (MHC) class I that are subsequently transported to and displayed at the cell surface, a process known as antigen presentation (211). Since the proteasome degrades the majority of intracellular proteins the presented peptides are representative for the total pool of proteins present within the cell. This pool also includes proteins that are derived from infectious agents and are consequently rapidly eliminated by cytotoxic T cells (CTLs). Antigen presentation is one of the fundamental phenomena of the immune system and is of major importance for the host response to viral infections. As a consequence of the host defence viruses have evolved sophisticated strategies to interfere at different steps of the antigen presentation pathway in an often-successful attempt to escape from recognition (51, 151, 197).

Although the proteasome generally degrades the protein into small peptides a few exceptions are known where the proteasome processes only a fraction of the
protein resulting in activation of the processed precursor that escapes from degradation. This mechanism is important for activation of precursor proteins of the mammalian transcription factor nuclear factor-κB (NF-κB) (117) and its distant yeast homologues Spt23 and Mga2 (201).

Beyond the functions of degradation there are also non-proteolytic functions associated with this system. There are reports suggesting that a subcomplex of the proteasome is recruited to some promoter regions and plays a role in transcriptional elongation (68, 84) or as a component in nuclear excision repair (78). Specific ubiquitin trees also play an important role in non-proteolytic mechanisms such as ribosomal function and viral budding or additional functions in DNA repair (81).
Figure 1. Overview of the ubiquitin-proteasome system. 1) Ubiquitin (Ub) is activated by a thiol ester linkage to the E1 enzyme. 2) Ub is then transferred to an E2 enzyme. 3) The E3 enzyme recognises and binds to a substrate. A RING finger E3 directly transfers the Ub from the E2 to a Lys residue in the substrate. In the case of a HECT domain E3 the Ub is first transferred to the E3 itself before conjugation to the substrate. The ubiquitination process is repeated and additional Ubs are covalently linked to the first Ub, forming a Ub-chain. 4) In some cases a U-box domain E4 enzyme is involved in elongation of Ub-chains. 5) At the proteasome the ubiquitinated substrate is deubiquitinated and unfolded. 6) The unfolded substrate is translocated through the proteasome and degraded into peptides. 7) The peptides can be used for antigen presentation or be further processed into amino acids by aminopeptidases.
3. UBIQUITINATION: TARGETING PROTEINS FOR DEGRADATION

3.1. Ubiquitin

Ubiquitin is encoded by a multigene family where monomeric ubiquitin genes are C-terminally linked to a sequence encoding a ribosomal protein, and multimeric ubiquitin genes that are encoded in a tandemly repetitive manner. The number of ubiquitin moieties in the precursors and the nucleotide sequence can vary substantially among different eukaryotic species. However the amino acid sequence shows only variation at three positions in the 76 residues long protein, making ubiquitin one of the most conserved eukaryotic proteins (112). Yeast and human ubiquitin differ at position 19, 24 and 28 (85). The newly synthesised ubiquitin fusion proteins or multimeric ubiquitin precursor are post-translationally processed by ubiquitin C-terminal hydrolases (UCH) that recognise and cleave after the last Gly residue of each ubiquitin moiety, releasing the functional single ubiquitin units (Fig. 2). The high expression levels and stability of ubiquitin combined with the fact that ubiquitin is continuously recycled after tagging proteins safeguards the high nuclear and cytosolic expression levels that are required for a functional ubiquitin-proteasome system.

Figure 2. Generation of single ubiquitin by deubiquitination enzymes. Ubiquitin (Ub) precursors are expressed as multimers or in fusion with ribosomal subunits, these are processed by ubiquitin C-terminal hydrolases (UCH) releasing functional ubiquitin monomers. There are also several ubiquitin-specific processing proteases (UBP) that can disassemble the free or attached ubiquitin chains.
There are two groups of proteins that share sequence homology with ubiquitin: 1) the ubiquitin-like proteins (UBL) and 2) proteins containing a ubiquitin-like domain (UDP) (111). UBLs, like SUMO or Nedd8, can be conjugated to other proteins in a mode resembling ubiquitin modifications. Yet they do not target for degradation but fulfil roles in for example inhibition of ubiquitination, regulation of protein trafficking and activation (115, 171). The UDPs comprise a diverse family of proteins that have an uncleavable ubiquitin-like domain most often located in their N-terminal part. It has been shown that the ubiquitin-like domain can interact with the proteasome but it is not believed that this interaction results in degradation of the UDPs.

3.2 Ubiquitination

Ubiquitination of a substrate is performed in a series of enzymatic steps that start with the ATP-dependent activation of ubiquitin by the formation of a thiol ester linkage between its C-terminal Gly and an active site Cys residue present within the ubiquitin-activating enzyme (E1). The ubiquitin is then transferred to a Cys residue of a ubiquitin-conjugating enzyme (E2) enzyme and finally covalently linked by an isopeptide bond to the e-NH$_2$ group of a Lys residue in the substrate, a process that is directed and catalysed by a substrate-specific ubiquitin ligase (E3). Once the first ubiquitin is linked to the substrate the procedure is repeated and additional ubiquitins are covalently linked by their C-terminal Gly76 to a specific Lys residues within the previous ubiquitin eventually forming a polyubiquitin chain. Long chains have higher affinity for the binding to the proteasome, and a minimum of four ubiquitins bind efficiently enough to serve as a recognition signal for degradation (249).

There is only a single E1 enzyme described and deletion of this gene in yeast is lethal (162). Mammalian cell lines expressing a temperature sensitive E1 tend to arrest in the G2 phase of the cell cycle and upregulate heat-shock proteins indicating a strong stress response at non-permissive temperatures (40, 70). The precise number of different E2 enzymes is not known but estimations based on sequence homology suggest that over twenty E2 enzymes are encoded in the mammalian genome (113). The E2s are often rather small proteins identified by a characteristic core of approximately 15 kDa that contains the active site Cys residue. Each E2 can interact with a number of different E3s, which comprises an even larger family of enzymes. The E3s, that can be single proteins or consist of multi-protein complexes, are responsible for the specificity of the ubiquitin-proteasome system since they recognise
the proteasome substrates through the presence of degradation signals (see chapter 3.5). There are multiple interactions between these enzymes: a single substrate can be ubiquitinated by different E2s and E3s and a single E3 can ubiquitinate different substrates and interact with different E2s.

The E3s can be divided into three major groups, the HECT (Homologues to E6-AP C-Terminus) domain family, the RING (Really Interesting New Gene) finger domain family and the recently identified U-box ligases (95, 193). There are some fundamental differences in the enzymatic activities between these E3 families. The E3s belonging to the HECT domain family first transfer the ubiquitin from the E2 to an internal Cys residue prior to the final conjugation to the substrate. On the contrary the RING finger domain E3s direct the specificity of the ubiquitination without forming an intermediate covalent linkage with the ubiquitin (113). Moreover RING finger, but not HECT domain, E3s are prone to auto-ubiquitination at an internal Lys residue, which is likely to have a regulatory function in some cases (193). Some of the RING finger E3s form multimeric protein complexes such as the Skp1/Cul1/F box protein (SCF) and the anaphase promoting complex (APC). The substrate specificity of the SCF is determined by F box proteins, an adaptor protein that recognises its substrate solely if the substrate is phosphorylated. On the other hand, the activity of the APC complex is regulated by phosphorylation of one of its own subunits. Hence temporal substrate degradation can be regulated by phosphorylation of either the E3 or the substrates. The SCF and APC ensure the timely degradation of many cyclins and other regulators of the cell cycle (183, 184).

The recently characterised U-box domain E3s contain of a conserved ~70 amino acid long stretch identified in at least six mammalian proteins (95). The domain is required for the E3 activity and share also some structural similarities to the E3 RING finger domains (6). The prototype U-box protein is the yeast Ufd2 protein, which contributes to ubiquitination of the ubiquitin fusion degradation (UFD) substrates (126). CHIP (C-terminus of the Hsc70 interacting protein) is also a member of this group of E3s and together with the E3 parkin it promotes ubiquitination of the unfolded Pael receptor involved in Parkinson’s disease (106). It should be noted that both Ufd2 and CHIP ubiquitinate substrates in concert with classic E3s and mainly promote elongation instead of initiation of the ubiquitin chain. Based on this unique feature of U-box proteins it has been proposed that they present a new class of
ubiquitination enzymes, the E4s. See Figure 1 for an overview of the ubiquitination process.

There are a few remarkable exceptions to the rule that proteasomal degradation requires ubiquitination of substrates. Ornithine decarboxylase, an enzyme involved in polyamine biosynthesis (172), and p21 a cyclin dependent kinase (228) are unstable proteins that can be degraded by the proteasome in a ubiquitin-independent manner. Ornithine decarboxylase is directed to the proteasome by a specific binding partner, antizyme, and subsequently degraded. Even though it has been shown with p21 constructs lacking Lys residues that it can be degraded in a ubiquitin-independent manner, its degradation in vivo can also involve polyubiquitination considering the interaction with the E3 complex SCF (283). However, p21 can also interact with an α-subunit of the proteasome and thereby mediate its degradation (251). Moreover a poor tertiary structure, as in the case of p21, or unfolded protein conformations, are known to preferentially bind the 19S regulatory complex of the proteasome (241) and it is possible that these proteins can bypass certain unfolding processes making ubiquitin-independent degradation more likely. This direct way of the proteasome to recognise non-ubiquitinated substrates could be more common than first thought as suggested by studies where various denatured proteins were shown to be degraded by the proteasome without prior ubiquitination (19).

3.3 Functions of ubiquitination

The primary function of ubiquitination is to serve as a reusable selective tag sorting out those proteins that are destined for proteasomal degradation, yet of crucial importance this is not its only function. Due to different modes of ubiquitination such diverse functions as DNA repair, protein activation and endocytosis can be accomplished. The different ubiquitination signals are a consequence of the various possibilities to build a Lys-linked chain (194). Ubiquitin carries seven Lys residues at position 6, 11, 27, 29, 33, 48 and 63 of which Lys29, Lys48 and Lys63 linked chains have been identified in vivo and are illustrated in Figure 3 (8, 35, 86, 236). The Lys48-linked chains, in which an isopeptide bond links the C-terminal Gly76 residues of ubiquitin and Lys48 of the previous ubiquitin, are the canonical signal for proteolysis. Overexpression of ubiquitin in which the Lys48 is substituted Arg is lethal in yeast (71). Though less common it has been shown that also Lys29-linked
ubiquitin chains can target a protein for degradation. Hitherto involvement of this type of polyubiquitin trees has only been shown for degradation of UFD substrates in which the N-terminally uncleavable ubiquitin moiety serves as an anchorage for both Lys29 and Lys48 chains leading to rapid degradation. We showed for the first time that both the Lys29 as well as the Lys48 ubiquitin trees can independently target a UFD substrate for degradation in mammalian cells (paper VI).  

Figure 3. Functions of different ubiquitin-chains. Ubiquitin contains seven Lys residues (K) and five of these, K6, K11, K29, K48, K63, can form ubiquitin-chains. No functions of the K6 and K11 trees have been described. To date K29 trees have only been demonstrated for degradation of UFD substrates. The canonical ubiquitin chain is K48 linked and targets the bulk of substrates for proteasomal degradation. K63 trees have different non-proteolytic functions. Mono-ubiquitination also serves non-proteolytic functions and is involved in many different biological processes.

Another well-characterised ubiquitin chain is linked through Lys63 and it has been shown that these trees can serve as a signal for a variety of events, however they do not play a role in targeting for proteasomal degradation. Lys63-linked polyubiquitin trees have been identified for example conjugated to the ribosomal subunit L28 mainly during S-phase when it enhances translation (235) and after proinflammatory stimuli on the IkB kinase (IKK) complex where it has an activating effect (55). Lys63 trees are also linked to post replicative DNA repair functions and in agreement with this notion cells expressing the Ub63R mutant show increased UV sensitivity (101, 236). Furthermore Lys63-linked ubiquitin chains have been implicated in endocytosis of cell surface receptors (99, 242) and in mitochondrial
DNA inheritance (72). How this chain exerts these functions is not known, possibly a unique structure of the polyubiquitin tree directs specific protein interactions. Finally, ubiquitin chains linked through Lys6 or Lys11 can also be generated but so far no biological functions of such chains have been reported. Interestingly it was shown that these Lys6 and Lys11-linked chains can bind the to the S5a subunit of the 19S regulatory complex of the proteasome in vitro with comparable affinity as the classical Lys48-linked chains (13).

In addition proteins can also be monoubiquitination or sometimes multimonoubiquitinated at several Lys residues. It is of note that monoubiquitin modifications are not efficiently recognised by the proteasome as compared to a chain of minimum four ubiquitins (249). Yet two other motifs have been described; the ubiquitin interacting motif (UIM) and ubiquitin associated domain (UBA) that do interact with ubiquitin monomers (29). It has been proposed that a UIM-ubiquitin-based intracellular network exist, that is important for trafficking of membrane proteins (198). In agreement with the involvement of monoubiquitination in the endocytic machinery it also plays an important role for proper viral budding (100). While the abundant presence of monoubiquitinated histones in nuclei was acknowledged more than 20 years ago (30), it was only recently found to be of importance for yeast meiosis (210) as well as development of Drosophila embryos (192)

3.4 Deubiquitination

Ubiquitination is a reversible protein modification since the conjugated polyubiquitin tree can be removed and disassembled by the action of deubiquitinating enzymes (DUBs). The DUBs comprise a large family of enzymes within the ubiquitin-proteasome system. All DUBs are cysteine proteases with the exception of one recently discovered DUB within the proteasome that is a metalloprotease (267, 279). DUBs can be divided into two groups, UCHs and ubiquitin-specific processing proteases (UBP). These two classes have some common residues in their active sites but besides that there is poor sequence homology between UCHs, UBPs and other cysteine proteases. UCHs are typically small proteins that remove peptides fused to the C-terminus of ubiquitin while UBPs make up a more diverse family of proteins that cleave isopeptide bonds between ubiquitin and a Lys residues within a substrate or between ubiquitins in a polyubiquitin chain.

11
DUBs play a vital role in maintaining a sufficient pool of free ubiquitin in the cell by processing the ubiquitin monomers from their precursors and by recycling polyubiquitin trees from substrates targeted for degradation (38, 274). Figure 2 illustrates how the pool of functional ubiquitin monomers is maintained by the activity of DUBs.

It is also believed that many of the DUBs play an important role in stabilisation of substrates by removing polyubiquitin trees in a regulated fashion. However the number of identified substrates of DUBs is very limited. The tumor suppressor protein p53 can be regulated by HAUSP (Herpes virus-Associated Ubiquitin Specific Protease), also known as human USP7, which rescues p53 from degradation through specific deubiquitination (147). Another example is the human homologue of the *Drosophila* deubiquitinating enzyme fat facet, named Fam, that can regulate both β-catenin (246) and the Ras target AF-P (247).

Most DUBs are soluble cytoplasmic or nuclear proteins but some are found in association with the 19S regulatory particle of the proteasome: the Rpn11/S13, Dna4/Ubp4, UCH37/Uch2, Ubp6/Usp14 and Ubp14/IsoT (134, 138, 148, 187, 267, 279). Most of these, except Rpn11, are probably only transiently associated with the proteasome as indicated by their non-stoichiometric interactions (187). UCH37 trims the polyubiquitin tree from its distal end and is believed to play an important role as a molecular clock that determines the residence time of an interacting protein at the proteasome. If the polyubiquitin tree is long enough the protein will stay associated for sufficient time in order to be unfolded and translocated into the proteasome (134). The deubiquitination activity of the proteasomal subunit Rpn11, is essential in yeast illustrating the importance of these regulatory function in cells (267, 279). It is tempting to speculate that Rpn11 complements the molecular clock function of UCH37 by its capability to cut off the remnant of the chain just prior to complementation of translocation of the substrate into the proteasome.

3.5 Degradation signals

The levels of many proteins must be highly regulated both spatially and temporally in order for cellular functions to proceed accurately. The steady state levels of proteins are a result of the rate of synthesis and the rate of breakdown. The turnover of proteins is generally determined by the presence of degradation signals, also known as degrons, which are domains or small motifs recognised by E3 enzymes leading to
ubiquitination and proteasomal degradation of the protein (136). Many different signals have been identified that confer instability to a protein: the N-end rule, UFD signal, PEST sequence, destruction box, protein misfolding and N-glycosylation are some examples of structures targeted by E3s.

The N-end rule is a universal degradation signal that is mediated in yeast (14) and mammalian cells (145) by the ubiquitin-proteasome system and in bacteria by an unrelated complex protease (250). It has been thoroughly studied and it refers to the principle that protein stability is a consequence of the identity of the N-terminal amino acid. An N-terminal Met is not recognised by the N-end rule pathway whereas proteins with basic (Arg, Lys, His) or bulky hydrophobic amino acids (Phe, Leu, Trp, Tyr, Ile) are generally recognised and ubiquitinated by an E3 enzyme called Ubr1 in yeast and E3α in mammals (265). It is noteworthy that the N-end rule requires in addition to a destabilising N-terminal residue also a Lys in close proximity to the N-terminus, which serves as the ubiquitin conjugation site (14, 243). The Ubr1 has in addition a separate binding capacity independent of the N-terminal residue of the substrate and can by this manner regulate the expression of a peptide transporter in yeast that controls peptide import (253). N-end rule degradation signals are responsible for degrading a number of cellular and viral proteins and one bacterial protein (265) and was recently suggested to have a function in cardiovascular development (130). Besides that, the functionality of the N-end rule pathway is crucial for maintaining chromosome stability. Sister chromatids are held together by a complex named cohesin until one of its subunits is cleaved by separin, an event that triggers anaphase and safeguards proper chromatide separation (177). It has been shown that the C-terminal fragment of the cleaved cohesin has a destabilising amino terminus and needs to be removed by the N-end rule pathway (199).

The UFD signal is based on the concept that if a ubiquitin-fusion protein can escape cleavage by UCHs, the ubiquitin will serve as a site for further ubiquitination followed by rapid degradation of the fusion protein (114). Substitution of the last Gly residue in the ubiquitin, within an artificial ubiquitin-fusion protein, blocks cleavage by UCHs and transforms the fusion in a UFD substrate. Like the N-end rule this degradation signal is widely used as a tool in molecular biology for targeting a protein of interest for degradation. The first and hitherto only naturally occurring UFD substrate is identified in paper VI as being the aberrant ubiquitin UBB™ that is associated with neurodegeneration. However there are indications that the
transcription factor MyoD (27), human papilloma virus oncoprotein E7 (204) and Epstein-Barr virus oncoprotein LMP-1 (12) are all ubiquitinated at the free N-terminal amino group instead of a Lys inside the proteins. It is tempting to speculate that upon ligation of the first N-terminal ubiquitin moiety the substrates are subsequently being recognised and processed identical to the artificial UFD substrates.

PEST sequences are another type of degradation signal, which are characterised by a high content of Pro, Glu, Asp, Ser, and Thr residues. Many well-known proteasome substrates contain such sequences like the tumor suppressor p53, the inhibitor of NF-κB, IκBα, the yeast transcription factor Gen4 and many cyclins (212). Degradation of IκBα is not solely dependent on the PEST sequence as its degradation can also be separately stimulated by TNFα induced phosphorylation. After TNFα stimulation a signalling cascade takes place that can result in stimulation of IKK by Lys63-linked ubiquitination (55). This is followed by phosphorylation of IκBα on two specific Ser residues that induce ubiquitination by the E3 SCF and subsequent degradation. Hence, IκBα is subject to both constitutive (PEST) and inducible (phosphorylation) proteasomal degradation (203).

The destruction boxes consist of a nine amino acids long consensus sequence found in many mitotic cyclins which presence is tightly regulated during the cell cycle. This motif is recognised by the E3 complex APC that in addition is regulated by phosphorylation before it can recognise its targets. The APC is also involved in timely regulated ubiquitination of many other cell cycle proteins (191).

Another signal for ubiquitin-dependent degradation comes from protein misfolding. Accumulation of proteins that fail to fold properly is often a sign of cellular stress induced by for example elevated temperatures, amino acid homologues or exposure to oxygen radicals. These hazardous conditions are counteracted by the cell with upregulation of chaperones and components of the ubiquitin-proteasome system (229). Recently a collaborative system involved in recognition and delivery of misfolded proteins to the proteasome has been detected. Here co-chaperones like BAG-1 or CHIP provide a link to degradation by their interaction with chaperones such as Hsp-70 that are specialised in recognising misfolded proteins (44, 164, 234). BAG-1 can with the help of its UBL domain deliver Hsp-70-bound proteins to the proteasome (152). CHIP on the other hand is an E3 enzyme, that by interacting with
Hsp-70 can specifically ubiquitinate the misfolded protein and promote its degradation (174).

The protection from aberrant folding by proteasomal degradation also accounts for proteins residing in the ER. Chaperones like Bip will promote retrograde transport of aberrant proteins through the Sec61 translocon back into the cytosol where ubiquitination and degradation takes place. This process of protein quality control is referred to as ER associated degradation (ERAD) (196). Control of ER proteins translocated to the cytosol are also performed by specific recognition of cytosolic proteins that are N-glycosylated. For these proteins the N-glycan serves as a degradation signal that is specifically recognised by the SCF<sub>Brc2</sub> E3 complex (281). Closely related to proteasomal degradation of misfolded proteins is the degradation of newly translated proteins that do not attain native structures due to improper posttranslational processes or errors in translation. These peptides were named DRiPs for Defective Ribosomal Products and make up as much as 30% of the newly synthesized proteins (219), and form a primary source for antigenic peptides (206).

### 3.6 Stabilisation signals

It has been postulated that some proteins carry also stabilisation signals that counter the targeting for destruction by degradation signals (51). This could add an extra dimension to the regulation of protein stability. Another reason for the existence of such signals could be that a stable protein that is targeted for degradation by ubiquitination exerts a specific function in its ubiquitinated state. One function could be to recruit either substrates or accessory proteins to the proteasome or to facilitate non-proteolytic functions of the proteasome, as has been proposed for Rad23. The stabilisation hypothesis was made based on the observation that the Epstein-Barr virus (EBV) nuclear antigen (EBNA)-1 resists proteasomal degradation due to the presence of an internal Gly-Ala repeat (51). The Gly-Ala repeat provides a stabilising domain which property remains when it is transferred to another destabilised protein (143). The exact underlying mechanism of this phenomenon is not yet known but is discussed in further detail in chapter 6.2.2 and paper IV.

There are hints that other low complexity sequences could have functions similar to the Gly-Ala repeat. The NF-κB precursor p105 contains a Gly-rich region (GRR) that share some similarities to the Gly-Ala repeat, but so far this sequence does not behave as a transferable stabilising element (181). The p105 is partially processed
from its C-terminus by the proteasome but stops at the GRR releasing the active p50 subunit (41). Similarly a low complexity Asn-rich sequence in the ER-bound yeast transcription factors Spt23 and Mga2 mediates partial digestion of the proteasome releasing the transcription factor from the membrane. Interestingly Spt23 forms a homodimer in the ER membrane that leads to monoubiquitination (200). After proteolysis the generated p90 retains the ubiquitin moiety and is translocated to nucleus where it functions as a transcription factor that controls membrane fluidity by regulating the levels of unsaturated fatty acids (102).

In independent studies in vitro degradation of a substrate could be prevented by the presence of a tightly folded domain indicating that there might exist conformational stabilisation signals. A large biotin-avidin complex linked to the C-terminus of a protein prevented its degradation in contrast to when it was placed in the N-terminus (178). Another example shows that unfolding by the proteasome is influenced by structural domains and that their localisation within the substrate can influence protein stability (137). Some studies have reported that the aggregation-prone expanded poly-Gln repeats characteristic for many neurodegenerative disorders (286) delay proteasomal degradation (47, 110). It has been postulated that their resistance may be related to their aggregation in analogy with the conformational stabilisation (51). In paper V we present evidence in favour of this hypothesis showing that the expanded poly-Gln repeats protect proteins from being degraded through the formation of insoluble aggregates, while the free pool of poly-Gln proteins is efficiently degraded. This might be explained by the tight structures in the aggregates that could hinder unfolding by the proteasome.
4. PROTEASOMAL DEGRADATION: PROTEIN DESTRUCTION BY A COMPLEX MACHINERY

4.1 The 26S proteasome

Proteasomes are very abundant in the cytosolic and nuclear compartments as they account for up to 1% of the cellular proteins (17). It has been shown that proteasomes can diffuse freely in the cytosol and the nucleus (205). Proteasomes enter the nucleus through the nuclear pore complex or they get entrapped within the nuclear compartment when the envelope is restored after cell division. In yeast proteasomes have also been detected associated with the nuclear and ER membranes and to have an altered distribution during mitosis (64).

The mammalian proteasome forms a barrel shaped complex consisting of the 20S core particle (CP) with the proteolytic activities facing inwards to the lumen of the proteasome. The CP is located in between two 19S regulatory particles (RP), also called PA700, that serve as proteasomal gatekeepers and control entry or exit of proteins and peptides (Fig. 4A). The CP consists of four heptameric rings, named α and β, forming the structure: α2β7-β7α2. The two inner rings that consist of β subunits contain three proteolytic activities each. The β1 subunit displays post-glutamyl peptide hydrolysing (PGPH) activity, cleaving preferably after acidic amino acids. The β2 subunit has a trypsin-like activity, cleaving preferably after basic amino acids. Finally the β5 subunit is responsible for the chymotrypsin-like activity, which cleaves preferably after hydrophobic amino acids (7). All three catalytic sites are activated through cleavage of an N-terminal leader sequence that occurs during the assembly of the proteasomal complex. The cleavage results in N-terminal Thr residues that exert the nucleophile attacks on peptide bonds (37).

In addition, two less characterised proteolytic activities have been described, namely the “branched-chain amino acid preferring” (BrAAP) activity and the “small neutral amino acid preferring” (SNAAP) activity (182). These additional activities are most likely accommodated by the already identified active β sites since the presence of additional proteases in the proteasome complex have been excluded by structural analysis (89), kinetic studies (31, 161) and site-directed mutagenesis (57). In the end the cleavage repertoire makes sure that virtually any peptide bond in a protein can be
potentially digested, thus the active sites of the proteasome show much broader specificity than their names reveal.

The generated peptides vary in size between 3-22 amino acids with a median of 6 (123). They are short lived in the cytosol and nucleus as they are subject to further degraded by other proteases and aminopeptidases (231, 244). A minority of peptides will also be translocated by the transporters for antigen presentation (TAP) into the ER, where they can be loaded on MHC class I molecules and further transported to the cell surface and displayed to the immune system (168).

Figure 4. The mammalian proteosome complexes and their proteolytic activities. A) The 26S proteosome consists of the 20S core particle (CP) and the 19S regulatory particle (RP). The CP is built up of two α rings and two β rings each containing 7 different subunits (14 different in total). The RP docks to one or both ends of the CP and can be divided in the base (9 subunits) and the lid (8 subunits). The proteolytic active sites are located in the β₁, β₂ and β₅ subunits facing the inside of the CP. These different activities are referred to the as post-glutamyl peptide hydrolyzing (PGPH), trypsin-like and chymotrypsin-like. B) The immunoproteasome induced by IFNγ has exchanged the β₁, β₂ and β₅ subunits to iβ₁/LMP2, iβ₂/MECL1 and iβ₅/LMP7 altering the cleavage specificities. There are also alternative activating complexes that can replace the RP. Here is the 11S or PA28 complex illustrated that consists of a hexameric ring of PA28 α and β subunits.
4.2 The 19S regulatory particle

The RP regulates the entry of substrates targeted for degradation. It docks to the CP and regulates the opening of a narrow hole of 10 Å in diameter, a size that allows only unfolded proteins to enter (266). The α rings form the antechamber leading into the CP, however without association to a RP the N-termini of the α-subunits serves as a plug held together by the α3 subunit keeping the CP closed (88). Electron microscopy studies have revealed that the RP is flexibly linked to the CP and probably continuously, in an uncorrelated manner, exerts wagging movements (254). If this is mechanistically important for proteasomal functions is not known.

The RP plays an important role in recruitment of the polyubiquitinated substrates, removal of the polyubiquitin tree, unfolding of the protein and tethering of the protein into the core chamber. The RP can be subdivided into two subcomplexes known as the base and the lid. The lid contains eight subunits, namely Rpn3/S3, Rpn5, Rpn6/S9, Rpn7/S10a, Rpn8/S12, Rpn9/S11, Rpn11/S13 and Rpn12/S14. (yeast/mammalian nomenclature). The precise role of the lid has been unclear but it is required for proper protein degradation and is important for substrate recognition (82).

Recent studies also revealed deubiquitination activity in the Rpn11/S12 subunit suggesting that it is important for removal of the ubiquitin tree prior to degradation (267, 279). The base contains nine subunits, the Rpt1/S7, Rpt2/S4, Rpt3/S6, Rpt4/S10b, Rpt5/S6’, Rpt6/S8, Rpn1/S2, Rpn2/S1 and Rpn10/S5a (yeast/mammalian nomenclature). Rpt1-6 are AAA (ATPases Associated with a variety of cellular Activities) ATPases that form a ring at the entrance. AAA ATPases are known to be involved in protein remodelling and in the context of the proteasome it has been postulated that these subunits are in charge of unfolding and translocation of the substrate into the proteolytically active chamber (26, 131). Interestingly it has been recently shown that one of the AAA ATPases, S6’, contain a polyubiquitin binding site that may be responsible for positioning the ubiquitinated substrate directly adjacent to the entrance (132). The other two base subunits, Rpn1 and 2, are the largest subunits of the RP and expose hydrophobic surfaces possibly functioning as a scaffold for interacting proteins. Recently Rpn1 was shown to interact with UDP proteins such as Rad23 and Dsk2 in yeast (63). The S5a subunit, that connects the base and the lid in mammalian cells, contains two UIM sequences that recognise polyubiquitin trees of at least four ubiquitins (282). S5a is the only RP subunit that is
largely present in the cell as a free protein with only a minor fraction incorporated in the proteasome (260). How this reflects its function in the RP is not well understood but it is not ruled out that the ubiquitin binding function is largely performed in its soluble state. It is noteworthy that also Rad23, that transiently associates with the RP but is mostly present as a free protein, contains two ubiquitin binding sites. It has been proposed that both Rpn10 and Rad23 are important for recruitment of ubiquitinated proteins to the proteasome (36, 135, 282). Deletion of Rpn10 causes only stabilisation of UFD substrates (282), and deletion of Rad23 results in UV sensitivity (90, 169) reflecting the pivotal role of Rad23 in nuclear excision repair. However the double Rpn10 and Rad23 mutant displays a severely affected ubiquitin-proteasome system, suggesting overlapping functions of these two proteins (135).

In addition the proteasome is also involved in processes that do not require proteolytic activities. The ATPase subunits in the RP are essential for proper nucleotide excision repair independent on proteolysis. For this function the binding of Rad23 mediated through its UBL domain to the proteasome is a prerequisite (213). Furthermore transcriptional regulation has also been implicated to require non-proteolytic functions of the proteasome. There are proofs that the RP activates transcriptional elongation by RNA polymerase II (68) and that a subcomplex containing some of the ATPase subunits of the base is directly recruited to Gal promoters and functions as a transcriptional activator in yeast (84). The underlying mechanism for these functions is not well understood.

4.3 Other proteasomal complexes
The mammalian proteasome has a unique feature in that it has three exchangeable proteolytic β subunits that are induced in the presence of the cytokine IFNγ. IFNγ is produced by activated CD4 and CD8 positive T-cells and natural killer cells and plays a major role in mobilising the host defence against infectious pathogens. The β1, β2 and β3 subunits are exchanged to iβ1/LMP2, iβ2/MECL1 and iβ3/LMP7, respectively, resulting in the formation of the immunoproteasome complexes (75) (Fig. 4). These modifications cause an altered preference of cleavage sites favouring the generation of peptides with hydrophobic C-termini and reducing the output of peptides with acidic C-termini. A peptide with a hydrophobic C-terminus is more likely to be suitable for antigen presentation by MHC class I. The alteration of cleavage
specificity for the C-terminus is of most importance for the proteasome (34) since the N-terminus can later be further trimmed to fit the binding groove of MHC class I by the recently identified ER-resident protease ERAAP (**ER Aminopeptidase Associated with antigen Processing**) (224).

Another event that is observed upon IFNγ stimulation is the exchange of the RP for another regulatory structure called 11S or PA28 (146). It has been shown that different complexes with PA28, the RP and the 20S core can be formed: 20S-PA28, 19S-20S-PA28 or PA28-20S-PA28 (245). Unlike the RP the PA28 contains only two different types of subunits, PA28α and PA28β, forming a hetero-hexameric ring structure that upon binding to the CP is believed to open the entrance hole to the inner proteolytic chamber as observed with an archeabacterial homologous activator (270). Both these subunits lack ATPase activities and ubiquitin binding sites and therefore it is likely that mainly already unfolded proteins or peptides will be allowed to enter (240). In addition PA28α/β stimulates the proteolytic activities and change the spectrum of peptides that are generated, even though the overall rate of protein breakdown remains and the peptides are similar in size to those generated from the 26S (33, 175). Besides that its expression is induced by IFNγ and the fact that this adaptor is only found in animals with an adaptive immune response also hints in the direct of an important role in antigen presentation. It has indeed been shown that PA28α/β can optimise presentation of some epitopes (87). A third PA28 related subunit is PA28γ that is constitutively expressed and forms a homohexameric ring that docks, similar to the PA28α/β, to the CP. In contrast to the PA28α/β ring that is predominantly located in the cytoplasm, the PA28γ is mainly nuclear (273). The function of PA28γ remains elusive but it is unlikely to play a role in antigen presentation considering the presence of functional homologues in insects (156). The only change found in PA28γ knockout mice was a slightly retarded growth and a delay in the cell cycle of primary cells in culture (173). It has also been shown that the *Drosophila* homologue of this regulator can modify the proteolytic activities of the proteasome (156).
4.4 Proteasome inhibitors

To date a number of synthetic and naturally occurring compounds have been identified that can inhibit or modulate proteasome activity. Most of these inhibitors are short, linear molecules in which a peptide backbone is linked to a pharmacophore resulting in a “suicide substrate”. The peptide fraction directs the inhibitor to the proteasome’s active site in the substrate-binding pocket where the pharmacophore exert the inhibitory effect by reversibly or irreversibly forming covalent adducts (176). Based on the pharmacophore most proteasome inhibitors can be divided into different groups as listed below in Table 1. The unique Thr-based proteolytic site of the proteasome subunits explains the high specificity of many of those inhibitors. Yet, most known natural and synthetic proteasome inhibitors have a predominant effect on the chymotrypsin-like activity, which is most likely due to two reasons: 1) hydrophobic inhibitors that block the chymotrypsin-like activity are more cell-permeable than those specific to the other activities, which often contain charged residues, and 2) inhibition of the chymotrypsin-like activity has the most dramatic effect on protein degradation.

There are several reasons that justify a major effort in the development of new and more specific proteasome inhibitors. First of all they can be used to study in greater detail the roles of the different proteolytic activities of the proteasome in biological processes. Detailed understanding of their roles in such events is of major importance for successful therapeutic interference with the ubiquitin-proteasome system. Secondly there is great therapeutic potential in the usage of proteasome inhibitors as drugs in treatment against cancer and immune and inflammatory responses.

Accelerated proteasomal degradation of tumor suppressors like p53 (214) and p27 (233) plays a role in malignant transformation of cells. Hence inhibition of the ubiquitin-proteasome system is anticipated to restore these activities resulting in induction of cell cycle arrest and apoptosis of malignant cells and may hence represent a new type of anti-cancer drugs. Moreover proteasome inhibition prevents degradation of the NF-κB repressor IkBα, which is also desirable for an anti-tumor effect. The inhibition would avoid that NF-κB stimulates expression of inhibitors of apoptosis (18, 255, 262) and expression of cell surface molecules that can be involved in tumor metastasis and angiogenesis (202, 284). In line with this notion it has been
Table 1. List of proteasome inhibitors and their inhibitory capacity. The activities of some proteasome inhibitors on the chymotrypsin-like activity (CT-L), trypsin-like activity (Tr-L) or PGPH activity are indicated by + or -. The + indicates the strength of inhibition and – indicates lack of inhibition.

<table>
<thead>
<tr>
<th>GROUP</th>
<th>Name of inhibitor</th>
<th>Mode of inhibition</th>
<th>Inhibiting capacity CT-L, Tr-L, PGPH</th>
<th>Inhibition of other proteases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non peptide inhibitors</td>
<td>Latacystin(^1) (Streptomyces)</td>
<td>irreversible</td>
<td>CT-L Tr-L PGPH</td>
<td>++++</td>
</tr>
<tr>
<td>Peptide boronates</td>
<td>PS-341(^1)</td>
<td>reversible (slow on and off rates)</td>
<td>CT-L Tr-L PGPH</td>
<td>++++ -</td>
</tr>
<tr>
<td></td>
<td>MG-262(^2)</td>
<td>reversible (slow on and off rates)</td>
<td>CT-L Tr-L PGPH</td>
<td>+++ +</td>
</tr>
<tr>
<td>Peptide aldehydes</td>
<td>MG-132(^2)</td>
<td>reversible</td>
<td>CT-L Tr-L PGPH</td>
<td>++++ +</td>
</tr>
<tr>
<td>Peptide vinylsulfones</td>
<td>NLVS(^3)</td>
<td>irreversible</td>
<td>CT-L Tr-L PGPH</td>
<td>++++ +</td>
</tr>
<tr>
<td></td>
<td>ZLVS(^3)</td>
<td>irreversible</td>
<td>CT-L Tr-L PGPH</td>
<td>++++ ++</td>
</tr>
<tr>
<td>Peptide epoxyketones</td>
<td>Epoxomicin(^1) (Actinomycetes)</td>
<td>irreversible</td>
<td>CT-L Tr-L PGPH</td>
<td>++++ -</td>
</tr>
<tr>
<td></td>
<td>Dihydroepotheonycin(^1) (Streptomyces)</td>
<td>Irreversible</td>
<td>CT-L Tr-L PGPH</td>
<td>++ +</td>
</tr>
<tr>
<td></td>
<td>Ac-GPFL-EX, YU102(^2)</td>
<td>irreversible</td>
<td>CT-L Tr-L PGPH</td>
<td>- - +++</td>
</tr>
<tr>
<td></td>
<td>Ac-NorL.PFL-EX(^2)</td>
<td>irreversible</td>
<td>CT-L Tr-L PGPH</td>
<td>- - +++</td>
</tr>
<tr>
<td>HIV-1 protease inhibitors</td>
<td>Ritonavir(^4)</td>
<td>reversible</td>
<td>CT-L Tr-L PGPH</td>
<td>+++ stimulates HIV-1 protease</td>
</tr>
<tr>
<td>cross-reacting with the</td>
<td>Saquinavir(^4)</td>
<td>reversible</td>
<td>CT-L Tr-L PGPH</td>
<td>+++ ? HIV-1 protease</td>
</tr>
<tr>
<td>proteasome</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

References:
\(^1\) Kisselev AF. et al.(2001) Chem. Biol. 8:739-758
\(^2\) Myung J. et al. (2001) Mol Cell. 7:411-420
\(^3\) Schmidte G. et al. (1999) J. Biol. Chem. 274:35734-35740
shown that the proteasome inhibitors eponemycin (166), epoxomicin (92) and PS-341 exert anti-tumor activities in animal models. The latter is presently being evaluated in phase I and II clinical trials. These trials comprise several haematological malignancies and solid tumors and are showing promising results but are accompanied with some side effects among which one of the most severe is painful sensory neuropathy (2). Some of the therapeutic drugs directed to the human immunodeficiency virus (HIV)-1 protease that are used for the treatment of AIDS patients modulate in addition the activity of the proteasome. This activity has been linked to the often-observed regression of HIV-1 associated Kaposi’s sarcoma in patients treated with those drugs (188, 225).

By using inhibitors more specific to one site it has been shown that the cleavage pattern of proteasome substrates can be altered generating a different peptide repertoire, which ultimately can lead to modifications of antigen presentation (5, 223). This opens up the possibility to use proteasome inhibitors as immuno-modulators in order to either provoke an immune response in the case of cancer or virus-infected cells, or to suppress the immune response in the case of autoimmune disorders. Indeed animal models support that proteasome inhibitors can have a potential therapeutic effect against unwanted immune and inflammatory responses, most likely also due to the inhibition of NF-κB activation, as shown in models for psoriasis (287), asthma (61), rheumatoid arthritis (186) and autoimmune encephalomyelitis (261).
5. GFP-BASED REPORTERS FOR THE UBIQUITIN-PROTEASOME SYSTEM

5.1 Generation of GFP-based proteasome substrates
The work of this thesis was initiated by the idea of using degradation signals to target the GFP for destruction by the ubiquitin-proteasome system. We reasoned that such reporter proteins would allow us to study the activity of this system in vivo. Traditionally the effect of compounds on the ubiquitin-proteasome system is analysed by determining their effect on the proteolytic activity of purified proteasomes, which can be measured with small fluorogenic substrates. Alternatively the turnover of an endogenous substrate can be followed by metabolic pulse-chase labelling of cells followed with immunoprecipitation of the protein-of-interest or by in vitro reconstitution of the system with individual components. Each method has its limitations. While the assay with the fluorogenic substrates has been a valuable tool for the identification of proteasome inhibitors, other compounds or proteins that affect different steps of this complex pathway cannot be identified. Moreover, it is performed with isolated proteasomes out of their natural microenvironment.

Monitoring of endogenous substrates unmasks another fundamental problem in quantification of the ubiquitin-proteasome system. Many, if not all, natural substrates are subject to different regulatory steps and hence stabilisation of an endogenous substrate does not have to be due to blockage of the ubiquitin-proteasome system. For these reasons we decided to target the well-established GFP (252) for proteasomal degradation using two different constitutive degradation signals: the N-end rule and UFD signal (paper I; see also commentary (83)). GFP is a small soluble protein that originates from the jellyfish Aequorea victoria. It does not have any known interacting partners and is relatively well tolerated at high concentrations in mammalian cells. Most importantly GFP can be easily monitored by its autofluorescence. Cervix carcinoma HeLa cells expressing the GFP reporters responded to treatment with proteasome inhibitors by a 100-1000 fold increase in fluorescence which could be conveniently detected and quantified by flow cytometry, fluorimetry or fluorescence microscopy. In contrast to the fluorogenic substrate assays that show to what extent certain proteolytic activities of the proteasome are inhibited, these reporters are designed to monitor if certain conditions or compounds result in a general functional impairment of the ubiquitin-proteasome system as a whole.
To ensure that we are following the general activity of the ubiquitin-proteasome system and not only one specific type of substrate we used the two conceptually different N-end rule and UFD degradation signals in the GFP reporters. These rely on different branches of the ubiquitination machinery and hence monitoring of both substrates makes it more likely that we unmask general impairment of the system affecting a large fraction of the intracellular proteasome substrates. **Figure 5** illustrates the GFP reporters that have been used throughout the studies in this thesis.

![Diagram of N-end rule and UFD degradation signals](image)

**Figure 5.** **GFP-based proteasomal substrates.** For generation of N-end rule substrates ubiquitin was cloned in frame with GFP so that cleavage by UCH exposes the amino acid “X” in the N-terminus. We generated a stable reporter with Met (M) at position “X” and destabilised N-end rule reporters with type I (Arg, R) or type II (Leu, L) amino acids at position “X”. Lys (K) residues were placed at position 3 and 17 to serve as anchorage sites for ubiquitin-chains of the N-end rule substrates. Two UFD substrates were generated, one with Pro (P) at position “X” resulting in poor cleavage by UCH, the other one with Val (V) instead of Gly (G) as the last ubiquitin residue, which abrogates cleavage by UCH. For UFD substrates the ubiquitination occurs on the uncleavable ubiquitin moiety.

Inhibition of the ubiquitin-proteasome system induces cellular stress responses and often has a pro-apoptotic effect that can be partly attributed to the accumulation of the normally short-lived tumor suppressor p53 (10) or the cyclin inhibitor p27 that prevent progression through the cell cycle (233). The apoptotic effect is in general more pronounced in rapidly dividing transformed cells, but there are exceptions where an anti-apoptotic effect is observed. For example thymocytes that are stimulated to undergo apoptosis are protected by proteasome inhibitor treatment by the stabilisation of the inhibitors of apoptosis (IAPs) (49, 278). Moreover differentiation of neurons was stimulated after treatment with very low concentrations of proteasome inhibitors (67). For this reason we analysed whether the inhibitor induced accumulation of the
GFP reporters were indicative for biologically relevant levels of inhibition. Indeed the accumulation of fluorescence correlated with the induction of G2/M cell cycle arrest and apoptosis. The detectable fluorescence however preceded the apoptotic effect indicating the usefulness of the GFP reporters for inhibited proteolysis.

More recently additional reporters for monitoring the ubiquitin-proteasome system in cells have been developed by other research groups. One is based on acceleration of the UFD signal by multiple N-terminal ubiquitin residues in β-lactamase or GFP (237), and another is based on insertion in GFP of the artificial degradation signal CL1 (20), originally identified in yeast (79).

5.2 A mouse model for the ubiquitin-proteasome system

Transgenic mouse models represent excellent tools for studying human pathogenesis and they can be suitable for identification and evaluation of new therapeutics. For this purpose we have generated a mouse model described in detail in paper II that expresses the UFD targeted GFP reporter Ub\(^{G76V}\)-GFP. This particular construct was chosen since it is our most destabilised GFP and it is very well tolerated in tissue culture. We expressed the Ub\(^{G76V}\)-GFP from a chicken β-actin promoter with a cytomegalovirus (CMV) immediate early enhancer, which is a promoter complex that is known to induce a constitutive and general expression pattern (180). Indeed we could detect the transcript in all tissues examined. However as anticipated, inhibition of proteolysis was required in order to detect the Ub\(^{G76V}\)-GFP protein. Primary cultures of fibroblasts, cardiomyocytes and neurons demonstrated a clear dose-dependent accumulation of the GFP reporter in response to inhibitor treatment. The functionality of the reporter was also confirmed in vivo by intraperitoneal injections with the different proteasome inhibitors MG-132, MG-262 and epoxomicin. A dramatic accumulation of GFP occurred in the liver, which is in line with previous data showing that proteasome inhibitor treatment of drug-primed mice induces hepatotoxicity in the form of accumulation of Mallory bodies (74). Injection of higher concentrations resulted in massive accumulation of fluorescence in the liver, small intestine, pancreas and kidney but was less pronounced in the lungs and just a few fluorescent cells were detected in the spleen whereas the heart, brain and skeletal muscle remained non fluorescent.
We believe that this mouse, which to our knowledge is the first in vivo model for proteasomal degradation, will provide an excellent tool for studies of the ubiquitin-proteasome system in human disorders, especially the neurodegenerative disorders characterised by pathologic accumulation of misfolded and aberrant proteins, which are potential proteasome substrates. Moreover we expect it to be of great value for determining the therapeutic potential of anti-cancer drugs and for development of tissue- or tumor-specific proteasome inhibitors.

5.3 Importance of individual proteolytic sites

With the help of proteasome inhibitors and our GFP reporter model we also investigated to what extent proteolysis can be blocked before functional impairment of the system is accomplished. Two surprising findings were made.

First in paper I we report that there is an excessive activity of the ubiquitin-proteasome system and that approximately 80% of the proteasome activity can be blocked without an overall impairment of the system. Two other laboratories presented more recently data with different assays that support this finding (1, 20). We postulated that this excessive proteasomal activity enables the cells to anticipate those stress responses where the amount of misfolded and aberrant protein substrates rapidly increases. Indeed we showed that in cells lacking this backup of proteasome activity a delayed clearance of experimentally induced accumulation of substrates occurred.

Second, we show in paper III that the contribution of different proteolytic sites in protein turnover varies to a large extent. The PGPH activity was shown to be redundant for general protein turnover while inhibition of the chymotrypsin-like activity resulted in impairment of the system (see also chapter 5.3). This is in line with the observation that selective inhibition of the chymotrypsin-like activity has an antiproliferative effect in mammalian cells (62). Moreover mutations of either the chymotrypsin-like or the PGPH activity in different yeast strains are still viable (57). Importantly even though these and earlier findings provides us with some guidelines to interpret how inhibition of proteolytic activities affect the ubiquitin-proteasome system we show in paper I with inhibitor combinations targeting different sites that the net outcome remains unpredictable, which underscores the importance of the reporter system that we developed.
5.4 A modifier site regulates the proteasome

A unique feature of the proteasome complex is that it harbours three proteases with each a distinct proteolytic activity (17). How the degradation by these sites is coordinated and the individual contribution of the sites is one of the major conundrums of the system. Based on detailed studies Kisselev and co-workers proposed a “bite and chew model” according to which the catalytically active sites of the proteasome are directed by an allosteric regulatory mechanism (122). The model was mainly based on the fact that peptide substrates of the PGPH activity had an inhibitory effect on the chymotrypsin-like activity, while vice versa a substrate of the chymotrypsin-like activity stimulated the PGPH activity. Based on these findings it was suggested that when an unfolded peptide enters the inner cavity of the proteasome it is first cleaved or “bitten” by the chymotrypsin-like activity. Once “bitten” the chymotrypsin-like activity allosterically stimulates the PGPH activity in order to promote the “chewing” by the PGPH, which while “chewing” in turn inhibit further “biting” by the chymotrypsin-like site in order to avoid “choking” of the proteasome (Fig. 6). Although the model elegantly explains the in vitro observation it appears to be in conflict with several earlier observations. For example, allosteric regulation of these distant sites would require cross modifications in the proteasome structure while X-ray diffraction showed that occupation of the catalytic site did not induce such structural changes (59, 89). Moreover yeast mutant strains lacking either the PGPH or chymotrypsin-like activity are viable which does not seem to favour a unique role of these sites in an orchestrated degradation mechanism (57).

In an attempt to shed light on these controversial data we studied the role of the PGPH activity in the mammalian ubiquitin-proteasome system (paper III). Instrumental in this study were a set of novel αβ'-epoxyketone based inhibitors with various specificities towards the PGPH activity. The two most active compounds, Ac-GPFL-EX and Ac-NorLPFL-EX, inhibited the PGPH activity ~50 to ~60 times more efficiently than the chymotrypsin-like activity, as analysed by k_{cat}/[I] values. It is noteworthy that these were the first and presently most-specific PGPH inhibitors available. We observed with the GFP-reporter expressing cells that the PGPH-inhibitors, despite their potency, were very poor inhibitors of the ubiquitin-proteasome system in cells. In line with this observation the inhibitors only poorly restrained growth. The observed inhibitory effect appeared to correlate with the
modest cross-inhibition of the chymotrypsin-like activity at higher concentrations. Most importantly regarding the allosteric regulation model, when the PGPH catalytic site was blocked in vitro by an irreversible PGPH-specific inhibitor the PGPH substrate was still able to inhibit the chymotrypsin-like activity. These data suggest that while the PGPH substrates do regulate the chymotrypsin-like activity this is not accomplished through binding to the PGPH catalytic site. Hence we postulated that the proteasome harbours an unidentified substrate-binding site that can modify the catalytic activity of the proteasome. This “modifier site model” not only explains the in vitro observation that led to the “bite-and-chew model” but at the same time clarifies the studies that on first site appeared to be in conflict with each other (Fig. 6).

![Diagram](image)

Figure 6. Regulation of the PGPH and chymotrypsin-like activities of the proteasome. The “bite and chew model” implies allosteric regulation of the catalytic activities of the proteasome. According to this model binding of a substrate to the chymotrypsin-like site (“bite”) stimulates the cleavage by the PGPH activity (“chew”). This in turn inhibits further “biting” until the “chewing” is finished. According to the “modifier site model”, proposed in this thesis, a substrate of the PGPH activity inhibits the chymotrypsin-like activity by binding to a yet unidentified modifier site (MS).

Similar conclusions were drawn from another research group that in addition studied trypsin-like activity by using the HIV-1 protease inhibitor ritonavir that cross-reacts with the proteasome (5). Ritonavir inhibits chymotrypsin-like activity and at the same time stimulates the trypsin-like activity irrelevant of inhibition of the other sites further strengthening the “modifier site model” (218). Recently also Kisselev and co-workers, who originally proposed the “bite-and-chew model”, presented evidence in support of a non-catalytic modifier site in the proteasome (124).
6. THE UBIQUITIN-PROTEASOME SYSTEM AND HUMAN DISORDERS

6.1 General overview

Considering the important role that the ubiquitin-proteasome system plays in various cellular processes it is not surprising that impairment or defects in selective components of the system have been put forward as contributing or causative factors in a broad array of human disorders (Table 2). Among these diseases are cancer, genetic disorders, neurodegenerative disorders, muscle atrophy and inflammation (221, 229).

The ubiquitin-proteasome system is also being creatively exploited by many pathogens in attempts to modify the cellular environment according to their needs (50, 197). For instance by accelerated degradation of a protein through misdirection of a cellular E3 enzyme (216), blockage of degradation of 1xBα (276), through a RING finger E3 encoded in the viral genome itself (25, 91) or monoubiquitination to aid viral budding (189, 220).

Some viruses have evolved strategies to interfere with cellular systems and increase stabilisation of an oncoprotein or destabilise a tumor suppressor protein in order to prolong the life of the host cell. In some cases this can result in human malignancies caused by viral interference within the ubiquitin-proteasome system. Some high-risk strains of the human papilloma virus (HPV) can cause uterine cervical carcinomas, which are associated with very low levels of the tumor suppressor p53. This is due to the HPV encoded oncoprotein E6 that directs the E3 activity of the cellular E6 associated protein (E6-AP) to target p53 for proteasomal degradation promoting virus-induced transformation (215).

Another common feature in malignancies involving p53, that is unrelated to pathogenic infections, is increased levels of Mdm2 (10), an E3 enzyme responsible for p53 ubiquitination in normal cells (129). If Mdm2 is amplified or overexpressed the levels of p53 will decrease rendering the cell susceptible to transformation. Similar mechanisms are also underlying the downregulation of p27 by increased levels of the E3 adaptor Skp2 (150). In addition a mutation in the cellular Von Hippel-Lindau (VHL) tumor suppressor gene predisposes for a wide range of cancers. VHL recognises and recruits the α subunit of the hypoxia-inducible factor (HIF) to an E3 complex for ubiquitination and degradation in the presence of oxygen (159).
Mutations that abrogate VHL function will result in stabilisation of HIF that is a transcription factor stimulating expression of several tumor-associated growth factors (116).

Table 2. Proteins within the ubiquitin-proteasome system and their association with human diseases.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Disease</th>
<th>Function in the ubiquitin-proteasome system</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>UCH-L1</td>
<td>Parkinson's disease</td>
<td>UCH and E3</td>
<td>• Liu et al 2002 Cell 111:209</td>
</tr>
<tr>
<td>UBB⁺⁺ (mutant ubiquitin)</td>
<td>Alzheimer's disease, Down's Syndrome, progressive supranuclear palsy, etc</td>
<td>ubiquitin</td>
<td>• Van Leeuwen et al 1998 Science 279:242</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Ferguson et al 2000 Neurosci Lett. 279:69</td>
</tr>
<tr>
<td>Mdm2</td>
<td>Broad variety of human tumors</td>
<td>E3</td>
<td>• Ashcroft and Vousden 1999 Oncogene 18:763?</td>
</tr>
<tr>
<td>BRCA1</td>
<td>Breast cancer</td>
<td>E3</td>
<td>• Scully et al 2002 Biochimie 84:95</td>
</tr>
<tr>
<td>Unp</td>
<td>Lung carcinoma</td>
<td>DUB</td>
<td>• Frederick 1998 Oncogene 16:153</td>
</tr>
<tr>
<td>BAP1</td>
<td>Lung carcinoma cell lines</td>
<td>DUB that binds BRCA1</td>
<td>• Jensen et al 1998 Oncogene 16:1097</td>
</tr>
<tr>
<td>E6 (HPV)</td>
<td>Cervical cancer</td>
<td>E3 together with E6-AP</td>
<td>• Sheffner 1998 Pharmacol Thr:76:129</td>
</tr>
<tr>
<td>Avp (Adenovirus)</td>
<td>Respiratory diseases</td>
<td>DUB</td>
<td>• Balakirev et al 2002 J Virol. 76:6323</td>
</tr>
<tr>
<td>US2 and 11 (CMV)</td>
<td>Various diseases in immuno compromised patients</td>
<td>Target MHC class I for proteasomal degradation</td>
<td>• Van der Waal et al 2000 Curr Top Microbiol. 269:37</td>
</tr>
</tbody>
</table>
A number of inheritable diseases also comprise aberrations in the ubiquitin-proteasome system. Liddle’s syndrome is a genetic disorder resulting in a form of hypertension caused by mutations in the epithelial sodium channel, ENaC. The mutations abrogate its recognition by Nedd4, an E3 enzyme that stabilises the protein causing excessive reabsorption of Na⁺ and H₂O and consequently increased blood pressure (238). An additional genetic disorder involving an E3 enzyme is Angelman’s syndrome, which is a form of mental retardation. Here a deletion of chromosome 15q11-q13 comprises a loss of the E3 enzyme E6-AP that normally has a unique expression pattern due to its tissue imprinting. The paternal allele is silenced in cerebellar Purkinje cells and a mouse models carrying a maternal E6-AP deficiency has some analogous phenotypes as compared to the Angelman’s syndrome. The substrate(s) that hereby avoids ubiquitination is not yet identified (121).

The ubiquitin-proteasome system generates peptides in a non-discriminatory way from both endogenous cellular substrates, “self”, as well as from foreign derived substrates, “non-self” (211). The peptides can be used for antigen presentation by MHC class I, where normally the “non-self” peptides can elicit an immune response. Foreign peptides presented by MHC class I reveal the presence of viral proteins in infected cells. Many viruses have therefore evolved strategies to avoid recognition by interfering with the MHC class I antigen presentation pathway (see also chapter 6.2).

A key player in immune and inflammatory responses is the transcription factor NF-κB that promotes expression of many cytokines and adhesion molecules. NF-κB is activated in two ways by the ubiquitin-proteasome system 1) the NF-κB precursor protein, p105, is partially digested from its C-terminus by the proteasome to generate the active subunit p50 (41). 2) The inhibitor of NF-κB (IκBα) is degraded by the proteasome after signal-induced phosphorylation, resulting in activation of NF-κB (117). It is believed that the studies where proteasome inhibitors have been successfully used in treatment against pathologic immune and inflammatory responses were based on functional inhibition of NF-κB activity.

Muscle atrophy caused by conditions such as starvation, sepsis, denervation or diabetes can lead to pathological muscle wasting, which involves accelerated proteasomal degradation. Both an increase in the amount of ubiquitin conjugates and an upregulation of components of the ubiquitin-proteasome system are associated with these conditions. It has also been indicated that the overall rate of ubiquitination
is significantly elevated in affected muscles. An interesting observation in relation to this are the increased levels of E2_{4K} and the E3_{α} (265), that promotes ubiquitination of N-end rule substrates found in most atrophying muscles (109). The functional significance of the upregulated N-end rule pathway in atrophying muscles is however presently not understood.

6.2 Viral immune escape and the ubiquitin-proteasome system

Some viruses have the capability of establishing latency and can persist throughout life in their natural host (157). In order to achieve this different viruses have during their evolution developed traits that help them to escape from detection by the immune system and allow the viruses to coexist within their host (51, 197, 277). A variety of strategies have evolved that primarily focus on the interference at different steps in the antigen presentation pathway of viral antigens to CTLs.

Several viruses use the ubiquitin-proteasome system to down-regulate cell surface expression of MHC class I molecules that may otherwise reveal their presence in the host cell. CMV encodes several proteins that interact with the ubiquitin-proteasome system among which US2 and US11 are two small ER-resident transmembrane proteins. These proteins interfere with the quality control process in the ER that reassures that improperly folded secretory proteins are retro-translocated to the cytosol for proteasomal degradation (271, 272). Their activity promotes dislocation of the newly synthesised MHC class I molecule from the ER into the cytosol where it is subjected to proteasomal degradation (256).

It is noteworthy that proteasomal degradation is just one of the means by which viruses compromise MHC class I cell surface expression. HIV-1 (222), murine CMV (207) and Kaposi sarcoma-associated herpes virus (KSHV) manipulate membrane trafficking resulting in internalisation of MHC class I. KSHV encodes two transmembrane E3 enzymes called modulators of immune recognition 1 and 2, MIR1 and MIR2, also known by the names K3 and K5. These E3s can ubiquitinate the cytoplasmic tail of MHC class I that in turn leads to endocytosis and further trafficking to the lysosomal compartment for destruction (45, 165).

To date there are only a few isolated examples of viral immune evasion strategies that through selective inhibition of the ubiquitin-proteasome system blocks the generation of antigenic peptides. It has been shown that the Hepatitis B virus X protein (HBX) has dual substrate/inhibitor properties of the proteasome (103). The
inhibitory activity is most likely due its ability to interact with two proteasomal α-subunits that might reduce the efficacy of the generation of antigenic peptides (285). The CMV encoded phosphoprotein pp65 prevent generation of viral derived epitopes at the immediate early stage of infection (77). The first identified viral modulator of the ubiquitin-proteasome system was the Epstein-Barr virus (EBV) nuclear antigen (EBNA)-1 that is our major research object in the study presented in paper IV and will be discussed in more detail in the following chapter.

6.2.1 EBV and EBNA-1
EBV is a human γ-herpes virus that is one of the most common and widespread viruses having approximately 95% of the human population as carriers. Primary infection generally occurs early in life and is asymptomatic, while infection delayed until adolescence or adulthood can cause infectious mononucleosis, a benign self-limiting lymphoproliferative disease. After primary infection EBV establishes a life long latent infection in memory B-cells. EBV is the first identified oncivirus and it is linked to a number of human malignancies like Burkitt’s lymphoma, nasopharyngeal carcinoma, Hodgkin’s disease, post-transplant lymphoproliferative disease (208). In vitro EBV has very potent transforming activities and it is known that the viral genome encodes for at least one oncoprotein, the latent membrane protein (LMP)-1 (185). LMP-1 is a transmembrane protein that acts as a constitutively activated CD40 receptor and stimulates a signalling pathway that results in NF-κB activation (80, 170). The malignancies are associated with latency programs that are defined by different expression patterns of latent genes (157) where EBNA-1 is the only EBV protein regularly detected in all EBV-associated malignancies (139). EBNA-1 is a multifunctional protein that binds the viral episomal DNA and safeguards its maintenance during cell division (280). Moreover it is involved in replication and serves as a regulator of viral transcription (139).

6.2.2 The Gly-Ala repeat of EBNA-1: a stabilisation signal
Several laboratories consistently failed to identify EBNA-1 specific CTLs in EBV-positive individuals or patients with infectious mononucleosis while CTLs specific for other latent or lytic EBV proteins could be identified (209). It was shown that the lack of recognition of EBNA-1 was attributed to the presence of an internal Gly-Ala repeat
(143), which varies between different viral isolates from approximately 60 to over 300 amino acids (66). The presence of the Gly-Ala repeat in an antigenic protein resulted in abrogation of the recognition of cells expressing this protein by specific CTLs. An *in vitro* study revealed that the Gly-Ala repeat blocks selectively ubiquitin-proteasome-dependent proteolysis of the Gly-Ala-containing protein without influencing its ubiquitination (144). Interestingly the repeat could be transferred to other viral, cellular or artificial proteasome substrates like EBNA-4 (143), IxBα (227), p53 (96) and the GFP-based substrates (paper IV) resulting in protection of the new host proteins. Hence the Gly-Ala repeat is a stabilising transferable element that functions exclusively in cis. The only reported exception for the stabilising effect is the transmembrane protein ErbB-2, which was not protected from degradation by insertion of a short Gly-Ala repeat (195).

When the question of the absence of EBNA-1-specific CTLs was revisited using a mutant EBNA-1 lacking the Gly-Ala repeat, specific CTLs could be detected in the blood of healthy carriers (23). However, in the context of wild type EBNA-1 the CTLs could only recognise cells when the protein was exogenously added. This suggests that their presence is due to cross-priming and will not target the true EBV infected cells (22, 23).

A thorough investigation of the requirements of Gly-Ala repeat composition and the length was performed with the IxBα model system, which comprises both a constitutive as well as a potent TNFα induced degradation signal. The minimal length of the repeat that still completely protected IxBα from degradation consisted of eight amino acids (227). The composition of such an octameric repeat is of importance, for being functional it requires at least three hydrophobic residues, like Ala or Val, interspersed with one, two or maximum three Gly, a pattern strictly conserved in the natural viral repeats (226). Artificial repeats containing eight Ala also prevented degradation of IxBα (226), however long Ala repeats tend to form aggregates (16). The alternating of Ala and Gly residues in the repeat could be explained by the requirement of hydrophobic residues for its inhibitory activity while the Gly are required to avoid aggregate formation which would otherwise interfere with EBNA-1’s dynamic role.

The full protective effect with an octameric repeat in IxBα brings up the question why the Gly-Ala repeat in EBNA-1, even though it has a variable length, is
consistently long in viral isolates. To address this question we performed a study with our GFP-based reporters where we show that a very strong degradation signal, like the UFD signal, can override to some extent even the full length Gly-Ala repeat of 239 residues (paper IV). Moreover we observed a clear length-dependent stabilising effect detected with both the UFD and the N-end rule GFP substrates. We indeed demonstrated that even EBNA-1 could be degraded by the proteasome once provided with a strong degradation signal. In an independent study Tellam and co-workers showed later that targeting of EBNA-1 with similar degradation signals also resulted in processing and recognition of EBNA-1-derived epitopes by CTLs (248). We concluded from these experiments that the turnover rate of a Gly-Ala repeat-containing protein is determined by the length of the inhibitory repeat in concert with the efficiency and nature of the degradation signal.

In agreement with the length-dependent effect of the Gly-Ala repeat in the GFP reporters similar results were obtained in a continuation of the IκBα study but this time with a suboptimal repeat containing an Ala to Ser substitution (226). In this case three repeats were required for blockage of the constitutive degradation but these repeats were insufficient to block the TNFα-induced degradation. The strength of the degradation signal is likely to determine the ubiquitination rate. Accordingly it was shown with p53/Gly-Ala repeat fusions that increasing amounts of its E3 enzyme, Mdm2, also override the stabilising effect of the Gly-Ala repeat, which could again be countered by increasing the repeat length (96).

Interestingly EBNA1 homologues are expressed in the baboon and rhesus monkey lymphocryptoviruses. The repeat lengths in these proteins are substantially shorter, ranging from approximately 50 to 70 amino acids, and they also include Ser residues. These repeats too have a protective effect on proteasomal degradation but to lesser extent resulting in antigen presentation of peptides derived from these EBNA-1s (24). Supposedly the stabilising effect of the Gly-Ala repeat or its homologues reflects an ancestral function for the EBNA-1 protein, whereas the immune evasion strategy of EBV evolved at a later stage by amplification of the repeat (51). Thus even though short repeats do delay protein turnover and can be fully protective in some proteins, EBNA-1 may carry consistently long repeats to ensure that no antigenic peptides will be generated.
The underlying mechanism of how the Gly-Ala repeat exerts its protective effect remains poorly understood. However, extensive studies have excluded some hypotheses and narrowed down the possible location where the inhibition takes place. In vitro (144) and tissue culture studies (96, 227), revealed that Gly-Ala repeat-containing proteins are still subject to ubiquitination placing the inhibitory effect downstream of ubiquitination. It is unlikely that the failure of the proteasome to degrade the Gly-Ala repeat is due to structural constraints as structural analysis of one of the stable IκBα/Gly-Ala repeat chimera revealed that the Gly-Ala repeat lacks a defined structure (140). Although formally it cannot be excluded, it appears that the Gly-Ala repeat does not affect the proteolytic activity. In vitro hydrolysis of fluorogenic substrates by the proteasome was not inhibited by the presence of a synthetic Gly-Ala repeat oligopeptide (141). Moreover, the study presented in paper IV also shows that once provided with a strong degradation signal the Gly-Ala repeat can be degraded by the proteasome arguing against structural constraints or inhibition of proteolytic activity as the mode of action. It is tempting to speculate that the interaction between the polyubiquitinated Gly-Ala repeat protein and the proteasome is modified or abrogated. Yet by looking at the possible binding of Gly-Ala repeat-chimeras to the proteasome some confusion was raised. First it was shown that IκBα but not IκBα/Gly-Ala repeat could be co-immunoprecipitated with the proteasome from cell lysates (227). By using a more sensitive in vitro method it was shown later that polyubiquitinated p53/Gly-Ala repeat fusion could still interact with S5a (96). However, for interpretations of proteins interacting with S5a it should be noted that most S5a exists in free form (260) and it is not ruled out that there might be qualitative differences between the interaction of the free S5a and the proteasome associated S5a.

A hypothetical mode of action of the Gly-Ala repeat could be that the ubiquitinated Gly-Ala containing protein reaches the proteasome where the long hydrophobic stretch may modify the interaction and consequently the sequence of events that follow upon binding. It is unlikely that the mechanism of stabilisation would be due to problematic unfolding of the protein by the ATPases of the proteasome since the Gly-Ala repeat has a poor structure. However the proteasomal ATPases have also been suggested to have chaperone activities (26, 241), so it is possible that the Gly-Ala repeat instead triggers refolding attempts of the proteasome.
It is tempting to speculate that due some altered interactions by the Gly-Ala repeat, the protein is deubiquitinated before it is irreversibly captured for degradation. By this manner the Gly-Ala repeat containing protein might get caught in a perpetuate cycle of ubiquitination, interaction with the proteasome, deubiquitination, and release from the proteasome.

Intriguingly the p105 precursor of the NF-κB subunit p50 is ubiquitinated and partially processed until approximately the centre of the protein where the GRR is located (41) (see also chapter 3.6). The GRR share some similarities in the sequence to the Gly-Ala repeat and it also stabilises p50 to some extent. So far this sequence has however not proven to be a transferable stabilising element like the Gly-Ala repeat (181).

6.3 The ubiquitin-proteasome system in conformational diseases

“Conformational diseases” is a recent denomination on a group of disorders that share a common feature in the accumulation of insoluble protein deposits in the affected cells. To this group belong α1-antitrypsin deficiency, amyloidoses, cystic fibrosis and a large group of neurodegenerative disorders such as Alzheimer’s disease, Parkinson’s disease, Huntington’s disease and prion encephalopathies (32, 127). The deposits are referred to as aggregates, aggresomes or inclusions and although they vary in protein composition, shape and localisation, each of these structures are mainly composed of insoluble misfolded proteins sequestered together with different molecular chaperones and components of the ubiquitin-proteasome system (3, 229). The presence of these factors suggests that the aggregated proteins are, or have been, targeted for degradation or refolding but instead of being inactivated by the proteolytic and refolding machineries they accumulate in insoluble protein deposits (73, 232). The colocalisation with the constituents of the ubiquitin-proteasome system is not surprising since one of their fundamental tasks is to free cells from damaged or abnormal proteins and to protect cells during stress responses. It has indeed been shown in several studies that overexpression of different chaperones in cellular or animal models counteract the formation of aggregates (11, 48, 94, 128, 264) and even increase their degradation by the proteasome (15) pointing out the critical role of chaperones in these disorders.
The largest group of conformational disorders are the neurodegenerative diseases. Each disease has its own characteristic type of protein deposits like the neurofibrillary tangles and plaques of Alzheimer’s disease (93), Lewy bodies of Parkinson’s disease (76), Bunina bodies of amyotrophic lateral sclerosis, Pick bodies of Pick’s disease and intranuclear inclusions of poly-Gln disorders (286). Parkinson’s disease is particularly interesting regarding the ubiquitin-proteasome system since some mutations associated with familial forms of the disease are located in proteins that are closely linked to the ubiquitin-protein system: Parkin (125, 230) and UCH-L1 (142, 149).

Parkin is an E3 enzyme that has a general role in the unfolded protein response and degradation of ER proteins (108). Mutations in parkin associated with disease development typically destroy its ubiquitin ligase activity. Recently it has also been shown that allelic variations in parkin or its promoter may underlie the more common non-familial forms of the disease (268). Amongst others parkin has been shown to ubiquitinate the membrane receptor Pael-R (107) and synphilin-1 (39), an α-synuclein-interacting protein and major component of the Lewy bodies. Mutations in UCH-L1, a ubiquitin C-terminal hydrolase, observed in a familial form of Parkinson (142) on the contrary cannot be explained by a complete loss-of-function phenotype. Instead it was recently explained by the important finding that UCH-L1 in addition harbours E3 activity in its dimeric form. Reduced ubiquitin ligase activity of UCH-L1 is correlated with a reduced risk for Parkinson’s disease (149).

6.3.1 Characteristics of poly-Gln proteins

There are several neurodegenerative disorders that are caused by an expansion of a poly-Gln repeat in different proteins for example Huntington’s disease, spinocerebellar ataxia (SCA) type 1, 2, 3, 6 and 7, dentatorubral-pallidoluysian atrophy (DRPLA) and spinal and bulbar muscular atrophy (SBMA) (286). They show different symptoms but are all progressive and fatal neurological diseases that typically are initiated in midlife. Despite their clinical differences, there are strong indications that they share a common pathogenic mechanism linked to a gain-of-function attributable to the expanded poly-Gln repeat.

First, a specific subset of neurons is for unknown reasons particularly vulnerable despite expression of the poly-Gln repeat protein in other neuronal and non-neuronal cells. Second, all poly-Gln diseases show a striking positive correlation
between the length of the repeat and the severity of the disease: the longer the repeat the earlier onset and more severe disease. Third, the longest natural occurring poly-Gln repeats in the different proteins cause overlapping symptoms with the other poly-Gln disorders, which probably can be explained by a partial loss of cell- or tissue-specificity (286). In the most extreme case with SCA7 also non-neuronal tissue, e.g. heart can get affected and lead to congestive heart failure (21). Fourth, all disease proteins with an expanded poly-Gln repeat, except for SCA2, tend to form intranuclear aggregates with fibrillar structure. The combination of a high concentration and very long repeat increase the tendency to form aggregates (263). Mathematical models have predicted that neurodegeneration is caused by a stochastic "single hit" event (42, 43) and it has been proposed that aggregate formation is a nucleation dependent process that would fulfil the requirements of the single hit (190). Fifth, the aggregates colocalise with proteins involved in the ubiquitin-proteasome system and with chaperones (46) (see also chapter 6.3). Sixth, most diseases have primarily dominant traits indicating a toxic gain-of-function effect. This is illustrated by the fact that a homozygote patient with Huntington’s disease has a similar phenotype as the heterozygote sibling (269). Moreover deletion of huntingtin, the responsible protein for Huntington’s disease, does not cause Huntington’s disease in humans or in mouse models (4, 60), whereas an expansion of the Gln-repeat in the same protein does. Further indications of a gain-of-function effect are the null mutation of ataxin-1, the protein responsible for SCA1, in mice or a large deletion in humans that do not cause ataxia, which is characteristic of this disease. The mice however displayed learning deficits (158) and humans suffered from mental retardation (52).

It is important to emphasise that considerable efforts have been focused on elucidating the role of the accumulation of aggregates in the neuropathology. So far there are no unanimous conclusions proving whether they are harmful or whether they represent a beneficial cellular response based sequestration of dangerous proteins. Figure 7 below show the possible relation between soluble and aggregated poly-Gln proteins.
Figure 7. Interrelation between soluble and aggregated poly-Gln proteins. Proteins with long poly-Gln repeats tend to form fibrillary aggregates in cells, the biological consequence of these aggregates is not yet clarified. This model gives an overview of the relation between the soluble and aggregated part of these proteins and what factors can influence these processes.

There are a number of indications that intranuclear aggregates are not required for toxicity. One example is that patients with SCA2 lack such aggregates, there is however an increase in the amount of soluble protein in the cytoplasm or formation of cytoplasmic microaggregates in neurons in the affected area (104, 105). Moreover in a mouse model the loss of the E3 ligase E6-AP activity inhibited the formation of ataxin-1 aggregates but accelerated the SCA1-associated neurodegeneration (47). Furthermore overexpression of a heat-shock protein in a mouse model for SCA1 did not influence the aggregate formation but caused a milder progression of the disease indicating a lack of toxicity of the aggregates (48).

Toxicity derived from aggregates was recently supported by a study using two non-disease related proteins that are prone to form fibrillary aggregates in vitro. The early aggregates have a granular conformation, which later transforms into fibrillar structures. In this study it was shown that the presence of such early developed aggregates were inherently toxic to cells (28). Another elegant study supporting the toxicity of aggregates was performed in yeast. Here a yeast prion protein was identified as an initiator that seeds the aggregates. With this model it was shown that not the amount of poly-Gln protein but their presence in insoluble aggregates determines their toxicity in yeast (167).

When proteins with long Gln-repeats aggregate they can sequester some of their binding partners, other Gln-containing proteins, chaperones, ubiquitin and
proteasomes. This is one of the characteristics of the aggregates that favour the hypothesis of a toxic effect of poly-Gln aggregation. Many transcription factors have Gln-rich sequences and are at risk of being captured within the aggregates. This has been demonstrated to be the case with the CREB- and TATA-binding proteins (160) (179). However recent data indicated that the aggregates do not form completely static structures and that proteasomes, chaperones and CREB-binding protein can be transiently associated with these structures while the proteins containing the expanded poly-Gln repeat mostly are less dynamic (120, 239). Theoretically the sequestration of transcription factors could modify gene expression and indeed microarray analysis of brains from mouse models of SCA1 and Huntington’s disease show altered expression patterns in the presence of aggregates (153-155). Thus, by being captured into aggregates essential proteins might indirectly lose their proper function eventually leading to cellular malfunctions or culminate in cell death.

6.3.2 Poly-Gln repeat expansions and proteasomal degradation
The hallmarks of the poly-Gln repeat disorders as well as other conformational disorders are the presence of aggregates that often stain positive for components of the ubiquitin-proteasome system (see also chapter 6.3). It has been suggested that the poly-Gln aggregates can directly inhibit proteolysis possibly by clogging the proteasome with poly-Gln repeat or indirectly by sequestering proteasomes, ubiquitin and other essential components into the aggregates (73, 229, 232). The presence of the ubiquitin-proteasome system in the aggregates can however be interpreted in two ways: the components are sequestered in the aggregates and therefore cause improper proteolysis at other locations in the cell or the components are recruited to the aggregates in order to help the cell to clear or minimise the toxicity of the aggregated proteins.

Inhibition of the ubiquitin-proteasome system by poly-Gln aggregates was observed in cellular models including a model based on another type of GFP-based proteasome substrate (20, 110). However, we investigated different poly-Gln aggregating proteins in cells expressing either the N-end rule or the UFD targeted GFP reporter but did not observe any general inhibition of proteasomal degradation (K. Lindsten, L.G.G.C. Verhoeef, M.G. Masucci, N.P. Dantuma unpublished data). Further studies to answer this very important question have high priority.
We investigated the effect of the poly-Gln repeat on ubiquitin-proteasome-dependent proteolysis. It had been shown previously in an \textit{in vitro} study that long poly-Gln repeats could delay protein turnover, providing a possible mechanism for their accumulation in neurodegeneration (47). When we analysed the effect of the poly-Gln repeat on degradation of the GFP-based reporters, similarly to the previously mentioned Gly-Ala repeat experiments, we did not observe such a stabilising effect (\textbf{paper V}). These experiments showed that when targeted with a degradation signal the proteins were efficiently degraded in cells regardless of the length of the poly-Gln repeat, which stands in strong contrast to the length-dependent inhibitory effect observed with the Gly-Ala repeat (\textbf{paper IV}). A different story emerged when we focused on the fraction of the cells that developed aggregates as a consequence of the expanded poly-Gln repeat. In contrast to the cells that displayed a soluble staining, these cells failed to degrade the poly-Gln proteins.

In agreement with previous studies where poly-Gln aggregation leads to sequestration of other proteins, we observed similar recruitment of proteins, which in addition resulted in stabilisation of the co-aggregated protein. Together these data demonstrate that aggregate formation in itself can result in protection of proteins from proteasomal degradation, an important finding in the light of the suspected role of the ubiquitin-proteasome system in conformational diseases.

There was in addition a positive correlation between the stability of the poly-Gln protein, the amount of aggregates and cytotoxicity. The reporters with a strong degradation signal could be degraded and as a consequence fewer cells with aggregates and reduced cytotoxicity was observed.

\textbf{6.3.3 Aberrant ubiquitin, UBB$^{\mathrm{14}}$, and conformational diseases}

An aberrant form of ubiquitin, referred to as UBB$^{\mathrm{14}}$ and a mutant variant of \(\beta\) amyloid precursor protein (APP$^{\mathrm{14}}$) were originally found in the neurofibrillary tangles, neuritic plaques and neutropil threads of patients with Alzheimer’s disease and Down’s Syndrome (258). The aberrations result in a ubiquitin protein with an extended C-terminus for UBB$^{\mathrm{14}}$ and a truncation of \(\beta\)APP$^{\mathrm{14}}$. These mutations occur by a process described 10 years ago designated “molecular misreading” (65). Molecular misreading imply a transcriptional error that occurs at low frequency resulting in RNA transcripts with dinucleotide deletions in or adjacent to short repetitive
nucleotide sequences (257). The exact mechanism causing the erroneous transcripts is not known but an attractive explanation is that the RNA polymerase has difficulties with certain repetitive sequences and slips over one or two nucleotides. Molecular misreading is likely to be a common event since it has been demonstrated to occur in different organs at different ages in a transgenic mouse model (259). The UBB\textsuperscript{+1} transcripts have also been detected in brains of non-demented controls (F.W. van Leeuwen, personal communication) but the presence of the UBB\textsuperscript{+1} protein and other +1 protein products is restricted to affected cells in conformational diseases.

Recently the presence of UBB\textsuperscript{+1} protein was shown to be a more widely spread phenomena in conformational diseases present also in the brains of patients with progressive supranuclear palsy (69), frontotemporal dementia, Pick’s disease, argyrophilic grain disease and Huntington’s disease. On the contrary UBB\textsuperscript{−1} was absent in major forms of so-called synucleinopathies, such as Parkinson’s disease and multisystem atrophy (F.W. van Leeuwen, personal communication). Interestingly UBB\textsuperscript{−1} accumulation is not restricted to neurodegenerative disorders as its presence has also been demonstrated in the so-called Mallory bodies in livers from patients with steatohepatitis and hepatocellular carcinoma (163). Mallory bodies are cytoplasmic aggregates that constitute a morphological hallmark of alcoholic and non-alcoholic steatohepatitis and is occasionally also found in hepatocellular carcinomas (56).

The ubiquitin B gene encodes one of the ubiquitin precursors, which in this case consists of three tandemly organised ubiquitin moieties. Once translated, the C-terminal Gly residue of each ubiquitin will be recognised and cleaved by ubiquitin C-terminal hydrolases that separates the ubiquitins into monomers (Fig. 2). Close to the 3' end of the first ubiquitin moiety a GAGAG nucleotide sequence can give rise to molecular misreading leading to a dinucleotide deletion in the transcript (ΔGU). The outcome is a frameshift of the open reading frame and the introduction of a premature stop codon. The encoded aberrant UBB\textsuperscript{+1} protein will consist of the wild-type sequence until the Gly76 residue where the frameshift occurs. The Gly is substituted to a Tyr followed by 19 extra amino acids encoded by the frame-shifted open reading frame (Fig. 8) (258). As a consequence of the amino acid substitution at position 76, the ubiquitin moiety cannot be removed from the 19-amino acid extension nor can it conjugate to other substrates or ubiquitins (54, 133). UBB\textsuperscript{+1} however contains all the Lys residues present in normal ubiquitin and hence UBB\textsuperscript{+1} itself can be
polyubiquitinated resulting in a tree that resembles an unanchored polyubiquitin chain (133).

![Diagram](image)

Figure 8. Molecular misreading of the ubiquitin B transcript. Molecular misreading can occur in or adjacent to repetitive nucleotide sequences, where it causes a small frame shift deletion (1-2 nucleotides) in the RNA. The ubiquitin B gene encodes three tandem repetitive ubiquitins with a GAGAG nucleotide repeat in the end of the first ubiquitin moiety. Molecular misreading (ΔGU) can occur adjacent to this repeat causing a frame shift exchanging the last Gly (G) residue to a Tyr (Y) and introducing a premature stop codon. The normal ubiquitin B polyprotein can be processed into functional ubiquitin monomers by cleavage of UCHs, which on the other hand will be incapable of recognizing and cleaving the mutated UBB^1.

6.3.5 Effect of UBB^1 on proteasomal degradation

The presence of the aberrant ubiquitin in the brains of patients with neurodegenerative disorders inspired us to investigate if the presence of this protein had any effect on the ubiquitin-proteasome system in cells (paper VI). It had been recently shown that overexpression of UBB^1 but not ubiquitin caused neuronal cell death (54). Moreover, in in vitro experiments it was shown that artificially ubiquitinated UBB^1 could block degradation of a model proteasome substrate (133). To address our question we transfected two neuroblastoma cell lines expressing different GFP-based proteasome substrates with UBB^1-expressing vectors. Indeed a strong inhibition of the ubiquitin-proteasome system was observed and the vast majority of the UBB^1 positive cells accumulated the GFP substrates indicative for a functional impaired ubiquitin-proteasome system. UBB^1 affected a post-ubiquitination event, which was concluded
by the observed accumulation of ubiquitin conjugates in cells accumulating the GFP substrates ruling out the possibility that UBB\(^{+1}\) inhibited the ubiquitination machinery.

It had been previously shown that the steady state levels of UBB\(^{+1}\) always contain a fraction of ubiquitinated UBB\(^{+1}\), bearing one to three additional conjugated ubiquitins. To detect the presence of clear ubiquitin conjugates is unusual due to the rapid deubiquitination of substrates by the cellular DUBs. In the case of UBB\(^{+1}\) this can be explained by the observation that polyubiquitinated UBB\(^{+1}\) is partially resistant to deubiquitination (133). We asked whether the ubiquitination of UBB\(^{+1}\) influenced the toxic effect seen in earlier studies. To this end we substituted the Lys residue at position 48, which is the most common ubiquitin conjugation site in ubiquitin. To our surprise this mutant was still subject to ubiquitination and therefore we continued and generated substitution of Lys29, which is known to be specifically involved in ubiquitination of UFD substrates. We showed that either Lys29 or Lys48 were required for ubiquitination of UBB\(^{+1}\), whereas the mutant lacking both Lys29 and Lys48 resisted ubiquitination.

Since UBB\(^{+1}\) in previous studies had been suggested to be a rather stable protein (54) and this ubiquitination pattern suggested that UBB\(^{+1}\) is recognised as a proteasome substrate, we decided to readdress this question. Indeed by using a double expression system in which UBB\(^{+1}\) and an unmodified GFP were expressed from two different promoters in the same plasmid, we could reveal that a large fraction of the transfected cells had undetectable levels of UBB\(^{+1}\). Yet treatment of the cells with specific proteasome inhibitors increased the steady state levels in the cells resulting in positive staining of nearly all transfected cells. This shows that UBB\(^{+1}\) is efficiently degraded in the vast majority of the transfected cells while it accumulates in a small population where it causes a general impairment of the ubiquitin-proteasome system.

The clues leading a step closer to elucidation of the inhibitory mechanism of UBB\(^{+1}\) came from results showing that its ubiquitination was essential for inhibition of the ubiquitin-proteasome system. This was further emphasised by adding multiple UFD signals to the N-terminus of UBB\(^{+1}\) in an attempt to degrade it more efficiently and reverse the inhibitory effect. This approach to accelerate protein turnover has been described to be effective for degradation of different proteins (237). Most likely a massive ubiquitination of such substrates targets them more efficiently for degradation. Surprisingly the additional UFD signals on UBB\(^{+1}\) had an opposing
effect since UBB$^\text{+1}$ started to accumulate instead of being degraded. The effect was dependent on the efficiency of the degradation signal. In agreement with the increased amount of UBB$^\text{+1}$ the inhibitory effect on the ubiquitin-proteasome system was enhanced with the increased targeting of UBB$^\text{+1}$ for degradation.

Based on these data, presented in paper VI, a hypothetical model of UBB$^\text{+1}$ inhibition of proteolysis can be put together. In healthy neurons the production of UBB$^\text{+1}$ occurs but is recognised as a substrate of the proteasome and is rapidly destroyed by the ubiquitin-proteasome system. For unknown reasons, possibly compromised proteasome activity due to age (118, 119), aggregate formation (20), or cellular stress responses (58), UBB$^\text{+1}$ passes a threshold level upon which it starts to accumulate. This could potentially initiate a negative feedback loop inhibiting its own as well as others proteasomal degradation. Due to the positive correlation between targeting for proteasomal degradation and inhibition it is tempting to speculate that the inhibitory effect occurs at the site of the proteasome. We indeed confirmed that the polyubiquitinated UBB$^\text{+1}$ interacts with subunits of the 19S regulator of the proteasome (K. Lindsten, A. Leonchiks, L.G.G.C. Verhoef, M.G. Masucci, N.P. Dantuma, unpublished data). Whether the inhibitory effect of UBB$^\text{+1}$ is of importance for neuropathogenesis or whether the production of UBB$^\text{+1}$ is more a bystander event in already severely affected cells remains to be investigated.
The work of this thesis was initiated with the successful development of GFP-based proteasome substrates that allow monitoring of inhibited activity of the ubiquitin-proteasome system described in paper I. By means of transiently or stably expressing the GFP reporters in cells, compromised ubiquitin-proteasome system can be followed by measuring the emitted fluorescence. These GFP reporters provided a valuable tool and laid the foundation for the studies that followed throughout this thesis. Beyond the use of the GFP reporters as research tools in cell biology they have the potential of being used for high throughput screenings for small compound inhibitors of the ubiquitin-proteasome system. This is of special interest since it is well documented that proteasome inhibitors have therapeutic potential, especially as anti-cancer drugs but also as anti-inflammatory agents or immunomodulators.

The GFP reporter system was recently expanded to a transgenic mouse model described in paper II. This mouse model provides a reliable and convenient way for detection of the activity of the ubiquitin-proteasome system in vivo. Administration of proteasome inhibitors results in accumulation of GFP in different organs that can be visualised by fluorescence microscopy. This in vivo model will be used in future studies for evaluating the activity of the ubiquitin-proteasome system in various human disorders. Inhibited proteasomal degradation is suspected to be involved in several neurodegenerative disorders. Crossing our transgenic mouse model with such neurodegenerative mouse models could provide a new approach to get closer to the understanding of the contribution of the ubiquitin-proteasome system in human neuropathogenesis. Another possible application of our transgenic mouse model is to study the bioavailability in combination with the therapeutic potential of proteasome inhibitors. Of special interest would be to develop tissue- or tumor-specific inhibitors with the help of this transgenic mouse model.

In paper III we used the GFP reporters to gain insight in the regulation of the proteolytic activities of the proteasome. With new inhibitors specific for the PGPH activity we demonstrated that blockage of this proteolytic site, in sharp contrast to inhibition of the chymotrypsin-like activity, does not have an overall inhibitory effect on the ubiquitin-proteasome system. Furthermore we suggested that substrates of the proteasome most likely interact with a yet unidentified “modifier site” different from the active sites, with the purpose of regulating the different proteasomal activities. It
would be interesting to identify such alternative binding sites in order to gain better understanding of proteasomal degradation.

Turnover of proteins by the ubiquitin-proteasome system is regulated at many different levels. Substrates can contain stabilising or destabilising domains that influence their rate of degradation. In this thesis two divergent repetitive sequences that influence protein stability have been studied, the Gly-Ala repeat of EBV and the expanded poly-Gln repeat present in different cellular proteins causing certain forms of neurodegeneration.

The Gly-Ala repeat, that was the focus of paper IV, is a transferable cis-acting stabilisation signal capable of protecting substrates from proteasomal degradation by a yet unknown mechanism. Here we demonstrated that there is a length-dependent effect behind this inhibitory mechanism. Modifications with a short repeat can be enough for stabilising a protein that is moderately unstable, however a highly destabilised substrate requires a longer repeat to prevent its degradation. Based on previous studies the protective mechanism of the Gly-Ala repeat can be narrowed down to an event between the ubiquitination and proteasomal degradation. We have postulated that the Gly-Ala repeat repeat promotes an interaction with a binding partner and that the length of the repeat strengthens this contact. We further speculate that this could lead to an altered interaction with the proteasome promoting the premature release of the protein containing the Gly-Ala repeat before it is unfolded and degraded.

Further studies will be focussed on gaining insight in the protective mechanism of the Gly-Ala repeat by searching for possible interacting partners in the regulatory particle of the proteasome. For such studies one approach is to create GST fusion proteins of the proteasomal subunits of interest and look for interactions with the Gly-Ala repeat containing protein in GST pull down experiments. Another approach to unravel a possible interacting protein is to develop a corresponding Gly-Ala repeat system in yeast and use it in various genetic screens.

In paper V we show that the expanded poly-Gln repeats in contrast to the Gly-Ala repeat confer a stabilising effect by the formation of aggregates. However in contrast to what have been suggested previously expanded poly-Gln repeats, as such, do not seem to hinder proteasomal degradation as the soluble portion of such proteins can be readily degraded. Furthermore, we showed that the sequestration of substrates with short poly-Gln repeat into the aggregates results in stabilisation of these proteins.
An important implication of this finding is that, regardless whether the aggregates have a protective or harmful effect, a possible therapeutic approach aimed at reducing the toxic protein should be aimed at events before aggregate formation. An interesting approach for this purpose would be to target the soluble poly-Gln containing protein for proteasomal degradation by artificially designed ubiquitin ligases. A very important question in this field is to determine the role of the aggregates in neuropathogenesis: are they toxic or do they represent a beneficial cellular response mechanism? A toxic effect could be due to inhibition of proteasomal degradation in the presence of aggregates as suggested by another research group. In contrast our preliminary data obtained with our different GFP reporters do not indicate a general impairment of the ubiquitin-proteasome system in aggregate-containing cells.

In paper VI we studied the aberrant form of ubiquitin, UBB*1, and took an important step when showing that it causes a general blockage of proteasomal degradation in neuronal cells. We also demonstrated that UBB*1 is both a substrate and an inhibitor of the ubiquitin-proteasome system and that the inhibitory effect is dependent on ubiquitination of UBB*1. These results contribute to a better understanding of the effect and mechanism of action of UBB*1 in cells. One key question for future studies is to reveal the complete underlying mechanism of its inhibitory function. For example it would be interesting to investigate if ubiquitinated UBB*1 by obstructive binding to the proteasome can inhibit its activity. As suggested for the future studies with the Gly-Ala repeat above, a sensitive way for detecting interactions with the proteasome is by using GST fusions to proteasomal subunits. Another interesting possibility for inhibition of a protein post-ubiquitination that is worth investigating would be the interference of UBB*1 with deubiquitination enzymes in cells. Finally high priority should be devoted to studies aimed at elucidation of the actual role of UBB*1 in neurodegeneration.
ACKNOWLEDGEMENTS

I would like to express my gratitude to all of those who in one way or another have helped me to finish my PhD studies. In particular I would like to thank:

Nico, my lovely boyfriend ♥... and supervisor. You are my best friend and ‘ik hou van jou’, you make my life wonderful. Thank you for the tremendous support and inspiration that you always give me. I would also like to thank you for being an excellent supervisor and scientist, all students’ scientific role model!

Maria Masucci, for accepting me as a PhD student and for showing such genuine ambition and scientific competence. Thank you also for pushing me into the mouse project... I am really happy that you did that.

Stijn Heessen alias ‘Mr July’ or the creative man who discovered that ubiquitin is actually a SUMO-like protein. Thank you first of all for friendship but also for gourmet dinners, parties, skiing and all the other good things we have experienced together. Thanks also for many very interesting discussions, private and scientific, and for always sharing your latest ‘hot’ results with great enthusiasm.

Lisette Verhoef for being a very close friend who is always ready for a chat, preferably for a long time at very late hours. It’s really great to have you around 😊. Thanks also for great collaborations in the lab, for inspiring veggie cooking and for teaching me to recycle paper.

Sweet Victoria Menéndez-Benito, thank you for bringing a crazy happy atmosphere to the lab and for being a very kind and good amiga. Also for your help and patience with all our joined projects and adventures with the transgenic mice.

Anna Cartier for great friendship and for understanding how wonderful dogs are! Thanks also for all the nice discussions that we had and will have in the future I hope.🐶

My favourite Marianne Jelline, the mother of the lab. Thanks for being so kind and understanding. I really appreciate all our nice discussions, including all the gossip that we warmly share between us. Thanks also for support and of course for teaching me many different techniques in the lab and giving a helping hand at any moment.

Yvonne Strandberg for your excellent organisation skills. You really take care of all of us so well. However thank you most of all for being a good friend with a very warm personality and positive attitude.

Victor Levitsky, my excellent supervisor in the MGM group when I first joined for a summer job. Thank you for being so kind and always ready to share your broad knowledge with the students.

The rest of the MGM group: Anders Bergqvist for nice discussions and technical support. Iris van Dijk Härd, thanks for friendship, your warm heart, nice dinners and good discussions. Piri for being very kind and helpful. Many thanks also to the younger generation of students in the group creating a very good atmosphere: André for bringing an exotic Brazilian feeling to the lab, but most of all for being such a nice guy. Micke for being a very kind and nice friend. Vaya for being so generous and always ready to share your tempting lunchboxes or cookies with the rest of us. Elena for many friendly chats, Sara for being a very sweet and kind friend. Ulrika for good chats.
Former members of the group:
Ainars for being a warm, pleasant and very kind person. Teresa Frisan for being so extremely nice and helpful. Robert Skraban for friendship and fun company, Jelena Petrovic for friendship and many enjoyable discussions and parties. Simona and Rikard Glas (basta lyckönskningar till er alla tre!). Thanks also to Anatoly, Jelena Levitskaya, Riccardo Gavioli, Rebeca Valero and Tomoaki for friendship and help in the lab. A very special thanks goes to Xu-Hong for being such a generous friend and for spreading smiles through the corridors.
You are like a walking sunshine!

Other friends at MTC: Jakob Michaëlsson for long lasting and very dear friendship!!! Thanks for all the good memories that we share from France. Those were great days I must say. Tanja Näslund, Årtan, thanks for great fun and a lot of nice parties, but mainly because I feel that you are a friend who is always there for me. Åsa Björndahl and Laszlo Szekely for friendship and for inspiring dives and discussions, but also for help with the microscope of course. Thanks also to Kerstin Andreasson, Birgitta Wester, Britt Samuelsson for all nice discussions in the lunch-room and for technical assistance. Petter Höglund for help with our mouse project. Marie Henriksson for nice discussions. All of you together really create a good atmosphere at MTC.

The Chinese community that left a very empty space behind: Cheng-Hong my dear friend, thanks for being such a warm and kind person. Daronng my special friend, you are also so kind and thoughtful. Thanks to both of you for close friendship and for all help and the absolutely delicious food… that would be with one tiny exception; the chicken feet… LiQi Li for friendship, you are my hero, so strong… I miss you. I hope to see you all soon again, either in the States or in China!

My collaborators and friends in The Netherlands; Fred van Leeuwen, Elly Hol, David Fischer, Femke De Vrij, Paula, Remko, Jacqueline and all the rest. I had a great time and learned a lot when I visited your lab. A special thanks goes to Femke for being a very good friend.

Jan Konvalinka and Tatana Uhlíkova our in Prague. Thank you for fruitful collaborations and nice times together. Hope to see you in Prague next time!

Craig Crews for a generous collaboration on a very interesting project.

Niklas för gamla goda tider.

Some other old dear friends; Thomas, Bettina or “Bettan”, Cilla, Kristina B, and Valli.

Åsa Hallqvist for being such a close friend for a long time, at all times. I’m so happy that we stick together! Thanks for all the inspiring and open discussions about life and for pep-talks hos ‘italianarn på hörnet’. I wish you, Janne and “Janne junior” all happiness in the world.

Nico’s wonderful and kind family in The Netherlands, Ineke, Bouwe, Gerben, Peter, Janny, Michel, Bianca, Milo, Patrick, Danielle, Koen, Bram, Inge, Eric and Romi. Thank you all for making me feel so welcome in your family.

My FAMILY: Mamma, Pappa, Luffi, Magnus, Marie and last but not least little Ricky boy, the sweetest dog alive, really! My love to all of you is beyond words, you are the heart in my life. Thank you for always supporting me.
REFERENCES


68


