Regulation of the ubiquitin-proteasome system: characterization of viral and cellular stabilization signals

Stijn Heessen

Stockholm 2003
All previously published papers were reproduced with permission from the publisher.

Published and printed by Karolinska University Press
Box 200, SE-171 77 Stockholm, Sweden
© Stijn Heessen, 2003
ISBN 91-7349-600-6
To Sandra and my family.

“The beauty of nature lies in detail; the message in generality”.

Stephen Jay Gould (Wonderful Life)

“I read somewhere that you’ve got to beware,
you can't believe anything you read”

Jack Johnson (It's all understood)
ABSTRACT

The ubiquitin-proteasome system plays a fundamental role in virtually every cellular process. Degradation of endogenous proteins by this system is the major source for antigenic peptides that are presented to MHC class I-restricted cytotoxic T cells. The Epstein-Barr virus (EBV) nuclear antigen-1 (EBNA1) contains a long glycine-alanine (GA) repeat that inhibits proteasomal processing, resulting in a blockage of antigen presentation. The GA repeat acts in cis and can be functionally transferred to other proteins. The aim of the work described in this thesis was to characterize the protective effect of this viral stabilization signal and to identify similar cellular stabilization signals.

To evaluate the inhibitory activity of the GA repeat, we used green fluorescent protein (GFP)-based reporters that were targeted for ubiquitin/proteasome-dependent proteolysis by various degrons. Introducing GA repeats of increasing length resulted in enhanced protection of the fluorescent reporter from proteolysis. When provided with a strong degradation signal even EBNA1 could be efficiently degraded. This study showed that a balance between the strength of the degradation signal and the length of the repeat determines the GA repeat-dependent stabilization effect.

Next, we tested the ability of the GA repeat to prevent degradation of the tumor suppressor p53 because inactivation of p53 by accelerated degradation is a common event in tumor development. P53-GA repeat chimeras were protected from degradation and showed improved growth inhibitory activity in tumor cells with impaired endogenous p53 activity, suggesting that insertion of the GA repeat could provide a convenient strategy for the stabilization of potential therapeutic proteins.

We used the aforementioned GFP reporters to test the protective effect of the GA repeat in the yeast Saccharomyces cerevisiae. Expression of proteins carrying GA repeats required the generation of yeast codon-optimized recombinant GA (rGA) repeats. We found that introduction of rGA repeats in the GFP substrates resulted in stabilization of the proteins in mammalian and yeast cells, indicating that the protective signal targets a conserved mechanism in the ubiquitin-proteasome system.

The yeast DNA-repair protein Rad23 is long-lived despite the fact that it is ubiquitinated and interacts with the proteasome. We investigated whether Rad23 contains domains that can protect it from proteasomal degradation. Disruption of the UBA2 domain converted Rad23 into a short-lived protein that is targeted for proteasomal degradation through its ubiquitin-like domain. Insertion of the UBA2 domain from Rad23 or its human homologue HHR23A prevented the degradation of destabilized GFP reporters without causing a general inhibition of proteolysis. We suggest that the Rad23 UBA2 domain functions as a novel cis-acting stabilization signal that confers protection against proteasomal degradation.
LIST OF PUBLICATIONS

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


III  Stijn Heessen*, Nico P. Dantuma*, Peter Tessarz, Marianne Jellne and Maria G. Masucci. Inhibition of ubiquitin/proteasome-dependent proteolysis in *Saccharomyces cerevisiae* by a Gly-Ala repeat. *Submitted*

IV  Stijn Heessen, Maria G. Masucci and Nico P. Dantuma. The UBA2 domain functions as an intrinsic stabilization signal that protects Rad23 from proteasomal degradation. *Manuscript*

* These authors contributed equally to the work
ABBREVIATIONS

APC  anaphase promoting complex
ATP  adenosine triphosphate
CMV  cytomegalovirus
CP   core particle
CTL  cytotoxic T lymphocyte
CUE  coupling of ubiquitin conjugation to ER degradation
DUB  deubiquitination enzyme
E1   ubiquitin activating enzyme
E2   ubiquitin conjugating enzyme
E3   ubiquitin ligase
E6-AP E6-associated protein
EBNA EBV nuclear antigen
EBV  Epstein- Barr virus
ER   endoplasmatic reticulum
ERAD ER-associated degradation
GAr  glycine-alanine repeat
GFP  green fluorescent protein
GRR  glycine-rich region
HIV  human immunodeficiency virus
HPV  human papilloma virus
HSV  herpes simplex virus
IFN  interferon
IxB  inhibitor of NFkB
IKK  IxB kinase
KSHV Kaposi sarcoma-associated herpes virus
LMP  latent membrane protein
MHC  major histocompatibility
Mdm2 mouse double minute 2
NFkB nuclear factor-κB
RP   regulatory particle
RUB  related to ubiquitin
SCF  Skp1/Cull1/F-box protein
SUMO small ubiquitin-related modifier
UBA  ubiquitin-associated
Ubb+1 ubiquitin B gene with +1 frame shift
Ubl  ubiquitin-like protein
UBP  ubiquitin-specific processing protease
UCH  ubiquitin C-terminal hydrolase
UDP  ubiquitin-like domain protein
UFD  ubiquitin fusion degradation
UIM  ubiquitin interacting motif
TAP  transporter associated with antigen presentation
XPC  xeroderma pigmentosum group C protein
# TABLE OF CONTENTS

ABSTRACT
LIST OF PUBLICATIONS
ABBREVIATIONS

## 1. AIMS OF THE STUDY

## 2. THE UBIQUITIN-PROTEASOME SYSTEM

### 3.1 UBIQUITIN
### 3.2 UBIQUITINATION
### 3.3 DEUBIQUITINATION
### 3.4 FUNCTIONS OF UBIQUITIN
### 3.5 UBIQUITIN-BINDING MOTIFS
### 3.6 UBIQUITIN-LIKE PROTEINS

## 4. THE PROTEASOME: A GIANT PROTEASE

### 4.1 GENERAL INTRODUCTION
### 4.2 THE 20S PROTEASOME
### 4.3 THE 19S REGULATORY PARTICLE
### 4.4 THE IMMUNOPROTEASOME
### 4.5 PROTEASOME ASSEMBLY AND LOCALIZATION
### 4.5 NON-PROTEOLYTIC FUNCTIONS OF THE PROTEASOME

## 5. PROTEIN MOTIFS INVOLVED IN UBIQUITIN/PROTEASOME-DEPENDENT PROTEOLYSIS

### 5.1 DEGRONS
### 5.2 STABILIZATION SIGNALS

## 6. THE UBIQUITIN-PROTEASOME SYSTEM AND VIRAL IMMUNE EVASION

### 6.1 ANTIGEN PRESENTATION
### 6.2 VIRAL INTERFERENCE WITH ANTIGEN PRESENTATION

## 7. THE EBSTEIN-BARR VIRUS GLYCINE-ALANINE REPEAT

### 7.1 EBSTEIN-BARR VIRUS
### 7.2 LATENCY PROGRAMS
### 7.3 EBNA1
### 7.4 THE GA REPEAT
### 7.5 INHIBITION OF UBIQUITIN-MEDIATED DEGRADATION BY THE GA REPEAT
### 7.6 BIOLOGICAL SIGNIFICANCE OF THE GA REPEAT
8. THE NUCLEOTIDE EXCISION REPAIR PROTEIN RAD23 56

8.1 Introduction 56
8.2 Functions of Rad23 56

9. GFP-BASED REPORTERS FOR MONITORING THE UBIQUITIN-PROTEASOME SYSTEM 60

10. RESULTS AND DISCUSSION 62

10.1 Length-dependent effect of the GA repeat 62
10.2 Functional stabilization of p53 63
10.3 Reconstitution of the GA repeat effect in yeast 65
10.4 The Rad23 stabilization signal 67

11. FUTURE PROSPECTS 70

ACKNOWLEDGEMENTS 73

REFERENCES 76

APPENDIX (Papers I-IV)
1. AIMS OF THE STUDY

The general aim of the studies presented in this thesis was to increase our understanding of the parameters that determine ubiquitin-dependent proteasomal degradation. In particular, we aimed to study elements that may delay or inhibit degradation.

The specific aims were to:

I  Evaluate the effect of the Epstein-Barr virus (EBV)-derived glycine-alanine (GA) repeat sequence on the ubiquitin-dependent proteasomal degradation of green fluorescent protein (GFP)-based reporter substrates.

II  Apply the inhibitory effect on proteasomal degradation of the EBV-derived GA repeat to functionally stabilize the tumor suppressor protein p53.

III  Study and functionally reconstitute the effect of the EBV GA repeat in the yeast Saccharomyces cerevisiae.

IV  Study whether the yeast DNA-repair protein Rad23 contains domains that can protect it from proteasomal degradation.
2. THE UBIQUITIN-PROTEASOME SYSTEM

One of the factors that determine the biological activity of proteins is their concentration, which in turn is controlled by a balance between their synthesis and degradation rates. Studies on protein turnover have shown that the half-life of proteins can range from seconds to several hours or even days, suggesting that protein breakdown is a dynamic process. The enzymatic cascade that is responsible for most cellular non-lysosomal protein degradation, the ubiquitin-proteasome system, will be the main focus of this thesis.

After the observation that the degradation of cellular proteins is an energy-dependent process (132), it was first shown using biochemical fractionation studies in the 1970s and 1980s that protein degradation required the 76-residue protein ubiquitin (58, 130, 362). Ubiquitin could be conjugated to proteins and form polymers in a sequential reaction that was dubbed ‘ubiquitination’. Hershko and Ciechanover were the first to identify the key enzymes required for this process (131), including a ubiquitin-activating enzyme (E1), ubiquitin-conjugating enzymes (E2) and ubiquitin ligases (E3). Proteins tagged with ubiquitin-polymers or poly-ubiquitin chains are subsequently recognized by a giant cellular protease, the proteasome, which upon deubiquitination and unfolding, degrades the targeted polypeptide in a processive manner (Figure 1). Apart from its evident role in targeting proteins for degradation, ubiquitin has been shown to regulate endocytosis, intracellular sorting of proteins, viral budding, DNA repair, ribosomal function, transcriptional activation and kinase activation (reviewed in 103).

The proteasome or prosome, as it was originally named in 1968, is the responsible protease for most non-lysosomal intracellular protein degradation, highlighting its central role in many cellular processes (7). The proteasome is a self-compartmentalized multi-subunit protease, which displays three main proteolytic activities, divided among six active sites that are shielded from the surrounding cyto- and nucleoplasm, thereby safeguarding the cell against unregulated and non-selective
proteolysis (reviewed in 347). Over the years, a number of rather specific proteasome inhibitors have been discovered, which have proven to be valuable tools for biologists.

![Diagram of the ubiquitin-proteasome system](image)

**Figure 1.** The ubiquitin-proteasome system. **Ubiquitination** of a lysine residue within a target substrate is carried out by the sequential action of a ubiquitin-activating enzyme E1, a ubiquitin-conjugating enzyme E2, a ubiquitin ligase E3 and sometimes a chain elongation factor E4. Successive conjugation rounds result in the formation of a poly-ubiquitin chain that can be recognized by the proteasome. Binding of the poly-ubiquitinated substrate to the 26S proteasome complex results in **deubiquitination**, **unfolding** and processive **degradation** of the substrate. While most short peptide products of the proteasome are further processed by cytosolic peptidases to single amino acid residues, some can serve as ligands in **MHC class I antigen presentation** pathway. Antigens generally require further N-terminal trimming. See text for further details.

Regulated ubiquitin/proteasome-dependent degradation is currently understood to be a central regulatory mechanism in all eukaryotes. Although prokaryotes also contain proteasomes or proteasome-like structures, ubiquitination seems to have appeared later in evolution.
Through targeting proteins for degradation by the proteasome, ubiquitin directs for example critical cell cycle transitions, the activity of regulatory proteins and the response to extra cellular signals. Because the ubiquitin-proteasome system controls the abundance of cell cycle regulatory proteins, apoptotic and anti-apoptotic proteins, aberrant activity of the system further masterminds crucial transformations in cancer development.

An important part of the adaptive immune response in higher eukaryotes, is the major histocompatibility complex (MHC) class I antigen presentation pathway (reviewed in 277). While self-peptides are neglected, peptide fragments derived from foreign endogenously expressed proteins trigger CD8+ cytotoxic T-lymphocyte (CTL) immune responses. Protein fragments produced by ubiquitin-dependent proteolysis are a major source of MHC I peptides and viruses have evolved numerous mechanisms to interfere with the antigen presentation pathway (reviewed in 255, 330, 373), including inhibition of viral protein degradation (196, 197).

Some of the protein motifs that target proteins for ubiquitination and degradation have been studied and understood in detail (188). These so-called degrons come in various shapes and sizes but all share the ability to recruit the ubiquitination machinery. It was recently hypothesized that degrons might not be the only signals that determine protein turnover and proteins might contain additional opposing signals that interfere with proteolysis in a constitutive or conditional manner (67). By counteracting degradation, such stabilization signals could protect certain proteins from processing by the proteasome and may thus form an additional level of regulation within the already complex world of protein degradation.

The work described in this thesis deals with the characterization of a virus-derived stabilization signal that by interfering with the degradation of a crucial viral protein forms part of a complex immune evasion strategy. Additionally, this thesis describes the identification of the first cellular stabilization signal, which is derived from a protein involved in DNA repair. Before discussing the experiments and results that are
described in the original papers, I will first review the regulatory mechanisms that govern intracellular protein degradation.
3. UBIQUITIN AND UBIQUITINATION

3.1 Ubiquitin
Ubiquitin is an evolutionary conserved 76 amino acids long protein that, apart from its key role in ‘garbage disposal’ of abnormal or damaged proteins, fulfils functions in a wide variety of cellular processes through its covalent conjugation to other intracellular proteins (reviewed in 129, 252, 354). Most substrates appear to be marked with polymers of ubiquitin, which is associated with proteolytic and non-proteolytic functions, depending on the conformation of the poly-ubiquitin chain (Figure 2). Other substrates are tagged with just one ubiquitin, better known as mono-ubiquitination, which is only involved in non-proteolytic processes (reviewed in 136).

Ubiquitin is expressed in two forms: as a multimeric ubiquitin precursor (89) or fused to ribosomal proteins (88). Both forms are post-translationally processed by ubiquitin C-terminal hydrolases (UCH) that upon cleavage after ubiquitin’s C-terminal Gly residue, release the functional ubiquitin moieties. Ubiquitin is heat-stable and a poor proteolytic substrate, hence allowing continuous recycling.

3.2 Ubiquitination
Ubiquitination, is the reaction that results in the formation of a covalent amide bond between the C-terminal Gly residue of ubiquitin and the ε-amino group of a Lys residue in the substrate (reviewed in 117, 129, 252). Upon attachment of the first substrate-linked ubiquitin, a Lys residue within this first ubiquitin moiety can become the acceptor for a second ubiquitination round. Whereas ubiquitin itself contains seven lysines that can theoretically be ubiquitinated, only Lys\textsuperscript{48}, Lys\textsuperscript{63} and Lys\textsuperscript{65} have been shown to serve as ubiquitination sites \textit{in vivo} (Figure 2). Repetitive ubiquitination reactions form a substrate-linked poly-ubiquitin
chain, whose conformation dictates its function (see also below and Figure 2).

![Diagram of ubiquitin with lysine residues](image)

**Figure 2. Outline and structure of ubiquitin.** A) Ubiquitin contains 7 lysine residues (Lys⁶, Lys¹¹, Lys²⁷, Lys²⁹, Lys³³, Lys⁴⁸ and Lys⁶³). The cellular processes associated with Lys²⁹, Lys⁴⁸ and Lys⁶³-linked ubiquitin chains are listed. B) Crystal structure of human ubiquitin (PDB file number: 1TBE). The image was generated with Swiss-PdbViewer 3.7. Residues Lys²⁹, Lys⁴⁸ and Lys⁶³ are highlighted.

Ubiquitination requires the sequential action of at least three enzymes: a ubiquitin activating enzyme (E1), a ubiquitin conjugating enzyme (E2) and a ubiquitin ligase (E3). In some cases an additional chain elongation factor (E4) is necessary to obtain the required poly-ubiquitin chain length (123). Cells typically express one E1, a few E2s and multiple families of E3s
(reviewed in 252). Activation of ubiquitin by the E1 is a two-step process (reviewed in 252). Upon binding of ATP and ubiquitin, the E1 first forms a ubiquitin adenylate intermediate that is transferred to the active site cysteine of the E1 as a thiolester. As such, one E1 molecule can bind two ubiquitin moieties: one in the intermediate adenylate form and one as thiolester. The E1 thiol-linked ubiquitin is subsequently transferred to the active site cysteine of E2 ubiquitin-conjugating enzyme to form another thiolester. E2 enzymes weakly interact with free ubiquitin or the E1, but tightly to ubiquitin-bound E1. Some E2s like Ubc4 can interact with a number of different E3s, while vice versa some E3s may interact with several E2s (reviewed in 103, 129).

The E3 enzymes are responsible for selection of the target substrate and provide the necessary selectivity within the ubiquitination reaction (252). Based on structural features, the E3 family can be divided into two groups: the HECT (homologues to E6-AP C-terminus) domain E3s and the RING (really interesting new gene) finger domain E3s. The RING domain group is itself comprised of classic RING fingers E3s and U-box E3s. RING fingers are defined by a pattern of conserved Cys and His residues that coordinate two zinc ions (93). Although all E3s share the ability to ligate activated ubiquitin moieties to substrates, their modes of action are quite different. For example, HECT domain E3s require the transfer of the activated ubiquitin from the E2 to an internal active cysteine residue followed by conjugation to the substrate’s lysine residue, whereas RING finger domain ligases do not require this intermediate step and rather aid the E2 to transfer its ubiquitin directly to the target (156). In addition, RING finger E3s are susceptible to auto-ubiquitination, although the regulatory function of that process is poorly understood. Another interesting feature of the RING finger domain E3s is their ability to form large multimeric complexes, with the anaphase promoting complex (APC) and the Skp1/Cul1/F-box (SCF) ligases as prime examples (336). The activity of the APC complex is primarily focussed towards cell cycle regulatory proteins, while SCF-mediated ubiquitin ligation has been shown to direct degradation of cell cycle-associated proteins and certain transcription factors and their inhibitors (see also Chapter 5). The SCF
ligase F-box protein functions as adaptor between the substrate and the ligase. A number of different F-box proteins have been identified and by alternating the F-box protein, the specificity of the SCF-mediated ubiquitin ligation can be changed (249). Typically, F-box proteins only recognize the phosphorylated forms of substrates, thereby linking kinase-based signalling networks to the ubiquitin system.

Although structurally related to the RING domain E3s, the proteins in the U-box ligase family contain a unique domain that is necessary for their E3 activity (123). The S. cerevisiae Ufd2 protein and the C-terminus of the Hsc70 interacting protein (CHIP) are members of this family. Ufd2 is required for the formation of poly-ubiquitin chains on ubiquitin fusion degradation (UFD) substrates (175) and CHIP ubiquitinates a Parkinson’s disease-associated substrate (152). Ufd2 and CHIP both require the cooperation of a putative E3 and are mainly involved in the elongation of short ubiquitin chains rather than in the initiation of ubiquitination. To distinguish these U-box ligases from the canonical ubiquitin ligases, they are also known as E4s. However, six other U-box ligases were shown to mediate poly-ubiquitination in the absence of an E3, suggesting that E4 activity may instead reflect a specialized type of E3 activity (124).

3.3 Deubiquitination

Like phosphorylation, ubiquitination is a reversible process that is intimately linked with a deconjugation reaction, which is carried out by so-called deubiquitination enzymes (DUBs). These enzymes are traditionally divided in two major families comprised of the ubiquitin processing proteases (UBPs) or ubiquitin specific proteases (USPs) and the aforementioned UCHs (reviewed in 64, 361). Recently, Evans and co-workers described a third DUB family member belonging to the ovarian tumor protein superfamily, which displays structural similarity to cysteine proteases (80). While most DUBs are in fact cysteine proteases that can be inhibited by ubiquitin-aldehydes, recent studies have revealed that a
metalloprotease activity located in the lid of the proteasome complex is also involved in deubiquitination (344, 369).

UBPs contain conserved active site cysteine, aspartic acid and histidine residues and variable N- and C-terminal regions that are important for protein-protein interactions and targeting to specific subcellular locations (reviewed in 64, 361). Most UBPs contain multiple ubiquitin binding sites and are generally involved in cleavage of substrate-ubiquitin isopeptide bonds or between ubiquitin moieties within a poly-ubiquitin chain. UCHs only have a single ubiquitin-binding site and are thought to preferentially separate ubiquitin from ubiquitin-precursors and ubiquitin-fusion proteins, and potentially from small adducts like peptides and amino acids.

The family of gene products that are involved in the ubiquitin-proteasome system contains a large number of DUBs and deubiquitination is likely to play essential roles in diverse cellular processes. Still, little is known about their exact substrate preference and biological relevance and as such deubiquitination represents an exciting and largely unexplored field of research. Indeed, only some functions have been assigned to individual DUBs. For example, USP7/HAUSP can deubiquitinate and stabilize the tumor suppressor p53, resulting in the induction of p53-dependent cell growth repression and apoptosis (199). Other DUBs that regulate gene transcription include UBP3 (228) and UBP10 (DOT4) (161), which are both implicated in gene silencing in yeast. The Drosophila Faf protein is a large UBP and loss-of-function leads to abnormal development of photoreceptors (148). Faf was recently shown to control deubiquitination of the Drosophila homologue of the endocytotic epsin protein (53). The Faf mouse homologue Fam regulates β-catenin (248, 323) and AF-6 levels (324) and can functionally substitute Faf mutants, suggesting a conserved function. Other DUBs are associated with cell growth. For example, USP8/UPBY (237) and USP16/UPB-M (43) have been implicated in direct regulation of the cell cycle, while the UCH BAP1 inhibits cell growth in conjunction with BRCA1 (154). Deubiquitination is also involved in the secretory pathway in yeast. Specifically, Ubp3 and the cofactor Bre5 form
a deubiquitination complex that can protect Sec23 from proteasomal degradation (60).

3.4 Functions of ubiquitin

Originally, ubiquitination was solely believed to be associated with degradation of intracellular polypeptides because poly-ubiquitin chains of minimally four ubiquitins efficiently interact with the proteasome (327). However, it has become clear that ubiquitin is a multifunctional protein, whose covalent conjugation to protein substrates can have very different outcomes. The composition of the ubiquitin tree appears to be a key determinant of the signal that it will provide (Figure 2). Proteins destined for degradation are generally conjugated to Lys\(^{48}\)-linked poly-ubiquitin chains, while only certain UFD substrates are targeted through chains linked to Lys\(^{59}\) in their N-terminal ubiquitin moiety. Ubiquitin chains linked through Lys\(^{63}\) have also been reported and play non-proteolytic roles in several processes including endocytosis (reviewed in 134, 311), DNA repair (142, 307) and activation of the IxB kinase complex by proinflammatoy stimuli (72). Although Lys\(^{5}\)- and Lys\(^{11}\)-linked chains can be generated in vitro, they have not been found in vivo (11). It is not known if under physiological circumstances all ubiquitins within a poly-ubiquitin chain are linked through the same Lys residue or whether some ubiquitins display distinct conformations. The structure of Lys\(^{48}\)-linked poly-ubiquitin chains has been solved (62). Based on the position of the different lysine residues within ubiquitin, it is possible that Lys\(^{29}\)-, and Lys\(^{63}\)-linked chains display a strikingly different structure, supporting the notion that the conformation of a poly-ubiquitin chain may be an important determinant of its biological function (Figure 2).

Apart from the ability to form ubiquitin chains by the sequential ligation of several ubiquitins, ubiquitin can also be ligated as a single moiety. Mono-ubiquitination does not target for degradation, mainly due to its inability to interact with the proteasome (327). Instead, mono-ubiquitination has been mainly associated with membrane trafficking
events as exemplified by its role in endocytosis (reviewed in 136), viral budding (reviewed in 135) and protein sorting in the late endosome (reviewed in 136). Another function of mono-ubiquitination is the regulation of transcriptional activators and histones. The latter modification is required for meiosis in yeast (275) and embryonic development in *Drosophila melanogaster* (250). Apparently, the attachment of a single ubiquitin unit can have various outcomes. Specificity in signalling by mono-ubiquitination may therefore rely on the relevant signal-recognizing components and their cellular localization.

### 3.5 Ubiquitin-binding motifs

To date, four domains have been identified that are capable of interacting with ubiquitin: the ubiquitin-associated (UBA) domain (140), the ubiquitin-interacting motif (UIM) (141), the coupling of ubiquitin conjugation to ER degradation (CUE) domain (303) and a putative zinc finger domain called NZF (223).

UBA domains are evolutionary conserved motifs of about 40 amino acids. They are present in proteins linked to the ubiquitin system and were originally identified in a genome sequence analysis (140). A number of studies have probed into the possible functions of UBA domains. Based on their three dimensional nuclear magnetic resonance structures it was suggested that UBA domains might form a common surface for protein-protein interactions (229). Indeed, UBA domains were shown to interact with ubiquitin and poly-ubiquitin chains (27, 51, 95, 359) and with several other proteins (reviewed in 215) (see also Chapter 8). Binding to ubiquitin has been shown to inhibit both elongation and disassembly of ubiquitin chains (51, 121, 246, 262).

The UIM or LALAL motif is a stretch of around 20 residues that was identified using a sequence analysis with the ubiquitin-interacting motif of the S5a/Rpn10 proteasome subunit as template (141). S5a/Rpn10 preferentially binds ubiquitin chains (377) and certain proteins containing a ubiquitin-like domain (see section 3.5) (349). In contrast to the
SSa/Rpn10 UIMs, the UIM motifs found in endocytotic proteins such as epsin (302), Eps15 (256), Vps27/Hrs (302) bind mono-ubiquitinated proteins. Interestingly, the Eps15 UIMs also play an essential role in ubiquitination of Eps15 itself (172).

CUE domains are comprised of around 50 residues and have so far been found in 50 proteins (18, 257). Originally, CUE domains were named after the yeast Cue1 protein that is essential in the Ubc7-mediated ubiquitin-conjugation of misfolded proteins (28). The sorting protein Vsp9 contains a CUE domain that, besides binding to mono-ubiquitin, plays a role in the endocytosis of two mating factor receptors (69) and, like the Esp15 UIMs, in ubiquitination of Vsp9 itself (303). Other CUE domains also interact with mono-ubiquitin although with different affinities and even binding to poly-ubiquitin has been observed in vitro (303). Two independent structural analyses recently probed into the mode of ubiquitin binding by CUE domains. Prag et al. showed that dimerization of the Vsp9 CUE domain allows a high-affinity interaction with a ubiquitin monomer (259), while Kang and co-workers, based on the structure of the CUE domain of the Cue2 protein with ubiquitin, suggested that its two CUE domains interact with ubiquitin independently and in a non-cooperative manner (163). However, neither study appears to provide a plausible mechanism for selective recognition of mono-ubiquitin, because in the complex the ubiquitin Lys$^{48}$ side chain is pointing towards the solvent and is therefore not necessarily blocked from further chain formation.

The zinc finger domain NZF was originally identified in the mammalian adaptor protein Ufd1-Np14, which is involved in the ER-associated degradation (ERAD) pathway by modifying the function of the p97/Cdc48 ATPase (reviewed in 20). The structure of the Np14 NZF was recently solved (350). Using GST-fused Ufd1 and Np14, Meyer and co-workers showed in in vitro binding assays with HeLa cell extracts that the C-terminal NZF of Np14 could interact with ubiquitinated proteins, without blocking further poly-ubiquitination (223). Because UBA can inhibit ubiquitin-chain elongation (see above), it was suggested that NZF might have an alternative interaction with the ubiquitinated substrate.
Interestingly, it was recently reported that deletion of the Np14 NZF domain did not completely abolish in vitro binding to poly-ubiquitin chains, suggesting that other parts of the Ufd1-Np14 may contain an unidentified poly-ubiquitin binding domain (370).

The identification of several domains and proteins that interact with ubiquitin and possibly ubiquitinated substrates could at least partially explain how ubiquitination can play simultaneous roles in a variety of cellular processes. It is also important to note that multiple copies of ubiquitin binding domains can be found within some proteins. For example, S5a and Eps15 both contain two UIMs and the S. cerevisiae Rad23 and its human homologues HHR23A and HHR23B contain two UBA domains. How this relates to their function and whether there are functional differences between similar ubiquitin-interacting domains within the same protein, are questions that remain to be answered.

3.6 Ubiquitin-like proteins

Proteins that resemble ubiquitin can be divided in two main groups: small ubiquitin-like modifiers (Ubls) and ubiquitin-domain proteins (UBDs). The Ubls function analogous to ubiquitin through their capacity to become covalently conjugated to lysine residues within target proteins in a process strongly resembling ubiquitination (reviewed in 138). A number of Ubls have been discovered of which SUMO (small-ubiquitin-related modifier), RUB1 (related-to-ubiquitin 1) and Apg12 have been studied in most detail (reviewed in 138, 155). SUMO is ligated to lysine residues within a consensus sequence in a variety of target proteins. Like ubiquitin, addition of SUMO influences many different biological processes, such as: the ability of proteins to interact with their partners, patterns of subcellular localization, protein stability, transcriptional regulation and DNA repair (reviewed in 155, 231). Modification by Ubls is generally characterized by the attachment of a monomer, although SUMO polymers have been described in vitro (251, 322). RUB1 can modify the cullin subunits of SCF ligases, which has been shown to be required for full activation of the E3
towards IκBα (139, 266). Two other Ubis, APG12 and the recently
discovered AUT7/APG8 are essential for membrane dynamics during
autophagy (150, 226, 243).

UBD proteins typically contain an N-terminal ubiquitin-like domain
but constitute a structurally and functionally heterogeneous protein
family, and are not found conjugated to other proteins (reviewed in 42,
155). UBD proteins like the yeast Rad23 (78, 290), Dsk2 (78, 95) and the
human BAG-1 (210), hPLIC proteins (173) and the ubiquitin ligase Parkin
(285) all share the ability to interact with the proteasome through their
ubiquitin-like domain (see also Chapter 4 and 8). Interestingly, in the
case of Parkin, the ubiquitin-like domain also controls its steady state
levels, suggesting that such domains may also be directly involved in the
degradation of their native protein (91).

A second group of UBD proteins contains the UBX domain that is
typically located at the C-terminus (reviewed in 42). Despite a lack of
significant sequence homology, UBX domains share a close structural
relationship with ubiquitin. Some UBX proteins even contain N-terminal
UBA domains, providing another link between UBX domain proteins and
ubiquitin.
4. THE PROTEASOME: A GIANT PROTEASE

4.1 General introduction
The 26S proteasome holoenzyme is an ATP-dependent multi-subunit protease generally comprised of two subcomplexes: a 20S core particle and two regulatory particles, the 19S caps (reviewed in 19, 347). In higher eukaryotes the 19S caps can be replaced by the interferon (IFN)-γ inducible PA28 (11S) particles (reviewed in 277). The proteolytic sites are sequestered inside the hollow 20S core particle and their activation is delayed until assembly of the 20S complex is complete. The 19S subcomplexes are thought to be responsible for recruitment, deubiquitination, unfolding and translocation of substrates into the proteolytic chamber, providing a direct link with the ubiquitin system. Although genetic, structural and biochemical analyses have elucidated a great deal regarding the enzymatic activities confined in the 20S proteasome, the functions of the individual subunits of the 19S caps are less well understood.

4.2 The 20S proteasome
The eukaryotic 20S proteasome is an essential complex, highlighting its role in every basic intracellular pathway (reviewed in 347). The 20S core particle consists of 28 subunits, 14α and 14β, that are arranged as a stack of four rings each containing seven subunits, which together form a typical hollow barrel-shaped structure with a α₃-β₁₂-β₁-α₁-γ stoichiometry (reviewed in 19). In the yeast proteasome the N-terminal sequences of the α-rings obstruct direct access to the inner proteolytic cavity by covering its narrow entrance (113). While both α- and β-subunits share the same fold, only three β-subunits (β1, β2 and β5) display proteolytic activity. Upon complete assembly of the 20S particle, the N-terminal
extensions of the three β-subunits are autocatalytically removed to allow the exposure of the N-terminal threonine as active-site nucleophile (52). The catalytic β-subunits each exhibit an individual cleaving preference, with β1 cleaving after acidic amino acids (post-glutamyl peptide hydrolysing (PGPH) or caspase-like activity), β2 cleaving after basic amino acids (trypsin-like activity) and β5 cleaving after hydrophobic residues (chymotrypsin-like activity) (6). Additionally, the 20S proteasome displays branched-chain amino acid preferring (BrAAP) and small neutral amino acid preferring (SNAAP) activities (245). A variety of reversible and irreversible proteasome inhibitors have been identified (reviewed in 35, 171). Most inhibitors mainly target the chymotrypsin-like activity, which can be diminished up to 80% without apparent effects on bulk protein turnover, indicating a great redundancy within the system. Based on their anti-tumor activity in animal models, proteasome inhibitors may exhibit a clear therapeutic future (119, 222). For example, the recently FDA-approved chymotrypsin activity-specific proteasome inhibitor, PS-341, has clear therapeutic potential in treatment of multiple myelomas (97).

The peptide fragments that are generated by the proteasome vary in size between 3 and 22 residues with a median of 6 (169, 170), suggesting that most peptides escape prior to complete digestion. It is not understood how peptide fragments leave the proteolytic chamber, although peptide escape is restricted if the entrance channel is in a closed state, suggesting that entrance and exit may occur at the same opening (176). Alternatively, small peptides may escape through the side-openings in the cylindrical walls of the core particle (114). Fragments that exit the proteasome can be further processed by several cytoplasmic amino- and endoproteases (reviewed in 288). In organisms possessing an adaptive immune system, a minor fraction of these peptide fragments will be loaded on MHC class I molecules (reviewed in 277) (see also Chapter 6).

In vitro, purified eukaryotic 20S proteasomes demonstrate poor proteolytic activity towards tightly folded and ubiquitinated substrates. Only certain unfolded polypeptides and short fluorogenic substrates are efficiently digested upon activation of the complex by exposure to low
levels of sodium dodecylsulfate. It is thus reasonable to assume that, to allow its function as the central protease in the degradation of ubiquitinated proteins, in vivo most 20S particles are capped at one end or sandwiched at both ends by regulatory complexes.

Figure 3. Subunit interaction map of the human 19S regulatory particle. The yeast/human nomenclature is used. The depicted interactions are based on genetical and protein-protein interaction studies. See Table I for information on non-subunit proteins that interact with the 19S complex. Adapted from Ferrell et al. (87).
<table>
<thead>
<tr>
<th>Yeast</th>
<th>Human</th>
<th>Cellular interacting proteins</th>
</tr>
</thead>
<tbody>
<tr>
<td>RPN1</td>
<td>S2</td>
<td>Rad23 (78, 284); Dsk2 (78, 284); TNFR (36, 77); Ste6 (205); KIAA10 (376)</td>
</tr>
<tr>
<td>RPN2</td>
<td>S1</td>
<td>bRANBP-2 (86); Gpa1 (367); Rad23 (284); Dsk2 (284); Ubr1 (365)</td>
</tr>
<tr>
<td>RPN3</td>
<td>S3</td>
<td>CDC28 (162); CKS (162)</td>
</tr>
<tr>
<td>RPN4</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>RPT1</td>
<td>S7</td>
<td>S5b (109); HEC (54); Yme1 (44); Ubr1 (365); Nob1 (328)</td>
</tr>
<tr>
<td>RPT2</td>
<td>S4</td>
<td>S5b (109), HPV E7 (25, 39); ataxin-7 (220); AdE1A (335)</td>
</tr>
<tr>
<td>RPT3</td>
<td>S6</td>
<td>Centractin (272); MB67 (56); p28 (70)</td>
</tr>
<tr>
<td>RPT4</td>
<td>S10b</td>
<td>Gal4 (47); Dynamitin (272); p27 (71); Ufd4 (366)</td>
</tr>
<tr>
<td>RPT5</td>
<td>S6'</td>
<td>HIV-1 tat (239); VHL (334); TBPIP (321); TBP (315); p27 (71); poly-Ub (185)</td>
</tr>
<tr>
<td>RPT6</td>
<td>S8</td>
<td>Gal4 (190, 317, 318); HEC (54); TR (190); Phd (16, 383); c-Fos (351); XBP (353); VP16 (190, 318); AdE1A (110, 335); SV40-T (110); TBBP (318); Ufd4 (366); RARα (348); ERα (348); VDR (348); RXR (190, 348);</td>
</tr>
<tr>
<td>RPN5</td>
<td>p55</td>
<td>-</td>
</tr>
<tr>
<td>RPN6</td>
<td>S9</td>
<td>-</td>
</tr>
<tr>
<td>RPN7</td>
<td>S10a</td>
<td>-</td>
</tr>
<tr>
<td>RPN8</td>
<td>S12</td>
<td>-</td>
</tr>
<tr>
<td>RPN9</td>
<td>S11</td>
<td>-</td>
</tr>
<tr>
<td>RPN10</td>
<td>S5a</td>
<td>E4 (175); poly-Ub (377); HHR23A (137); RyR (214); ID1 (5); MyoD (5); E12 (5)</td>
</tr>
<tr>
<td>RPN11</td>
<td>Poh1/S13</td>
<td>PAP1 (304)</td>
</tr>
<tr>
<td>RPN12</td>
<td>S14</td>
<td>Cdc28 (177); Nob1 (328)</td>
</tr>
</tbody>
</table>

Table I. Non-Subunit interactions with subunits of 19S regulatory particle. The yeast and human nomenclature are used. The listed interactions with cellular proteins are based on genetical and/or protein-protein interaction studies. See Figure 3 for information on subunit-subunit interactions. Additional information is available on www.proteasome.com.
4.3 The 19S Regulatory Particle

Human and yeast 19S regulatory particles (RPs) are comprised of 17 and 18 subunits, respectively, which in turn can be subdivided into a base and a lid subcomplex that are located proximally and distally with respect to the 20S core particle (reviewed in (104, 347) (Figure 3 and Table I)).

The base complex contains eight subunits and using budding yeast/human nomenclature (90) these include a ring of six AAA (ATP-ases associated with a wide variety of cellular activities)-ATPases, Rpt1/S7, Rpt2/S4, Rpt3/S6, Rpt4/S10b, Rpt5/S6', Rpt6/S8, plus the two largest subunits Rpn1/S2 and Rpn2/S1 (Figure 3 and Table I). Despite their sequence similarity the AAA-ATPases are not functionally redundant (105, 280). Collectively, the 19S base has been associated with several regulatory functions, including unfolding (40), translocation (176) and opening of the gate that covers the axial channel of the 20S particle (176). It is important to note that, while a 20S proteasome capped with only the base subcomplex is unable to degrade ubiquitinated substrates, its Rpt5/S6' subunit interacts with poly-ubiquitin chains, suggesting an additional role for the base complex in the recruitment of ubiquitin-conjugated substrates (185).

Another ubiquitin-binding subunit Rpn10/S5a is believed to stabilize the interaction between the base and lid complexes. Rpn10 interacts with several 19S RP subunits and deletion of yeast Rpn10 causes dissociation of the 19S RP into separate base and lid subcomplexes (104). However, in yeast (340, 358), *Drosophila* (120) and *Arabidopsis* (339) Rpn10 is primarily found unassociated with the proteasome, indicating that it may perform important functions outside of the proteasome. Yeast Rpn10 contains one UIM that has a binding preference for chains of minimally four ubiquitins. It is important to note that deletion of Rpn10 is not lethal and only slight defects in protein degradation were observed (340), suggesting that other factors may take over and present ubiquitinated substrates to the proteasome. Such additional factors might include the UBD proteins Rad23 and Dsk2, as deletion of either gene together with deletion of Rpn10 results in strong phenotypes in budding yeast and is lethal in fission yeast (122) (see also Chapter 8). Intriguingly, in HeLa
cells S5a was not found in a free state (127). Moreover, S5a contains two ubiquitin-interacting motifs, suggesting possible functional differences between S5a and its non-human counterparts. Indeed, S5a directly interacts with the ubiquitin-like domains of the human Rad23 homologues (349), while yeast Rpn10 is incapable of binding to Rad23 (78).

The 19S RP lid is minimally comprised of eight subunits: Rpn3/S3, Rpn5, Rpn6/S9, Rpn7/S10a, Rpn8/S12, Rpn9/S11, Rpn11/S13 and Rpn12/S14 (105) (Figure 3 and Table I). Specific functions have only been assigned to a few 19S subunits. For instance, Rpn3 is required for the degradation of the S-phase cyclin Clb5, while mutations in Rpn12 severely inhibit the turnover of the cyclin-kinase inhibitor Sic1 and activation of the Cdc28 kinase (15). Fission yeast Rpn5 is required for proper assembly and localization around the inner nuclear membrane (371) and Rpn6 affects both the structure and the peptidase activity of the proteasome (286), underscoring its essential role in Drosophila development (200). Already in 1993 the 26S proteasome was shown to exhibit an ATP-dependent ubiquitin hydrolase activity that could be blocked by a heavy metal chelator but not by ubiquitin aldehyde (81). Two independent studies recently assigned this activity to a metalloprotease motif found in the Rpn11 lid-subunit (344, 369). Importantly, Rpn11/S13 also contains a putative active-site cysteine that is conserved in other classical DUBs (105), suggesting that it may exhibit a dual deubiquitination activity.

The precise sequence of events that lead to proteolysis of a substrate following its ubiquitination remain elusive. It is possible that not all proteins are degraded similarly and that individual proteins may require accessory proteasome-interacting proteins, that could either aid in targeting to the proteasome or help in the ubiquitination/deubiquitination or unfolding processes. A number of recent studies have indeed challenged the idea that the 26S proteasome solely consists of stochiometric subunits that resist stringent purification. Using a variety of methods several other proteins were identified that form transient interactions with the proteasome, although only some of those binders may be major components of the proteasome complex (191) (Table I).
Apart from Rad23 and other Ubl domain-containing proteins (78), proteins that interact with the proteasome include for instance ubiquitin-conjugating enzymes (329), ubiquitin ligases (191, 365, 366, 376), deubiquitination enzymes (38, 191, 345), adaptor proteins (191) and cell cycle regulators (177). Understanding the functional significance of these binding partners will be an important task.

4.4 The immunoproteasome

In organisms with an adaptive immune system, the three catalytic β-subunits (β1, β2 and β5) of the 20S particle can be replaced by LMP2, MECL1 and LMP7, respectively (96). The resulting proteasome was appropriately named the immunoproteasome, as these substitutions take place upon induction by the immune-regulatory cytokine IFN-γ. A functional consequence of this conversion is an increased output of fragments with hydrophobic C-termini versus acidic C-termini. The notion that MHC class I molecules strongly favour peptide fragments with a hydrophobic C-terminus provides a logical explanation for this observation. Suitable C-termini are indeed commonly generated by the peptidase activity of the proteasome (46), while the appropriate N-terminus can be produced by the ER aminopeptidase associated with antigen processing (ERAAP) (297) (see also Chapter 6).

Besides a change in organization of the catalytic β-subunits, IFN-γ also induces the expression of an alternative proteasome regulatory cap, PA28 (REG or 11S) (reviewed in 277). PA28-20S-19S hybrid proteasomes are very labile and only represented 25% of the total proteasome content in HeLa cells that underwent prolonged IFN-γ treatment (320). Like the 19S cap, PA28 has been suggested to both allow entrance of polypeptides and exit of degradation fragments by opening of the α-ring channel (213, 355). However, in sharp contrast to the 19S particle, PA28 does not contain binding sites for ubiquitin chains and hence does not support proteolysis of ubiquitin-conjugated substrates (309).
PA28 is composed of two homologous subunits PA28α (REG α) and PA28β (REG β) that preferentially form a hexameric α3-β3 ring structure (2). Alternative heptameric α3-β4 or α7 rings have also been observed in vitro, but their physiological relevance is questionable (174, 381). Unlike the 19S cap, association of PA28 with the 20S does not require ATP. Considering that the mean product size of PA28-20S-19S and 19S-20S-19S proteasomes is equal, it has been proposed that PA28 may rather enhance certain cleavages and facilitate the release of a broader range of peptide products, ultimately leading to a different spectrum of potential antigenic peptides (45). This supports the finding that PA28α+/β−/ mice only show defects in the processing of certain antigens (234). It is not understood how PA28 may modify the cleavage properties of the proteasome other than the suggestion that PA28 binding may induce subtle conformational changes within the 20S complex that could change the accessibility of the active sites (45).

While cytoplasmic PA28 is mainly formed by PA28α and PA28β, nuclear PA28 consists of PA28γ (REGγ) subunits that form a homomeric ring (357). Binding of PA28γ enhances the cleavage of peptide bonds following basic residues, but a single K188E mutation converts the cleavage pattern to that of PA28α and PA28β (198). PA28γ-deficient mice display minor cell proliferation defects, suggesting that some PA28γ functions are not rescued by PA28α or PA28β (233), providing further support for the idea that PA28γ may function independently from PA28α and PA28β.

4.5 Proteasome assembly and localization
Proteasome assembly is a slow process that involves intermediate complexes containing one complete α-ring, an incomplete and variable number of unprocessed β-subunits and an assembly chaperone called Ump1 in yeast (263), whose mammalian homologue is referred to as proteasome maturation protein (364) or Proteassemblin (112). Complete assembly occurs upon assemblage of the remaining β-subunits, followed
by docking of two intermediate complexes at the β-ring interface, autocatalytic removal of the β-propetides and degradation of the chaperone (52). How the 19S cap subunits assemble into a complex is largely unknown, apart from the observation that Rpn9 is required for accommodation of Rpn10 into the 26S proteasome (319). Further assembly of pre-assembled 20S and 19S sub-complexes into an active 26S complex requires proteins such as Pno1 (328), Nob1 (328), Ecm29 (191) and Hsp90 (151).

Green fluorescent protein (GFP)-tagged proteasome subunits have been used to study the subcellular localization of 26S proteasomes. In yeast, proteasomes seem to be primarily associated with the endoplasmatic reticulum (ER)/nuclear envelope interface (79, 360). Human 26S proteasomes are found throughout the nuclear and cytoplasmic compartments, and are excluded from the nucleoli and the lumen of the ER (270). While an intact nuclear membrane only allows slow import of assembled proteasomes, they acquire access to the nucleus during mitosis when the nuclear membrane is disintegrated. Fluorescence recovery after photobleaching experiments showed that GFP-tagged proteasomes diffuse rapidly within the nucleus and cytoplasm in living cells, suggesting that they interact with substrates by collision (270).

4.5 Non-proteolytic functions of the proteasome

The role of the ubiquitin-proteasome system in transcriptional regulation has long been appreciated (reviewed in 61). In fact, histone H2A was the first example of an ubiquitin-conjugated protein in vivo (106, 107, 149). Additionally, the activity of many transcription factors is tightly coupled to their proteolytic destruction (reviewed in 235). A number of recent studies have however suggested that the 19S RP may be involved in degradation-independent control of transcription (235) and DNA repair (100, 282).

Studying the yeast activator Gal4, Johnston and co-workers noticed that recessive mutations in the 19S AAA-ATPases Rpt4 and Rpt6 but not
mutations in subunits of the 20S proteasome could suppress the effects of deleting the Gal4 transcriptional activation domain (TAD) (281, 283, 317). Rpt6 physically interacts with the Gal4 TAD (318) and with several other TADs. Based on these observations, it was suggested that the Rpt6 and possibly other components of the 19S might have a direct and proteolysis-independent role in transcription. This view is further supported by the observation that mutations or otherwise inactivation of Rpt4 and Rpt6 alleles exhibits transcription elongation defects in yeast and humans, which can be restored by the addition of purified 19S (84, 85). Importantly, inhibition of the catalytic activity of the 20S has no effect on elongation. Further in vivo evidence of a non-proteolytic role of the 19S complex was provided by elegant chromatin immunoprecipitation assays, which showed that a complex of six 19S AAA-ATPases is recruited to transcriptionally active genes in yeast (108). Because these studies failed to show the presence of any 20S subunits, this assembly was coined API5 (AAA proteins independent of 20S) complex. However, it is not known if such a complex exists within cells or if it may contain other proteins. Based on their observations, Gonzalez et al. suggested a model in which only the initiation of transcription may require the activity of the activator-API5 complex, while the transcriptional termination in turn may depend on the destruction of the activator by the complete 26S proteasome (108). Muratani and Tansey preferred an alternative and API5-independent model, in which recruitment of the 26S proteasome is both necessary for the initiation, elongation and termination of transcription (235).

Another non-proteolytic role for the 19S complex was proposed based on studies on nucleotide excision repair (NER) (100, 282), which repairs base damage in DNA that is caused by ultraviolet (UV) light or exposure to certain chemicals (reviewed in (94)). The Rad23 protein is required for efficient NER in yeast (352). Rad23 interacts with the proteasome through its N-terminal ubiquitin-like domain (78, 290) and deletion of this domain confers UV-sensitivity that is intermediate between that of wild-type cells and rad23 deletion strains (352). Using an in vitro NER system, Russell et al. found that the physical interaction between the
Rad23 ubiquitin-like domain and the 19S proteasome complex is required for optimal NER activity (282). It is noteworthy, that the Rad23-proteasome interaction can also be detected in the absence of active NER. While chemical and antibody-mediated inhibition of the Rpt4 and Rpt6 subunits of the 19S complex caused a decrease in NER efficiency in vitro, blocking the catalytic activity of the 20S proteasome had no effect. This suggests that the 19S RP may indeed have an as yet unidentified stimulatory function in NER independent of the proteolytic activity of the proteasome. However, using a genetic approach, evidence was presented regarding a putative negative role of the 19S complex in NER in vivo (100). First, mutant alleles of Rpt4 and Rpt6 displayed enhanced NER rates, while mutations in 20S subunits had no effect, which is a partial confirmation of an earlier report by Lommel and colleagues (207). Second, temperature-regulated degradation of Rpt4 resulted in a significant increase in NER in living cells. Interestingly, mutations in Rpt4 and Rpt6 selectively suppressed the NER defect in yeast strains expressing Rad23 lacking its ubiquitin-like domain. These observations demonstrated that the negative effect on NER might result from a direct Rad23-independent interaction between the NER machinery and the proteasome.

Although these studies have clearly defined novel non-proteolytic functions for the 19S RP, its mode of action in transcription and DNA repair has not been elucidated. However, because the AAA-ATPase superfamily can perform chaperone-like functions in both assembly and disassembly of protein complexes (40, 240), it is tempting to speculate that such activity may be involved in the degradation-independent roles of the 19S RP.
5. PROTEIN MOTIFS INVOLVED IN UBIQUITIN/PROTEASOME-DEPENDENT PROTEOLYSIS

5.1 Degrons

Recruitment of the appropriate E3 enzyme is believed to be the rate-limiting step in the ubiquitination reaction. The protein motifs that trigger the engagement of a ubiquitin ligase are designated degradation signals (188) or degrons (342), some of which are understood in detail. A common characteristic is their ability to be functionally transferred to other proteins.

A prime example is the PEST-sequence (reviewed in 267) that accommodates a phosphorylation-dependent degron that can be recognized by different ubiquitin ligases, such as the SCF complex and mouse double minute (MDM) 2 (reviewed in 129). Due to its intimate connection with phosphorylation, PEST sequences are mostly found in proteins whose degradation is tightly coupled to environmental status, cell cycle stage or developmental state (188). A degron commonly found in cyclins is the destruction box (368), which is recognized by the APC ubiquitin ligase whose own activity is regulated by phosphorylation (379). Alternative post-translational modifications that trigger ubiquitination and degradation include acetylation (115), prolyl isomerization (378), and hydroxylation (144). Recently, SCF-mediated ubiquitination was also shown to play a role in ERAD of glycoproteins by virtue of the F-box protein Fbx2 that specifically recognizes N-linked glycans (375).

The N-end rule represents another type of degron. According to the N-end rule, the identity of a polypeptide’s N-terminal amino acid mediates interaction with the E3 Ubr1/E3a and subsequent ubiquitination on an internal lysine (12, 13). As such, N-terminal amino acids can be categorized according to their stabilizing or destabilizing property, respectively (reviewed in 341). Artificial N-end rule substrates can be constructed by expressing proteins with an N-terminal ubiquitin moiety.
that is cotranslationally cleaved off by cellular UCHs, thereby exposing the new N-terminal residue. N-end rule-dependent ubiquitination and degradation have recently been shown to control fundamental cellular processes including chromatid separation, peptide uptake, and programmed cell death or apoptosis (181, 264). Three putative N-end rule ubiquitin ligases have been identified in mouse, namely Ubr1, Ubr2 and Ubr3. Mice lacking Ubr1 are viable and only display minor defects, while Ubr1−/−Ubr2−/− mice are embryonic lethal (182).

Johnson et al. first described the UFD degron (157, 158). The presence of an N-terminal ubiquitin moiety, which is insensitive to cleavage by UCHs or other proteases, serves as a potent degron, which was designated UFD signal. Ubiquitination occurs within the ubiquitin domain. The only natural UFD substrate identified so far is the ubiquitin B gene with +1 frame shift (UBB+1) (201). UBB+1 is generated by molecular misreading within the ubiquitin B mRNA, which results in an aberrant protein that is found accumulated in certain neurodegenerative disorders (337, 338). A number of proteins are degraded by the proteasome upon conjugation of a poly-ubiquitin chain to their free amino terminus (8, 41, 268), suggesting that ligation of the first ubiquitin moiety may in fact confer these proteins into substrates for UFD-mediated ubiquitination. It is not understood what triggers ubiquitination at free N-terminal amino groups rather than on internal lysine residues.

The Deg1 degron typifies a conceptually different degron because it relies on the exposure of a hydrophobic protein-protein interaction surface (159). A screen in S. cerevisiae identified several other hydrophobic peptide-degrons (101). Generally, exposure of hydrophobic segments in misfolded or otherwise abnormal protein leads to their rapid elimination by the ubiquitin-proteasome system (reviewed in 118, 178), suggesting that the ubiquitination machinery may indeed efficiently recognize hydrophobic sequences.

It is noteworthy that proteasomal degradation of the cyclin-dependent kinase (Cdk) inhibitor p21 (300) and ornithine decarboxylase (ODC) (232) does not absolutely require ubiquitination. Instead, ODC degradation is mediated by interaction with its partner antizyme. Although
p21 contains six lysine residues that form potential targets for SCF-mediated ubiquitination (37), a C-terminal p21 domain can directly bind to an α-subunit of the proteasome, resulting in the degradation of p21 (331). Generally, the exposure of a hydrophobic surface may result in a direct interaction with the 19S cap of the proteasome (310), thereby bypassing the ubiquitination requirement. This phenomenon was shown to occur for various denatured proteins (22), but the biological significance of those observations awaits further clarification. Interestingly, the ability of p21 to recognize and inhibit multiple cyclin-Cdk complexes derives from its capacity to adopt both disordered and ordered conformations (179), which may be directly related to the ubiquitination requirement for its degradation. In other words, proteasomal degradation of p21 lacking a stable structure may not require ubiquitination because the C-terminal proteasome interacting domain is exposed, while in the ordered conformation this domain is shielded and ubiquitination may be necessary to allow a long enough interaction with the 19S to permit unfolding and degradation.

It is well appreciated that the turnover of proteins is generally determined by the presence of degrons, which target substrates for ubiquitination and subsequent degradation by the proteasome. The large number and diversity of the degrons known to date underscores their crucial role in the ubiquitin-proteasome system.

5.2 Stabilization signals
It has been recently postulated that apart from degrons certain proteins may contain so-called stabilization signals that delay or inhibit their degradation (67). Dantuma and Masucci postulated that stabilization signals might function either by decreasing the rate of poly-ubiquitination, increasing the rate of deubiquitination, modification of substrate-proteasome interaction or interfering with substrate unfolding (67).

The stabilization signal hypothesis was originally based on the observation that the Epstein-Barr virus nuclear antigen 1 (EBNA1)
contains an internal Gly-Ala (GA) repeat sequence that prevented the MHC class I presentation of EBNA1-derived antigenic peptides (196). Because the GA repeat was subsequently shown to interfere with the proteasomal processing of EBNA1 (197), it is the first example of a viral stabilization signal (67). Importantly, the stabilizing effect can be transferred to other natural and artificial proteolytic substrates (reviewed in 68). The mechanism by which the GA repeat inhibits proteasomal degradation is poorly understood (see Chapter 10 and Paper I, II and III).

Another example comes from studies on the partial processing of the NF-κB precursor protein p105 (83, 247). Upon ubiquitination in its C-terminus p105 is targeted for proteasomal degradation but an internal glycine-rich region (GRR) serves as a stop signal, resulting in complete degradation of the C-terminal domain while preserving and releasing the N-terminal p50 subunit (57, 244). The GRR contains a GAGAG sequence reminiscent of the EBV GA repeat but, although the GRR may protect the p50 product from further degradation, it does not obstruct proteasomal degradation when transferred to other proteolytic substrates. Together with the finding that partial processing of NF-κB in yeast cells occurs independent of the GGR, this implies that other domains may be involved (294). The active forms of two latent membrane-bound yeast transcription factors Spt23 and Mga2 are also generated through a partial processing mechanism that requires the ubiquitin-proteasome system (146). Both proteins are structurally related to NF-κB p105 and contain immunoglobulin-like putative DNA binding domains and C-terminal ankyrin repeats. Based on the observation that Spt23 and Mga2 lack the p105 GGR but have several asparagine-rich stretches whose locations correspond to the p105 GGR, it is tempting to speculate that these domains allow the formation of a flexible loop that is required for endoproteolytic processing of the substrate by the proteasome (265). It would be interesting to test whether these asparagine motifs function as transferable elements or if they can be functionally replaced by the p105 GGR.

It has been postulated that tightly folded protein domains may be a type of conformational stabilization signal (67). Using in vitro degradation
assays several labs have probed the inability of the proteasome to efficiently hydrolyse compact domains. Based on these studies it was proposed that unfolded polypeptides are fed to the proteasome from an accessible terminus and as such domains that are resistant to unfolding may serve as a stop signal releasing the remaining peptide chain (189, 238). In fact, the partial processing of p105 is prevented by destabilization of the Rel homology domain in p50, suggesting that the resistance against unfolding of this domain may provide the required stop signal (189). Using fusions of a stable GFP domain and two natively disordered proteasome substrates, it was recently shown that in vitro the proteasome possesses an additional proteolytic activity towards internal peptide bonds, providing a possible mechanistic basis for the endoproteolytic activation of transcription factors and the accessibility of internal folding defects in misfolded proteins (203). This further implies that the proteasome could indeed accommodate an unfolded polypeptide loop. It is noteworthy that tightly folded domains may lead to a general inhibition of the proteasome probably by physically obstructing its narrow entrance channel (see also Chapter 4).

A number of neurodegenerative diseases are characterized by the aggregation of proteins containing expanded polyglutamine or poly(Q) repeats (reviewed in 301). It has been shown that soluble poly(Q) containing proteins are efficiently degraded by the proteasome while they resist proteolysis once aggregated, suggesting that poly(Q) repeats may represent another type of conformational stabilization signal (343). It is not understood whether aggregates are formed as a cellular protection mechanism or whether they represent the early stages of the pathology associated with neurodegeneration. It is important to note that different groups have reported conflicting data on whether overexpression of poly(Q) proteins may result in a general impairment of the ubiquitin-proteasome system. While Bence and coworkers found that poly(Q) aggregates caused some inhibition of the system (23), Verhoef et al. reported that, while aggregated poly(Q) proteins resisted degradation, they observed no obvious effect on the turnover of another proteolytic substrate (343). It remains to be tested if this discrepancy reflects
differences between the used reporters systems, cell lines or experimental set-ups.

It is important to note that no cellular stabilization signals have been identified so far and it remains to be proven that the examples described above indeed function as stabilization signals. However, the possible combination of a stabilization signal and a degron within one protein may allow proteins to become ubiquitinated and targeted to the proteasome without proteolysis as a necessary consequence. Such proteins may be well suited to perform important functions in for instance recruiting proteasome substrates or co-factors that regulate the formation of the proteasome or even facilitate non-proteolytic activities of the proteasome.
6. THE UBIQUITIN-PROTEASOME SYSTEM AND VIRAL IMMUNE EVASION

6.1 Antigen presentation

As part of their adaptive immune system, higher eukaryotes have acquired an antigen presentation pathway that enables them to sample the dynamic contents of their intracellular protein pool by presenting peptide fragments on MHC class I molecules to surveying CD8+ CTLs (reviewed in 126, 372). The classical MHC class I antigen presentation pathway is composed of the following steps: degradation of proteins into peptide fragments, transport of the peptide across the ER membrane, assembly of peptide loaded MHC class I molecules in the ER and subsequent translocation to the cell surface (Figure 4).

The major source of class I antigenic peptides is degradation products derived from the ubiquitin-proteasome-mediated degradation of endogenous proteins (discussed in Chapter 3 and 4) (reviewed in 277). The cytosolic tripeptidyl peptidase II (TPPII) also appears to have the capacity to generate certain peptides for MHC class I molecules (295), although the TPPII-dependent antigen presentation pathway is rather inefficient (166). Exceptionally, even endocytosed material can be presented on MHC class I by a process called cross presentation (276). Most antigenic peptides are derived from newly synthesized proteins that form the major pool of proteasome substrates (271, 293). From the MHC class I perspective, proteasomes and immunoproteasomes (see Chapter 4) mainly produce N-terminally extended peptides whose further trimming by cytosolic- and ER-resident peptidases is necessary to allow efficient binding to class I molecules (46). Cytosolic peptides that do not immediately enter the ER or nuclear compartment are attacked by peptidases (269), including the leucine amino peptidase (LAP) (24), thimet oligopeptidase (TOP) (227) and TPPII (98). Interestingly, TPPII may partially compensate proteasome function in cells that are adjusted
to permanent proteasome inhibition (102), although Princiotta et al. later suggested that it may be the residual proteasome activity in these adapted cells that supports survival (261). Because TPPII activity was recently shown to be required for the generation of critical human immunodeficiency virus (HIV)-1 epitopes irrespective of proteasome inhibition, it was suggested that TPPII may not only function downstream of the proteasome but also in parallel (295).

The presence of peptidases such as LAP, TOP and TPPII provides a partial explanation for the observation that the MHC class I antigen presentation pathway is rather inefficient, with the formation of only one peptide-MHC complex for every $10^2$ to $10^4$ degraded proteins, arguing that escape from complete degradation is the exception rather than the rule (260). How do some peptides escape further degradation by cytosolic peptidases? It may be possible that proteasomes directly associate with the peptide transporter in the ER membrane (see below). Another suggestion comes from studies on heat shock proteins that may shuttle peptides through the cytoplasm, thereby offering protection from potential peptidase activity (30). However, compelling evidence supporting any of these scenarios is lacking. Interestingly fluorescent peptides compete with endogenous peptides for binding to chromatin-associated proteins such as histone H2B and H4, thereby reducing their availability to peptidases (269). It remains to be tested whether chromatin influences delivery of a specific peptide pool or if protection occurs temporarily to assure proper sampling of peptides that are produced during mitosis.

In order to serve as MHC class I ligands, peptides acquire access to the ER by the transporter associated with antigen processing (TAP) (reviewed in 126, 372). TAP is a heterodimer composed of TAP1 and TAP2 subunits, which are members of the ATP-binding cassette transporter family and peptide transport indeed occurs in an ATP-dependent manner (1). Peptides transported by TAP range from 8 to 40 amino acids in length, although shorter peptides are most efficiently transported. This broad size specificity suggests the presence of an ER peptidase that trims elongated peptides. Indeed, such a peptidase, named ERAAP was recently identified (287, 297, 374). ERAAP efficiently trims N-terminally extended
peptides and like other components of the antigen presentation pathway, ERAAP expression is upregulated by IFN-γ. 

Mature class I molecules form a trimeric complex containing a heavy chain, a light chain and a peptide that is embedded in the heavy chain peptide groove (31). The immense allelic variation within the binding groove results in the presentation of a unique set of antigens to specific CTLs and provides a basis for the diversity of individual immune responses. In the ER, MHC class I molecules undergo a series of maturation steps that require interactions with TAP and three ER-localized chaperones calnexin, calreticulin and tapasin (reviewed in 63, 111). Exclusively upon peptide binding, MHC class I molecules are authorized to exit the ER and subsequent transportation through the ER-Golgi intermediate compartment and the Golgi-network finally results in expression at the cell surface, where circulating CD8+ CTLs survey the presented peptides for fragments of viral or pathogenic origin.

6.2 Viral interference with antigen presentation
The presentation of viral peptides by MHC class I molecules triggers CD8+ CTLs to eliminate the infected cell by inducing apoptosis (reviewed in 17, 209, 305). It is well documented that several viruses have evolved elaborate mechanisms to selectively frustrate virtually every step in the antigen presentation pathway, from processing of viral peptides to interfering with the expression of MHC class I molecules (reviewed in 255, 330, 373). Notably, the identification and characterization of viral gene products that interfere with MHC class I presentation have greatly enhanced our understanding of this pathway. The next sections describe some examples of human viral proteins that obstruct class I antigen presentation (Figure 4).

To date, two examples are known of viral proteins that inhibit the generation of virus-derived antigenic peptides. The first example is provided by EBV, which selectively blocks the proteasomal degradation of one of its nuclear proteins, thereby preventing the generation and
presentation of peptides derived from this protein (196, 197) (discussed in detail in Chapter 7). The inhibitory effect locates to an internal stretch of glycines and alanines. Interestingly, the domain can be functionally

![Diagram](image)

**Figure 4.** Viral interference with antigen presentation. Peptide fragments that are products from ubiquitin-mediated proteasomal degradation are translocated to the ER by TAP. In the ER, nascent MHC class I molecules are loaded with antigenic peptides which signals the transport of the loaded class I molecule to the cell surface. Several viral proteins interfere with this process. 1) The EBV GA repeat and the HCMV pp65 protect viral proteins from degradation. 2) HSV ICP47 and HCMV US6 interfere with TAP-dependent peptide transport. 3) HCMV US2 and US11 cause the Sec61-dependent but proteasome independent retrograde translocation of class I molecules to the cytosol resulting in ubiquitination and proteasomal degradation of MHC class I. 4) The adenovirus E19 and HCMV US3 proteins retain peptide-loaded class I molecules in the ER. 5) The HIV-1 and HIV-2 Nef protein accelerates endocytosis and lysosomal degradation of MHC class I complexes. The KSHV K3 and K5 proteins induce ubiquitination of the class I tail and induce targeting to the lysosome. 6) The human pox virus and HCMV prevent NK-cell recognition by expressing decoy MHC class I molecules.

transferred to other proteins, causing the selective inhibition of proteasomal degradation of the new host protein (reviewed in 68). A
conceptually different example comes from studies on CTL responses against infection with the human cytomegalovirus (HCMV). During the early stages of infection, simultaneous expression of the kinase phosphoprotein 65 (pp65) and a 72 kDa transcription factor interferes with the generation of peptides from the 72 kDa protein that otherwise provokes a strong CTL response (99). Whether pp65 phosphorylates the transcription factor and thereby prevents its ubiquitination or if phosphorylation interferes with the proteasomal cleavage pattern is not understood. No examples are known of viral proteins that specifically target the cytosolic peptidases, even though upregulation of their activity may destroy potential deleterious antigens and may diminish the overall capacity of the infected cells to present antigens.

Viruses have, however, acquired mechanisms to inhibit ER import of peptides by interfering with the TAP complex. Herpes simplex virus (HSV)-1 and HSV-2 prevent TAP-mediated peptide transport by expression of the infected-cell protein 47 that acts by weakening the interaction between TAP1 and TAP2 (183). The HCMV unique short (US) 6 protein binds TAP at the ER luminal side thereby probably blocking the peptide exit pore of the TAP complex (3, 128, 192).

Inhibition of MHC class I surface expression is a common evasion strategy. The adenovirus E3/19-kDa protein (E19) contains a dilyamine ER retention signal and through binding to class I molecules retains those in the ER (216). Apart from interfering with the generation of antigenic peptides and their transport by TAP, HCMV deploys additional tactics to evade CTL recognition. Like adenovirus E19, HCMV’s US3 protein retains peptide-loaded class I molecules in the ER (160). US3 itself lacks an obvious ER retention signal, suggesting that a putative interacting protein that contains a retention signal may be required. Additionally, HCMV encodes the US2 and US11 proteins that target class I molecules for degradation by the ubiquitin-proteasome system by dislocating class I through the Sec61p translocon from the ER into the cytoplasm (356). It is tempting to speculate that the targeting of class I molecules by US2 and US11 may resemble the retrotranslocation and degradation of misfolded proteins by the ERAD pathway. It should, however, be noted that in
contrast to yeast, the human ERAD system is not well understood. HIV deploys another strategy to reduce MHC class I expression. The HIV-1 and HIV-2 Nef protein localizes to the cell membrane where it accelerates endocytosis of MHC class I complexes, followed by targeting to the lysosome (254). However, association with the endocytic machinery is not absolutely required for Nef-mediated downregulation of class I molecules, suggesting alternative targeting mechanisms. The Kaposi sarcoma herpes virus (KSHV) uses a similar strategy. The KSHV K3 and K5 proteins induce ubiquitination of the cytoplasmic tail of MHC class I molecules, which triggers their destruction in the lysosome (133). Another HIV protein, Vpu, selectively prevents the cell surface expression of newly synthesized class I molecules by inducing a destabilization event (165).

Although prevention of surface expression of peptide loaded class I molecules prevents CTL-mediated recognition, the immune system has evolved mechanisms to detect the loss of MHC class I expression. According to the missing-self hypothesis, natural killer (NK) cells recognize the absence of cell surface-expressed self-MHC class I molecules as a signal for destruction of a target cell (204). Expression of surface expressed class I homologues may thus provide a means to avoid detection by NK cells. Indeed, HCMV (21) and the human poxvirus molluscum contagiosum (296) express decoy MHC class I molecules that protect infected cells from NK-mediated lysis.
7. THE EPSTEIN-BARR VIRUS GLYCINE-ALANINE REPEAT

7.1 Epstein-Barr virus
The Epstein-Barr virus (EBV) is a human DNA virus, belonging to the γ-herpes virus subfamily and is closely related to EBV-like viruses found in Old World non-human primates like chimpanzees, baboons and rhesus monkeys. Less than 10% of the human population worldwide is EBV seronegative, highlighting its evolutionary success (reviewed in 273). Infection in children is largely asymptomatic. Only when primary infection occurs during adolescence or adulthood, EBV causes a benign self-limiting lymphoproliferative disease, known as infectious mononucleosis. Thereafter, the virus maintains an asymptomatic latent infection although the primary infection evokes a stringent immune response (reviewed in 274). Strong epidemiological and molecular evidence has linked EBV with a number of malignancies, including Burkitt’s lymphoma (BL), Hodgkin’s disease (HD), nasopharyngeal carcinoma (NPC), post-transplant lymphoproliferative disease (PTLD), X-linked lymphoproliferative syndrome and oral hairy leukopilakia (273).

B-lymphocytes form the reservoir of latent (non-productive) EBV infection, but lytic (productive) infection is believed to occur both in B-cells and epithelial cells of the oropharynx (reviewed in 167). EBV infection of B-cells in vitro produces immortalized lymphoblastoid cell lines, while in vitro infection of epithelial cells results in active replication followed by virus production and cell lysis (306).

7.2 Latency programs
Although the spectrum of viral antigens expressed in lytic and latent infection is dramatically different, strong CD8+ CTL responses are mounted in both cases, which persist in the T-cell memory compartment throughout life (reviewed in 274). While almost hundred EBV gene
products are expressed during replication, different levels of viral gene silencing are associated with the so-called latency programs of EBV (reviewed in 217). Besides two small non-polyadenylated EBV encoding RNA (EBER) transcripts that are found in all latency types, infected cells representing latency III maximally express eleven different antigens, including six nuclear proteins, the EBV-encoded nuclear antigens (EBNA) 1-6, three latent membrane proteins (LMP) 1 and 2A/2B, plus the poorly characterized BamHI-A rightward transcripts (BARTS) and a transcript from the BamHI A region, BARF0. The latency III program represents the least restrictive program and is associated with autonomous B-cell expansion that is probably required to increase the initial pool of infected cells (217). In latency II, viral gene expression is restricted to the LMPs and EBNA1. This program is found in circulating B-cells that subsequently home to the germinal centres of lymphoid follicles providing a pool of rapidly expandable cells that may periodically refresh the viral stock in healthy carriers (9). Latency I represents the most restrictive program and is characterized by expression of LMP2A alone or in combination with EBNA1. EBV-carrying resting memory B-cells either express latency I or latency 0 programs (49, 225, 279). The latter is characterized by complete silencing of the viral genome.

Like the correlation of the latency types with B-cell differentiation state, EBV-associated diseases can also be categorized based on their latency type. For example, malignant lymphoproliferative disorders, like PTLID in immunocompromised EBV carriers, display a latency III virus expression pattern (326), while latency II programs are found in HD, T-cell lymphomas and epithelial tumors such as NPC. BL is traditionally regarded as the prime example of a latency I malignancy, though EBNA1 is exclusively expressed, and not LMP2A (167). However, Kelly and co-workers recently reported on the identification of a subset of BL tumors that express most EBNAs except EBNA2 and the LMPs, thereby partially resembling latency III (164).
7.3 EBNA1

Experiments to test the specificity of dominant CD8⁺ CTLs against latent EBV proteins showed that their reactivity is predominantly directed towards epitopes derived from EBNA3, 4, 6 and sometimes EBNA2, 5 and the LMPs, while EBNA1 is not recognized when expressed as an endogenous antigen (reviewed in 274). The original observation that EBNA1-specific immune responses were lacking, was made using a tumor-rejection model in 1991 (332) and was later corroborated by the finding that EBNA1 was indeed unable to provoke CTL responses even when overexpressed in highly competent antigen presenting cells (196).

EBNA1 is a multi-domain phosphoprotein consisting of around 600 amino acids and it performs a wide range of functions, some of which have been characterized in detail (reviewed in 193) (Figure 5). The N-terminal part contains a GA repeat region (see below) flanked by two short positively charged repeats that are involved in DNA binding. The C-terminal region contains a nuclear localization signal and oligomerization and DNA-binding domains (4). Once dimerized, it can bind the origin of plasmid replication (OriP) and latent Q promoter (Qp) DNA sequences and act as transcriptional enhancer or repressor of these promoters, respectively. Binding to the Qp promoter represses EBNA1 transcription and has been suggested to decrease the total cellular immune activation state (289). Binding of EBNA1 to OriP’s tandem repeats and dyad symmetry sequences is required for plasmid maintenance (180), replication (314) and DNA distortion (147). Maintenance of the EBV episome in proliferating cells is critically dependent on EBNA1, which explains the observation that EBNA1 is the only viral antigen consistently found in EBV-associated malignancies (reviewed in 193).

Using yeast one- and two-hybrid approaches plus a recently reported affinity chromatography assay, a number of cellular EBNA1-interacting proteins have been identified. For example, the EBNA1 C-terminus interacts with the nuclear import factors karyopherin α1, karyopherin α2 (also known as importin α and Rch1), karyopherin β2, karyopherin β3 and importin β (92, 168). These interactions with EBNA1 may thus play a role in EBNA1 entry into the nucleus. EBNA1’s interaction
Figure 5. Outline and functions of EBNA1. EBNA1 is a phosphoprotein that in the B95.8 EBV sequence is composed of 607 amino acids. The N-terminal part contains two short glycine-arginine (GR) motifs that are interspaced by a 238 residue glycine-alanine (GA) repeat. GR1 and GR2 are involved in DNA binding and the GAR protects EBNA1 from ubiquitin-dependent proteasomal degradation. The C-terminal region contains a nuclear localization signal (NLS) and a dimerization and DNA-binding domain. EBNA1 dimers can bind the origin of plasmid replication (OriP) and the latent Q promoter (Qp). Binding of EBNA1 dimers to the tandem repeats (TR) and dyad symmetry (DS) sequences in the OriP is associated with maintenance of the viral episome, DNA distortion, replication start and transcriptional enhancement, while binding to Qp represses transcription of viral genes from this promoter, including EBNA1 itself.
with the EBNA1 binding protein 2 (Ebp2) plays a role in maintenance of the viral episome because Ebp2 acts as a physical linker between EBNA1, the viral episome and chromosomal DNA. The precise molecular mechanism behind viral episome maintenance remains obscure, although recent advances suggest critical roles for the telomeric repeat binding factor 2 (TRF2), the TRF2-interacting protein hRap1, and the telomere-associated poly (ADP-ribose) polymerase, which all bind to OriP in an EBNA1-dependent manner (73). Others have found interactions between EBNA1 and the replication protein A and the origin recognition complex, suggesting that they may be involved in the replication function of EBNA1 (74, 380). Interestingly, EBNA1 also interacts with the HSV-associated DUB Usp7/HAUSP, which modulates the proteasomal turnover of p53 (199). While Usp7 does not affect the turnover of EBNA1, a mutant EBNA1 defective in binding to Usp7 showed increased replication activity, suggesting that Usp7 may have a negative role in the replication function of EBNA1 (143).

Other EBNA1 interacting proteins include the nucleosome assembly protein 1 and the template activating factors-I α and β (143), which are involved in proper spacing of the nucleosomes and activation of replication and transcription through histone interactions, respectively. EBNA1 further interacts with two of the three subunits of the casein kinase II serine/threonine kinase and with the mitochondrial protein P32/TAP. The significance of these interactions remains obscure.

### 7.4 The GA repeat

A unique feature of the EBNA1 protein is the presence of an internal repeat motif solely consisting of glycine and alanine residues (14) (Figure 5). This GA repeat of the prototype B95.8 EBV virus is 238 residues long but the length varies between different virus strains between 60 and more than 300 residues (82). It is noteworthy that the GA repeat is the major target of EBNA1-specific antibody responses (76).
A first clue of the function of the GA repeat came from experiments using recombinant vaccinia virus to overexpress full length EBNA1 or EBNA1 lacking the GA repeat, linked to an immunodominant EBNA4 epitope (196). Deletion of the GA repeat resulted in efficient CD8+ CTL-induced killing, suggesting that the domain interferes with MHC class I presentation of EBNA1. Further, EBNA4 fused to a GA repeat was no longer recognized by EBNA4 specific CTLs, indicating that the GA repeat is a transferable element capable of conferring protection against CTL-dependent immune responses, a characteristic that was later supported by data from murine immunisation-models (230) and expression studies in simian cells (34). Intriguingly, EBNA1-specific CTLs can be demonstrated in all EBV carriers. However, CTL-induced killing requires internalization of exogenous EBNA1, while cells expressing endogenous EBNA1 remain unaffected (32,33). The presentation of exogenous EBNA1 occurs in a TAP-independent manner and the presence of EBNA1-specific CTLs may thus be explained by a cross-priming mechanism in which phagocytosed antigens gain access to the MHC class I loading compartment in professional antigen presenting cells (32, 33).

### 7.5 Inhibition of ubiquitin-mediated degradation by the GA repeat

A possible function for the GA repeat was identified using *in vitro* degradation assays. Levitskaya *et al.* showed that EBNA1 lacking the GA repeat domain and EBNA4 were efficiently degraded by the proteasome in an ATP-dependent manner (197). However, insertion of a 17 amino acid long GA repeat in EBNA4 resulted in complete stabilization of the protein (197). Experiments on the inhibitor of NF-κB, IκBα, showed that a short 8 residue GA repeat could completely block constitutive and signal-dependent degradation of the substrate (299), without affecting its phosphorylation, ubiquitination, subcellular localization, function or folding properties (194, 299). Co-immunoprecipitation experiments failed, however, to detect interaction of ubiquitinated IκBα-GA repeat chimeras with the proteasome, supporting a model in which the GA repeat could
affect events downstream of ubiquitination and preclude the substrate from stably interacting with the proteasome to allow substrate degradation. The IκBα model was also used to test for possible structural constraints within the octamer GA repeat (298). Replacing the alanines with valines showed no reduction of the inhibitory effect, while elimination of the glycine spacer or spacers with more than 3 glycines abolished the effect, as well as replacing one alanine with a polar or charged residue. A serine containing octamer that resembles simian repeats (see below) was inactive, although inhibition could be partially restored by three consecutive repeats, suggesting that the length of the repeat may correlate with the efficiency of inhibition (see also Paper I). Intriguingly, a repeat containing 8 consecutive alanines was as efficient as the original octameric GA repeat. As poly-alanine stretches have been shown to cause aggregation due to their hydrophobic nature, the additional glycine residues in EBV GA repeat may serve to avoid potential aggregation. Based on these findings, a model was proposed where efficient inhibition could occur upon interaction with a putative receptor, which would require at least three alanine residues of the GA repeat in a beta-strand conformation with adjacent hydrophobic binding pockets (298). This model has not been tested in interaction experiments since an alleged binding partner has yet to be identified (further discussed in Chapter 9). It is noteworthy, that introduction of a 24 amino acid GA repeat in the cell surface protein Erb2 could not rescue the chimeric protein from degradation (253). It is not known if the lack of protection is due to the membrane association of Erb2 or whether the tested GA repeat was not of sufficient length.

7.6 Biological significance of the GA repeat

Taken together, the studies described above have identified the GA repeat as a viral stabilization signal. Interestingly, the second human γ-herpesvirus KSHV encodes a latency-associated nuclear antigen (LANA) that contains a long highly acidic repetitive sequence rich in aspartic acid,
glutamic acid and glutamine, which has a comparable length and position to the GA repeat (48). Although LANA is a functional EBNA1 homologue, it remains to be tested if the LANA repeat functions analogous to the GA repeat. The EBNA1 homologous of EBV-like viruses found in baboon and rhesus monkeys also contain repetitive domains, which vary in length between 50 and 70 residues, and beside glycine and alanine contain serine residues (34). Blake and co-workers probed the evolutionary conservation of the GA repeat and found that substitution of the EBNA1 GA repeat with these simian GA repeats did not prevent the MHC class I presentation of EBNA1 epitopes, suggesting that the protection from CTL recognition may not be a strictly conserved function of the repeats and may instead represent a relatively recent evolutionary acquisition (34). In fact, EBV-infected cells express several viral antigens that efficiently mount specific immune responses (9, 10), indicating that a failure to present EBNA1-derived peptides may only contribute to escape from recognition by the immune system in cells with restricted viral gene expression. Instead, even a limited inhibition of EBNA1 proteolysis may serve another important role in survival of the virus. As mentioned above, maintenance of the viral episome is critically dependent on EBNA1 and the GA repeat may serve a non-immunological role and assures sufficient EBNA1 protein levels to allow preservation of the viral DNA. In line with this hypothesis, it has been observed that only a minority of freshly separated EBV-positive memory B-cells express EBNA1 mRNA (49). This suggests that while these cells exhibit low metabolic activity and lack ongoing transcription/translation, a stable EBNA1 protein may persist long enough to maintain the EBV episome.

The idea that the various repetitive sequences are absolutely required for latent infection of γ-herpesviruses has recently been challenged by the finding that a distant relative of EBV that is endemic in New-world primate marmosets, contains an EBNA1 homologue lacking any repetitive sequence (55). However, no function has been assigned to the homologue and whether it is resistant to proteasomal degradation is not known.
Taken together, it seems that protection of EBNA1 from ubiquitin-proteasome-dependent degradation may provide the virus with two possible advantages in establishing latent infection. First, the lack of EBNA1-derived antigens could prevent certain latently infected cells from detection by the immune system, which would present an especially stringent strategy considering the obligatory presence of EBNA1 in all virus-infected cells. Second, the lack of degradation results in EBNA1 being a long-lived protein, which may prevent loss of the viral episome even in cells lacking most metabolic activity. Additionally, this may allow further silencing of viral gene expression without losing the viral genome.
8. THE NUCLEOTIDE EXCISION REPAIR PROTEIN RAD23

8.1 Introduction

The *S. cerevisiae* Rad23 protein is a multi-domain protein composed of 398 amino acids (352). Rad23 is highly conserved during evolution and has several homologues, including the fission yeast Rhp23 (206) and human HHR23A and HHR23B proteins (219). Rad23 contains an N-terminal ubiquitin-like domain and two UBA domains; one internal (UBA1) and one located at the extreme C-terminus (UBA2) (Figure 6). The ubiquitin-like domain shares 22% identical and 43% similar (identical plus conserved changes) residues with yeast ubiquitin (352). In total, Rad23 contains 14 lysine residues and although Rad23 is indeed ubiquitinated it is a surprisingly stable protein with a half-life that exceeds 6 hours (352).

![Rad23](image)

**Figure 6. Outline of the Rad23 protein.** The location of the ubiquitin-like domain, two UBA domains (UBA1 and UBA2). The Rad4 interacting region is based on the HHR23B XPC binding domain described by Masutani et al. (218).

8.2 Functions of Rad23

A first clue to a possible function of Rad23 came from the observation that deletion of Rad23 results in UV sensitivity and inefficient NER (352). The same group also found that Rad23 lacking its ubiquitin-like domain could not suppress these phenotypes, suggesting an important role for this domain in DNA repair. Interestingly, replacement of the ubiquitin-like domain by ubiquitin functionally rescues Rad23-mediated DNA repair in a Rad23 deletion mutant (186, 352). The ubiquitin-like domain was shown to interact with catalytically active proteasomes (290) and the binding site
seems to locate to the Rpn1 and Rpn2 subunits of the 19S base complex (78, 284). Importantly, 19S RP appears to have a dual role in DNA repair (see also Chapter 4). The physical interaction between the Rad23 ubiquitin-like domain and the 19S RP was shown to be required for optimal NER in vitro (282), while in vivo the 19S RP can negatively regulate the NER rate independent of Rad23 (100). Based on these observations, it was suggested that the exclusive role of the Rad23 ubiquitin-like domain in NER might be to attenuate the negative effect of the 19S RP on NER (100), while the interaction of other NER components with the Rad23 C-terminal part would enhance NER. It is important to note that Rad23 lacking both UBA domains supports DNA repair (27), suggesting that other domains may be involved. In line with this hypothesis, it was shown that binding of Rad4 to a C-terminal domain of Rad23 stimulates the assembly of a multi-subunit DNA repair complex on UV-damaged DNA (116, 153), which in humans occurs upon interaction of HHR23A or HHR23B with the xeroderma pigmentosum group C protein (XPC) (312, 313). Importantly, two groups recently reported that Rad4 and XPC are unstable proteins that are degraded by the proteasome in the absence of Rad23 and HHR23, respectively, resulting in a NER-deficient phenotype (208, 241). Ng et al. proposed a model in which NER in humans may be regulated by HHR23-dependent control of XPC stability (241). Specifically, under normal conditions HHR23 may partially inhibit XPC proteolysis leading to proficient NER and solely upon DNA damage a further transient increase in XPC levels would result in enhanced NER capacity.

A direct role for Rad23 in the ubiquitin-proteasome system was suggested earlier. Genetic data showed that Rad23 has overlapping functions with the ubiquitin-binding proteins Rpn10, Ddi1 and Dsk2 (29, 59, 187). For example, double mutant yeast lacking Rad23 and Ddi1 displays growth defects that are not apparent in single mutant strains (59). Rad23 and Ddi1 together assure the legitimate proteasomal degradation of the anaphase inhibitor Pds1 (26). Simultaneous mutation of Rad23 and Rpn10 causes proteolytic defects and accumulation of ubiquitinated proteins (187). Importantly, alleviation of the phenotypes in
rad23Δ rpn10Δ yeast required both the ubiquitin-like domain and two intact UBA domains of Rad23, suggesting a functional interplay between these domains (187). While both Rad23 and Dsk2 are not essential genes, a double mutant is synthetically lethal at 35°C (95), suggesting a functional redundancy. Both proteins bind the proteasome through their ubiquitin-like domains and to poly-ubiquitin chains by virtue of their UBA domains.

The precise role of Rad23 in the ubiquitin-proteasome system is controversial. First, Rad23 has been suggested to promote degradation. By simultaneously binding to the ubiquitinated substrates and the proteasome, Rad23 could act as a putative shuttle factor (50), a role that was also proposed for Dsk2 (95). Rad23 (262) and the fission yeast homologues of Dsk2 and Rpn10 (121) can inhibit the disassembly of ubiquitin chains, which in a transient manner could stimulate degradation by allowing ubiquitinated proteins to interact longer with the proteasome. Second, Rad23 has been shown to inhibit the elongation of ubiquitin chains in vitro and overexpression of Rad23 stabilizes certain proteolytic substrates in ubiquitin-conjugated forms in vivo, suggesting that Rad23 may selectively inhibit degradation (51, 246). Indeed, HHR23 can stabilize XPC (241) and Rad23 expression stabilizes Rad4 (208) and the anaphase promoter protein Pds-1 (26). Raasi and Pickart recently provided additional experimental evidence for an inhibitory role by showing that the Rad23 UBA domains inhibited the in vitro degradation of Lys48-linked poly-ubiquitinated substrates by purified proteasomes, suggesting that the UBA domains might compete with the proteasome for binding to substrate-linked poly-ubiquitin chains (262). Is it possible that Rad23 both promotes and inhibits proteolysis? Besides ubiquitin, UBA domains bind to several other proteins (see Chapter 3 and below) and based on the structure of UBA domains it was suggested that they might contain several potential interacting surfaces (229). It is thus tempting to speculate that UBA domains can simultaneously bind different proteins, and based on the individual UBA-substrate interaction, a substrate may be protected from or targeted for proteolysis.
The involvement of human and mouse Rad23 homologues in NER has been studied in detail. However, a number of observations suggest additional roles for these proteins. For example, mice lacking the mouse HHR23B homologue (mHR23B) display strong developmental defects and surviving animals show impaired growth and male sterility (242). Together with the observation that embryonic fibroblast from these mice are not UV sensitive, it was suggested that mHR23B has a separate role in development that cannot be functionally compensated for by mHR23A. The finding that HHR23A binds several proteins through specific interactions with its UBA2 domain, may also point to alternative functions. These proteins include the HIV-1 Vpr protein (75, 363), the transcriptional adaptor protein p300 (382), the N-glycanase Pgn1 (316) and the DNA repair specific 3-methyladenine DNA glycosylase MPG (224).

Finally, it should be noted that there might be subtle functional differences between the yeast and human Rad23 proteins. For instance, the HHR23 ubiquitin-like domain mediates a direct interaction with the S5a subunit of the proteasome (137), while yeast Rad23 is incapable to interact with the yeast S5a counterpart Rpn10 (78).
9. GFP-BASED REPORTERS FOR MONITORING THE UBIQUITIN-PROTEASOME SYSTEM

A major part of the work described in this thesis is based on the development of green fluorescent protein (GFP) proteasome substrates, which allow monitoring of the ubiquitin-proteasome system in living cells (66). GFP was originally derived from the jellyfish *Aequorea victoria* and is a small 26kD soluble, stable protein with a half-life exceeding 24 hours (333). Importantly, GFP has no known interacting partners and by virtue of its autofluorescence GFP is a widely used tool in molecular and cell biology (333).

![GFP-based proteasome reporters](image)

**Figure 7.** GFP-based proteasome reporters. To generate N-end rule substrates ubiquitin was cloned in frame with GFP, and post-translational cleavage by UCHs exposes the new N-terminal "X" residue (66). Methionine (M) at position "X" results in a stable reporter whereas arginine (R) generates a destabilized GFP. The lysine (K) residues that serve as ubiquitination sites are underlined. UFD GFP was generated by a proline (P) at the position "X" or replacement of ubiquitin’s C-terminal glycine (G) with valine (V), which allows poor UCH cleavage and abolishes UCH cleavage, respectively. UFD GFP is ubiquitinated in the N-terminal ubiquitin moiety.

The GFP-based proteasome substrates were generated by insertion of different N-end rule and UFD degrons (Figure 7) that converted GFP into a substrate of the ubiquitin-proteasome system, although not all degrons targeted GFP with the same efficiency (66). Degradation of the substrates resulted in low fluorescence intensities and pharmacological inhibition of the proteasome led to the accumulation of GFP-fluorescent
cells, which could be monitored by fluorescence microscopy, flow cytometry and fluorimetry. Although the ubiquitination of GFP by the N-end rule and UFD degrons relies on different enzymes, both pathways converge at the level of the proteasome. This implies that these reporters cannot be used to test the individual catalytic activities of the proteasome, but rather monitor the general status of the ubiquitin-proteasome system. The GFP-reporters have proven their general applicability in several species including mammals (65, 201, 236, 343) (see also Paper I), Drosophila (212) and yeast (Paper III and IV).

The recent generation of transgenic mice that constitutively express a UFD targeted GFP substrate will allow in vivo studies on the role of the ubiquitin-proteasome system in various disorders (202). Other groups have reported on the generation of related reporter substrates. Bence and co-workers described a GFP-based reporter that is targeted through a short hydrophobic sequence (23), which was originally identified in a screen for degrons in yeast (101). Stack et al. amplified the UFD signal by insertion of multiple N-terminal ubiquitin moieties to GFP or β-lactamase (308). Some functional differences have been observed between these various reporters, suggesting caution when generalizing the effects with individual experimental set-ups. Additionally, Luker et al. recently reported on the generation a UFD-targeted luciferase-based reporter that allows bioluminescence imaging (211).
10. RESULTS AND DISCUSSION

10.1 Length-dependent effect of the GA repeat

Previous studies suggested that the GA repeat could inhibit ubiquitin-dependent proteasomal degradation of EBNA1. When introduced in other viral or non-viral proteasome substrates the GA repeat exclusively inhibited the degradation of its new host protein, suggesting its protective capacity functions in cis and can be transferred (see Chapter 6). The aim of the work described in Paper I was to further characterize the protective effect of this viral stabilization signal. We used different GFP-based proteasome substrates that were targeted for proteolysis with different efficiencies by either N-end rule or UFD signals (Figure 7). GA repeats of 25 and 239 residues were fused to the C-terminus of the GFP reporters and GFP levels were monitored by Western blotting and flow cytometric analyses. We found that introduction of GA repeats of increasing lengths resulted in enhanced protection of the fluorescent reporter from proteolysis, suggesting a direct correlation between the length of the repeat and the stability of the protein. Such a length-dependent effect was later also observed with the simian-like repeats in 1xBα (298). Interestingly, we observed that a full-length GA repeat could only partially protect GFP substrates targeted with strong degrons, while it fully protected substrates that contained less efficient degrons, such as a modestly destabilized UFD substrate. This finding raised the possibility that the presence of the full-length GA repeat in EBNA1 might not confer absolute protection if EBNA1 were to be efficiently targeted for proteasomal degradation. Indeed, when provided with a strong N-end rule degron we observed that EBNA1 could be degraded. Together, this study shows that a balance between the strength of the degron or ubiquitination kinetics and the length of the repeat determines the GA repeat-dependent protein stability. This further suggests that degradation by the ubiquitin-proteasome system cannot only be blocked but may also be fine-tuned by the GA repeat or similar stabilization signals.
Of particular interest was the finding that the addition of a strong degron resulted in proteasomal degradation of EBNA1 and could overcome the protection acquired by its internal GA repeat, suggesting that degradation of EBNA1 could in fact result in the production of EBNA1-derived antigenic peptides. Using CTL-based assays Tellam and colleagues subsequently showed that cells expressing various destabilized EBNA1 proteins could indeed be efficiently recognized and killed by EBNA1-specific CTLs (325). Of note, compared to wild type EBNA1, the N-end rule-targeted EBNA1 was less efficient when tested for plasmid maintenance and Qp transcriptional repression, likely due to increased turnover (Imreh, M., Heessen, S., Dantuma, N.P., Masucci, M.G., Sharipo, A., unpublished observations). The length of the GA repeat differs between viral isolates and it is not understood whether this reflects differences in efficiency of resistance to proteolysis or if the shortest repeats provide the minimally required protection. While the GA repeat equips EBNA1 with a clear stabilization signal, it is not known if EBNA1 also contains a degron. The simultaneous presence of a degron and a stabilization signal could have important regulatory functions, as discussed in Chapter 5 and Paper IV. If EBNA1 indeed contains a degron this would imply that it might recruit an E3 ubiquitin ligase. However, interaction studies have so far failed to identify a putative E3 among the EBNA1 binding proteins.

10.2 Functional stabilization of p53

The tumor suppressor p53 is a natural proteasome substrate targeted for ubiquitination by the RING-finger E3 Mdm2 (125, 145) or in human papilloma virus (HPV)-infected cells by the HECT domain ligase E6-AP (291, 292). The latter ubiquitin ligase is directed to p53 by the HPV E6 protein. By activating the transcription of several targets genes, p53 plays a crucial role in the regulation of cell cycle arrest and apoptosis (reviewed in 195). Inactivation of p53 by accelerated degradation by overexpressed Mdm2 or HPV infection is a common event in tumor development.
(reviewed in 346). Restoring p53’s tumor suppression function in such tumors may thus require that a therapeutically active p53 should resist the activity of these ligases.

In Paper II we applied the stabilization property of the GA repeat and tested its effect on p53 turnover. To this end p53-GA repeat chimeras were generated containing N- and C-terminal insertions of GA repeats of different lengths. All chimeras were protected from Mdm2- or E6-mediated degradation but this protection could be partially overcome by increasing the amount of ligase, although chimeras containing the full-length GA repeat were more resistant to higher levels of Mdm2 ligase compared to chimeras with short repeats, which is in agreement with our observations that the effect of the GA repeat is length-dependent (Paper I) (298).

We further demonstrate that p53-GA repeat chimeras remain functionally competent and efficiently induce cell cycle arrest and apoptosis, although chimeras containing the longest repeat showed a 50% reduction in transcriptional activation. When we tested the turnover of the chimeras in tumor cells expressing high levels of Mdm2 and in cells infected with the oncogenic HPV types, we found that all chimeras were stabilized while endogenous p53 was rapidly degraded. Finally, these cells were used to compare the anti-growth properties of the p53-GA repeat chimeras with p53 lacking the GA repeat. The chimeras clearly manifested improved growth inhibitory activity in these tumor cells. These results thus suggest that insertion of the viral repeat could provide a convenient strategy for stabilization of a wide variety of proteins that are of potential interest for gene replacement therapies.

Based on earlier work on the effect of the GA repeat on 1xBα turnover, it was proposed that 1xBα-GA repeat chimeras were ubiquitinated but could no longer interact with the proteasome, as monitored by co-immunoprecipitation analyses (299). Interestingly, when we co-expressed p53-GA239/C with Mdm2, we observed a distinct ladder of regularly spaced high-molecular weight species, suggesting that this chimera was subjected to ubiquitination. We exploited the easy access of ubiquitinated GA repeat-containing proteins and tested these for possible
interaction with the poly-ubiquitin chain receptor S5a. Using a sensitive in vitro binding assay we found that the ubiquitinated chimera retained the capacity to interact with immobilized S5a, suggesting that the GA repeat may not preclude the binding of a ubiquitinated substrate to the proteasome, but instead may affect the quality of this interaction. The interaction of GA repeat-containing proteins and the proteasome may thus have a lower affinity or higher dissociation rate ultimately resulting in premature release of the substrate. It was recently observed that the base of the proteasome 19S RP contains a chaperone-like activity that could provoke refolding rather than degradation of substrates, suggesting that GA repeat-containing substrates may even undergo partial unfolding and refolding during their limited interaction with the proteasome. Whether the GA repeat inhibits unfolding or stimulates refolding is not known. Does the GA repeat compete for interaction with an ubiquitin receptor? At present, no putative GA repeat interaction surface on the proteasome has been identified, although its presence was suggested based on the observation that a GA repeat peptide can compete with the degradation of ubiquitinated lysozyme in vitro without affecting interaction with S5a (194). Clearly, identification of an interacting partner could aid in understanding the inhibitory mechanism. It is noteworthy that a yeast two-hybrid screen using the yeast codon-optimised GA repeat (Paper III) as bait failed to identify any cellular interacting proteins (Dantuma, N.P. unpublished observation). Together, the studies described in Paper I and Paper II demonstrated that the GA repeat could efficiently counteract proteolysis of ubiquitinated substrates regardless of the degron or E3 ubiquitin-ligase that induces ubiquitination, although the precise mode of action remains poorly understood.

### 10.3 Reconstitution of the GA repeat effect in yeast

The ubiquitin-proteasome system has traditionally been studied intensively in the budding yeast *S. cerevisiae*. Since both the N-end rule and UFD signals were originally identified in this yeast, we exploited the
GFP reporters harbouring these signals in Paper III to probe further into the mechanism underlying the protective effect of the GA repeat and tested its functionality in yeast.

Using a panel of yeast deletion mutants, we first confirmed that the GFP-based proteasome substrates were targeted through independent ubiquitination pathways. Notably, the Ub\textsubscript{G76V}-GFP reporter was only partially stabilized in various UFD deletion mutant yeasts. The E2s Ubc4/Ubc5, E3 Ufd4 and E4 Ufd2 ubiquitination proteins are differently involved in the formation of Lys\textsuperscript{29} and Lys\textsuperscript{48}-linked poly-ubiquitin trees linked to Lys\textsuperscript{29} and Lys\textsuperscript{48} of the N-terminal ubiquitin moiety, respectively. Specifically, Ubc4/Ubc5 and Ufd4 are required for the formation of Lys\textsuperscript{29} chains, while Ufd2 extends Lys\textsuperscript{48} ubiquitin trees. Pioneering work by Varshavsky and co-workers showed that a β-galactosidase-based UFD substrate required both Lys\textsuperscript{29} and Lys\textsuperscript{48}-linked poly-ubiquitin trees while ubiquitination of a dihydrofolate reductase UFD substrate strictly depended on Lys\textsuperscript{29} linked chains (158). It was recently reported that in mammalian cells the Ub\textsubscript{G76V}-GFP substrate can be targeted independently by the Lys\textsuperscript{29} and Lys\textsuperscript{48}-linked poly-ubiquitin trees (201), suggesting that the partial stabilization observed in UFD deletion mutants may be explained by the fact that only a single ubiquitin tree is affected in each of these mutants while the remaining tree can still trigger degradation of the Ub\textsubscript{G76V}-GFP substrate.

In the second part of this study we aimed to test the effect of the EBV GA repeat on the turnover of the GFP substrates. However, because the codon usage of the EBV GA repeat is sub-optimal in yeast, we observed low expression levels of GFP substrates fused to the EBV repeat. Expression of proteins carrying GA repeats thus required the generation of a yeast codon-optimized recombinant GA (rGA) repeat that was generated by sequential amplification of a minimal Gly-Gly-Ala-Gly-Ala-Gly-Ala-Gly octamer shown to stabilize I\textsubscript{xBα} (299). We then first confirmed that introduction of 53, 107 and 251 amino acids long rGA repeats in the GFP substrates resulted in stabilization of the proteins in mammalian cells, resembling the effect of the EBV GA repeat (Paper I). Next, we found that the rGA repeat was expressed efficiently although we still observed a
modest length-dependent decrease in expression levels. When tested in pulse-chase analyses we found that the rGA repeat could efficiently stabilize the N-end rule and UFD targeted GFP substrates in yeast. Notably, rGA repeat-mediated protection of the N-end rule substrate was achieved in yeast lacking Rpn10, which is the yeast homologue of the human ubiquitin-chain interacting protein S5a, providing conclusive evidence that this protein is not required for the protective effect in yeast. As mentioned in Chapter 4, Rpn10 is mainly present in a non-proteasome associated form (340), while its mammalian homologue S5a appears to be constantly in complex with the 26S proteasome (127). This suggests potential functional differences between the yeast and mammalian S5a proteins, and we can thus not formally exclude the possibility that in human cells S5a may be involved in the GA repeat-mediated stabilization effect. Nevertheless, the functional reconstitution of the protective effect of the GA repeat in yeast implies that this stabilization signal targets an evolutionary conserved mechanism in the ubiquitin-proteasome system and now allows us to employ genetic yeast screens to unravel its mode of action.

10.4 The Rad23 stabilization signal
In paper IV we describe the identification and characterization of an intrinsic stabilization signal in the multifunctional budding yeast Rad23 protein. Rad23 is a long-lived protein despite the fact that it is ubiquitinlated and physically interacts with the proteasome (290, 352) (see also Chapter 8). These characteristics normally predispose proteins for proteolysis, suggesting that Rad23 may have acquired the ability to counteract its destruction by the proteasome. To assess whether Rad23 contains any stabilization signal, we fused truncated forms of the Rad23 protein to the C-terminus of the N-end rule GFP substrate and expressed those in yeast. Surprisingly, insertion of the C-terminal UBA2 domain but not UBA1 was sufficient to significantly increase the steady state level of the reporter as quantified by the fluorescent intensity and to delay

- 67 -
proteasomal degradation as measured by shut-off experiments. An UBA2 domain that contained a single L392A mutation that destroys its structural integrity could no longer stabilize the GFP substrate. A small but significant stabilization effect was also observed upon insertion of a region that contains the alanine-rich stretch Ala-Ser-Ala-Ala-Pro-Ala-Ala-Thr-Ala. It is of interest that this region fulfils the GA repeat consensus as well as the minimal length found to be protective in the IxBα model (see Chapter 7). If the alanine-rich motif is indeed responsible for this additional stabilization effect, the native Rad23 protein would appear to contain multiple stabilization signals that together rescue the protein from proteasomal degradation.

Importantly, we found that the protective effect of the UBA2 domain was restricted to the protein harbouring the stabilization signal, implying that the UBA2 domain behaves as a cis-acting protective motif. High-level expression of an UBA2-containing GFP substrate does not cause general inhibition of the ubiquitin-proteasome system, although we cannot exclude that the turnover of certain proteins may have been affected.

The GFP reporter protein linked to the Rad23-derived stabilization signal accumulated as ubiquitinated species conjugated to short ubiquitin chains, implying that the UBA2 domain may not interfere with the initial ubiquitination of the substrate and rather freezes short ubiquitin trees, which is in line with the reported ability of UBA domains to interfere with elongation and disassembly of ubiquitin trees. However, we found that the UBA1 domain lacked any stabilizing activity. This indicates that the general inhibitory activity of UBA domains towards ubiquitin chains is not the sole explanation for UBA2-mediated protein stability, though it may well be a contributing factor.

In the context of the native Rad23, distortion of the UBA2 domain converts Rad23 into a short-lived protein that is targeted for degradation through its ubiquitin-like domain, thereby excluding the possibility that misfolding caused the increased degradation of mutant Rad23 lacking a functional UBA2 domain. It thus seems that the Rad23 ubiquitin-like domain only behaves as a proteasome interacting domain but can also function as an authentic degron.
Finally, the UBA2 domain derived from the human Rad23 homologue HHR23A also confers stability upon the GFP-based proteasome substrate in yeast, while its UBA1 domain lacks any protective effect. When tested in human cells, we found that both the Rad23- and the HHR23A-derived UBA2 domains efficiently stabilized the GFP reporter, while both UBA1 domains displayed no protective effect, indicating that the UBA2-mediated stabilization effect is conserved in yeast and humans.

Our observations suggest that Rad23 and HHR23A combine degrons with stabilization signals in order to evade proteolytic inactivation while maintaining a transient association with the proteasome. We suggest two models that may explain how limited ubiquitination may play a role in the protein stability caused by UBA2. First, interfering with the formation of poly-ubiquitin trees may secure an alternative interaction between the Ubl domain and the Rpn1/Rpn2 subunits of the 19S cap, because the binding of poly-ubiquitin trees present on Rad23 to the Rpn10 and Rpt5 subunits would otherwise support degradation. Second, it has been suggested that the length of a poly-ubiquitin chain is directly proportional to the residence time of the substrate at the proteasome as DUBs will eventually terminate the interaction by disassembly of the poly-ubiquitin tree. As such, 19S RP-bound DUBs directly control how long the proteasome is occupied by the substrate and disassembly of short poly-ubiquitin trees prior to translocation in the 20S CP can result in release of the recruited ‘substrate’ (184). Thus, conjugation of a few ubiquitins onto Rad23 might only support a transient Rad23-proteasome interaction that would allow Rad23 to escape degradation. Both models predict that inactivation of the UBA2 domain will allow conjugation of long poly-ubiquitin chains to Rad23 and binding to the Rpn10 and Rpt5 subunits will result in a classical substrate-proteasome interaction followed by degradation.

Together, our data on Rad23 suggest that this protein may have evolved from a classical proteasome substrate into a specialized protected ‘substrate’ that assists in ubiquitin-proteasome-dependent functions including escorting ubiquitinated substrates to the proteasome or interacting with the 19S RP during DNA repair.
11. FUTURE PROSPECTS

The studies described in this thesis have revealed that the degradation of proteins by the ubiquitin-proteasome system is controlled at more levels than previously appreciated. We have shown that a viral repetitive sequence can inhibit proteasomal degradation of artificial and natural proteasome substrates, regardless of the ubiquitin ligase involved. While its mode of action has not yet been resolved, it is has been proposed that the repeat changes the quality of the interaction of the ubiquitinated substrate with the 19S cap of the proteasome resulting in the premature release of the substrate. Immunoprecipitation studies have failed to show interaction of GA repeat containing 1xBα, suggesting that more sensitive methods are required to determine putative interaction partners. Our observation that the protective effect of the GA repeat is functional in yeast now allows the design of genetic screens that can identify gene products involved in the protective effect (Paper III). As mentioned earlier, a pilot yeast-two hybrid screen failed to identify rGA repeat binding partners (N.P. Dantuma, unpublished observation), suggesting that other approaches should be considered. A second approach could thus involve the subdivision of the individual 19S cap subunits as GST-fusion proteins to determine if the GA repeat specifically interacts with any subunit in vitro. However, this method has some serious drawbacks as subunits may not acquire their original fold as single subunits or a functional repeat may require simultaneous interaction with multiple partners. These potential problems can be circumvented by applying a method described by Leggett and co-workers, who reported on the affinity purification of 26S proteasome holoenzyme and regulatory particle (191). Both assemblies could be individually tested for interactions with the GA repeat. Possible GA repeat interacting subunits may also be identified using an UV-inducible cross-linking method, which has for instance been instrumental in the identification of the Rpt5/S6’ subunit as a poly-ubiquitin chain receptor.
Based on the original findings that the GA repeat exhibits immune-evasion properties and the subsequent observations that the repeat functionally stabilizes a wide variety of proteasome substrates (Paper I and II), it has been suggested that the repeat may have possible applications in gene transfer settings that may require protection of the newly introduced protein from ubiquitin-dependent proteasomal degradation and from strong immune reactions against the products of the transfer vector or the transferred gene itself (258). An illustrative example of the first setting is based on the tumor suppressor p53 (Paper II). Most tumors exhibit impaired regulation of p53, either through loss-of-function mutations or by accelerated proteasomal degradation (346), which is accomplished by increased levels of the natural ubiquitin ligase Mdm2 (125, 145) or by redirection of the cellular ubiquitin ligase E6-AP to p53 by the E6 proteins of the oncogenic HPV variants 16 and 18 (291, 292). Importantly, p53 is an attractive candidate for gene replacement studies and in Phase I clinical trials with p53-encoding adenoviral vectors treatment resulted in decreased tumor progression (reviewed in 278). While tumors expressing mutant p53 may certainly benefit from this approach, in tumors that are HPV-positive or express high Mdm2 levels the newly introduced p53 will likely succumb to rapid proteolytic inactivation just like the endogenous protein. In Paper II we show that p53-GA repeat chimeras are stable, resist high levels of Mdm2 or expression of the HPV E6 protein and retain their capacity to induce apoptosis and cell cycle arrest, suggesting that these chimeric proteins may provide an excellent strategy for gene therapy applications in tumor cells that display increased proteasomal destruction of the endogenous product. It is tempting to speculate that this approach may also be applied to stabilize other proteins with therapeutic potential but that display accelerated proteasomal degradation in malignant cells.

A second application of the GA repeat may be to confer immunoprotection on the gene-of-interest or the transfer vector. However, this approach has not yet been put to the test and some considerations apply regarding the position and length of the repeat. Most MHC class I presented peptides are derived from defective ribosomal
products that are generated by premature termination of translation (293). As such, positioning of the GA repeat close to the N-terminus may prevent the generation of antigenic peptides from the full-length and possible truncated products. Further, the data on the degradation of N-end rule targeted EBNA1 (Paper I) and its subsequent immune recognition (325) stresses that the repeat must be of sufficient length to minimize any CTL recognition. The final consideration deals with the lack of inhibition of cross-priming by the GA repeat. It has been reported that the internalisation of EBNA1 by antigen presenting cells results in the generation of antigenic peptides derived from EBNA1 (33). It is thus possible that CTL responses may be generated to the transferred gene although CTLs will most likely fail to recognize cells that would attempt to generate epitopes from the GA repeat containing protein through the classical MHC class I antigen presentation pathway.

Finally, our work on the Rad23 UBA2 domain has defined the first example of a cellular stabilization signal. Biochemical and structural analysis of UIM, CUE and NZF domains have shown that also these modules can bind ubiquitin chains. It will be interesting to test if the other ubiquitin-binding domains also contain stabilization signals and whether these signals allow the formation of transient non-proteolytic interactions with the proteasome. An important open question is the role of ubiquitination in the regulation of Rad23 function. It will be interesting to identify which lysine residue(s) are ubiquitinated in Rad23. Of interest is further the functional significance of this ubiquitination and whether it is required for efficient DNA repair.
ACKNOWLEDGEMENTS

Welcome to the certainly most read and probably most enjoyable part of this thesis. Although the appearance of one author name on the cover may suggest that this thesis is the product of a single individual, the help of others was more than essential. It is here that I would like to express my sincere gratitude and thanks to all the 'others'. Special thanks to:

**Nico:** Since our first email contact in 1998, you have been the main driving force behind all the projects that I’ve been involved in. Your knowledge, unstoppable enthusiasm and working spirit have helped me to scientifically mature since then and have inspired me to further pursue a scientific career. Outside the lab you’ve been a good friend and I won’t forget our Friday afternoon football matches in the old gym, the climbing, swimming, skiing, skating and early morning fishing expeditions in Djurgården.

**Maria:** Thank you for your continuous support from the first day I arrived at MTC. From you I definitely learned not to be afraid to always be critical and ask questions. Some time ago I asked a question during a seminar and someone commented that it sounded like they heard Maria talking. I regarded that as a compliment.

**Kristian:** The first time I met you I thought you were completely nuts (probably you were, at that point). Since then I’ve had to revise my original opinion and you’ve never ceased to amaze me. Together we enjoyed so many parties, lunches, boat trips, game nights, saunas, BBQs and more recently Jeu des Boules, which somehow always were accompanied by sometimes serious but most often completely irrelevant and ridiculous discussions. The point being: it is great to have you as a friend.

**Richard S.**: Having you around has been a constant inspiration. You are probably one of the most generous people I know and I’ve appreciated all
your help and friendship tremendously. Thank you for introducing me to
Jack Johnson, Mason Jennings and last but not least shrimp burritos.

Rutger: What better to have a friend who appreciates Dutch cheese, beer
and humor as much as I do? Your enthusiasm and laughs always create a
good atmosphere and jag hoppas att vi ska skiljas inte, vitlök!

Kristina: Thank you for being a great colleague and friend. I very much
enjoyed our many discussions about life, work and science. By the way,
for a first-timer with Sinterklaas in ’99 you did very well.

Lisette: It’s always nice to have chat in Dutch. Thanks for always being
there and helping out, especially during the last year.

Victoria: Your special vibe has boosted the atmosphere in lab a great
deal. I’ve always appreciated your enthusiasm and support.

Anna C.: We both spend a small fortune in that stupid coffee machine.
Special thanks for your help with improving my English vocabulary.

Fausto: Thank you for all the gourmet dinners we’ve enjoyed together. I
never thought I would cook ‘North-African’ one day.

Yvonne & Marianne: It cannot be said enough. Without you two the lab
would easily deteriorate into an unorganised chaos. Your input and
commitment makes life a lot easier for everyone. Stor Tack!

Former Lab members: Thanks to Anatoly, Ainars, Teresa, Jelena
Petrovic, Xu Hong, LiQi, Daorong, Cheng Hong, Robert, Jelena Levitskaya,
Rikard, Simona, Tomoaki, Isodoro, Riccardo, Piri, Madeleine, Rebecca,
Rosario and Iris.

Current Lab members: Elena, Ulrika, Jianjun, André, Michael, Vaya,
Victor, Anders, Sara, Sei and Abela. Thanks for creating an improved and
relaxed lab-atmosphere and helping out whenever needed.

Project students: Niloo & Peter. Thanks for your assistance and teaching
me that good supervision requires special skills.
**Galina & Natalia:** Your help during the p53 project has been crucial.

**The Band:** Thanks for all the tunes. Let’s hope now the file-sharing business is dying, that a major label will pick up our demo instead. Yèh, right!?

**Walter, His Dudeness, Quintana, Bunny, Maude & Duderina:** Bowling should be an Olympic sport. Over the line!

**Richard M.:** Our continuous friendship proves that Stockholm is not that far from Haarlem.

**Shahab & Luigi:** Thanks Dudes!

**Astrid, René & Ilse:** Thanks for all your help, trust and support.

**Aniol, Zheer, Monica, Carla & Deirdre:** Thanks for all the memorable ‘Wednesday dinners’ we shared in Pax back in ‘98/’99.

**Mettine:** Your cards, emails and packages with ‘lectuur’ are always greatly appreciated.

**Martijn:** Although our email contacts seem to suffer from major interval problems, I’ve always enjoyed them. Leve Toeter ’94!

**Anna N., Wiktoria, Charlotte, Johanna & Johan:** Det är roligt att ha några svenska vänner.

**Jantina & Ramón:** Wanneer gaan we weer een keertje pimpampetten?

**Ward:** Thanks to Ali for turning you around. You are a true brother.

**Ineke & Ton:** Hassan die wiel bedanke vuur al die goeihe zorghe.

**Het Thuisfront:** My parents Maartje and Henk. Thanks for your unconditional support and for introducing me to the world of biology. Your genuine interest in all living creatures is a great inspiration.

**Sandra:** What would I do without you? You shared my tears and joys during the last six years. You are the sea, upon which I float.
REFERENCES


binding protein Mcb1 is a component of the 26S proteasome in Saccharomyces cerevisiae and plays a nonessential, substrate-specific role in protein turnover. Mol Cell Biol 16:6020-8.


