expressed in low abundance) although the suppression techniques described above help decrease this problem significantly. The use of FlAsH as an FRET acceptor from CFP (or the red version from GFP) suffers much less from this problem as only specifically bound FlAsH is excited through FRET. Background FlAsH staining is not significantly excited by the wavelengths used for CFP. Further optimization of the FlAsH target site, perhaps through screening of peptide libraries, should still allow strong binding at concentrations of dithiol that minimize background staining.

[41] Ubiquitin Fusion Technique and Its Descendants

By ALEXANDER VARSHAVSKY

The ubiquitin (Ub) fusion technique was developed in 1985–1986, through experiments in which a segment of DNA encoding the 76-residue Ub was joined, in frame, to DNA encoding *Escherichia coli* β-galactosidase (βgal).[1,2] When the resulting protein fusion was expressed in the yeast *Saccharomyces cerevisiae* and detected by radiolabeling and immunoprecipitation with an anti-βgal antibody, only the moiety of βgal was observed, even if the labeling time was short enough to be comparable to the time (1–2 min) required for translation of the Ub-βgal open reading frame (ORF). It was found that in eukaryotic cells the Ub moiety of the fusion was rapidly cleaved off after the last residue of Ub (Fig. 1).[1] The proteases involved are called deubiquitylating[3] enzymes (DUBs) or Ub-specific processing proteases (UBPs).[4–7] A eukaryotic cell contains more than 10 distinct DUBs, all of which are highly specific for the Ub moiety. The *in vivo*

[1] A. Bachmair, D. Finley, and A. Varshavsky, *Science* **234**, 179 (1986).
[2] A. Varshavsky, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 12142 (1996).
[3] Ubiquitin whose C-terminal (Gly-76) carboxyl group is covalently linked to another compound is called the *ubiquityl* moiety, the derivative terms being *ubiquitylation* and *ubiquitylated*. The term *Ub* refers to both free ubiquitin and the ubiquityl moiety. This nomenclature, which is also recommended by the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology,[19] brings Ub-related terms in line with the standard chemical terminology.
[4] K. Wilkinson and M. Hochstrasser, *in* "Ubiquitin and the Biology of the Cell" (J.-M. Peters, J. R. Harris, and D. Finley, eds.). Plenum Press, New York, 1998.
[5] J. W. Tobias and A. Varshavsky, *J. Biol. Chem.* **266**, 12021 (1991).
[6] R. T. Baker, J. W. Tobias, and A. Varshavsky, *J. Biol. Chem.* **267**, 23364 (1992).
[7] C. A. Gilchrist, D. A. Gray, and R. T. Baker, *J. Biol. Chem.* **272**, 32280 (1997).

0076-6879/00 $30.00

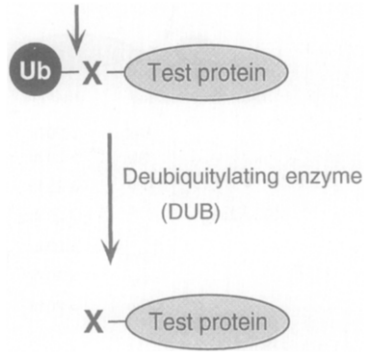

FIG. 1. The ubiquitin fusion technique. Linear fusions of Ub to other proteins are cleaved after the last residue of Ub by deubiquitylating enzymes (DUBs) (see text).[1,2]

cleavage at the Ub–polypeptide junction of a Ub fusion has been shown to be largely cotranslational.[8,9]

One physiological function of the cleavage reaction (Fig. 1) is to mediate the excision of Ub from its natural DNA-encoded fusions either to itself (poly-Ub)[10] or to specific ribosomal proteins.[11,12] Many of the DUB proteases that catalyze the cleavage of linear Ub fusions can also cleave Ub off its branched, posttranslationally formed conjugates, in which Ub is joined either to itself, as in a multi-Ub chain, or to other proteins.[4,13] A branched Ub–protein conjugate usually comprises a multi-Ub chain covalently linked to an internal lysine residue of a substrate protein. The ubiquitylated substrate is processively degraded by the 26S proteasome, an ATP-dependent multisubunit protease.[14–17] For reviews of the Ub system, see Refs. 18–22.

Another finding about the DUB-mediated cleavage reaction (Fig. 1)

[8] N. Johnsson and A. Varshavsky, *EMBO J.* **13**, 2686 (1994).
[9] G. C. Turner and A. Varshavsky, submitted (2000).
[10] D. Finley, E. Özkaynak, and A. Varshavsky, *Cell* **48,** 1035 (1987).
[11] K. L. Redman and M. Rechsteiner, *Nature* (*London*) **338**, 438 (1989).
[12] D. Finley, B. Bartel, and A. Varshavsky, *Nature* (*London*) **338**, 394 (1989).
[13] C. M. Pickart, *FASEB J.* **11**, 1055 (1997).
[14] O. Coux, K. Tanaka, and A. L. Goldberg, *Annu. Rev. Biochem.* **65**, 801 (1996).
[15] W. Baumeister, J. Walz, F. Zühl, and E. Seemüller, *Cell* **92**, 367 (1998).
[16] M. Rechsteiner, *in* "Ubiquitin and the Biology of the Cell" (J. M. Peters, J. R. Harris, and D. Finley, eds.), pp. 147–189. Plenum Press, New York, 1998.
[17] G. N. DeMartino and C. A. Slaughter, *J. Biol. Chem.* **274**, 22123 (1999).
[18] M. Hochstrasser, *Annu. Rev. Genet.* **30**, 405 (1996).
[19] A. Varshavsky, *Trends Biochem. Sci.* **22**, 383 (1997).
[20] A. Hershko and A. Ciechanover, *Annu. Rev. Biochem.* **76**, 425 (1998).
[21] T. Maniatis, *Genes Dev.* **13**, 505 (1999).
[22] L. Hicke, *Trends Cell Biol.* **9**, 107 (1999).

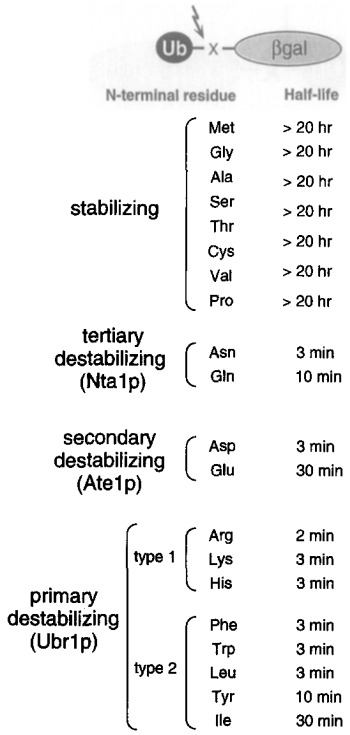

FIG. 2. The N-end rule of the yeast *S. cerevisiae*.[2] Specific residues at the N terminus of a test protein such as βgal are produced by the Ub fusion technique (Fig. 1 and text). The *in vivo* half-lives of the corresponding *X*-βgal proteins are indicated on the right. Stabilizing N-terminal residues (Met, Gly, Ala, Ser, Thr, Cys, Val, and Pro) are not recognized by Ubr1p (N-recognin), the E3 component of the N-end rule pathway. Primary destabilizing N-terminal residues (Arg, Lys, His, Phe, Trp, Leu, Tyr, and Ile) are directly bound by either type 1 or type 2 substrate-binding sites of Ubr1p. Secondary destabilizing N-terminal residues are arginylated by the *ATE1*-encoded Arg-tRNA-protein transferase (R-transferase), yielding the N-terminal Arg, a primary destabilizing residue. Tertiary destabilizing N-terminal residues Asn and Gln are deamidated by the *NTA1*-encoded N-terminal amidohydrolase (Nt-amidase), yielding the secondary destabilizing residues Asp and Glu, respectively. The N-end rule of mammalian cells is similar but contains fewer stabilizing residues.[2]

led to the discovery of the N-end rule, a relation between the *in vivo* half-life of a protein and the identity of its N-terminal residue (Fig. 2).[1] First, it was shown that the cleavage of a Ub–*X*–polypeptide fusion after the last residue of Ub takes place regardless of the identity of a residue *X* at the C-terminal side of the cleavage site, proline being the single exception. By allowing a bypass of the "normal" N-terminal processing of a newly formed protein, this result yielded an *in vivo* method for placing different

residues at the N termini of otherwise identical proteins. Second, it was found that the *in vivo* half-lives of the resulting test proteins were determined by the identities of their N-terminal residues, a relation referred to as the N-end rule (Fig. 2).[1] The N-end rule pathway, which targets the N terminus-specific degradation signals, called the N-degrons, is one pathway of the Ub system. For a review and work on the N-end rule pathway, see Refs. 2 and 23–31.

The Ub fusion technique (Figs. 1 and 2) remains the method of choice for producing, *in vivo*, the desired N-terminal residue in a protein of interest. Owing to the constraints of the genetic code, nascent proteins bear N-terminal methionine (formyl-Met in prokaryotes). The known methionine aminopeptidases (MAPs), which remove N-terminal Met, do so only if the residue to be exposed is stabilizing according to the yeast-type N-end rule.[2,32] In other words, MAPs do not cleave off N-terminal methionine if it is followed by any of the 12 destabilizing residues (Fig. 2). The Ub-specific DUB proteases are free of this constraint, except when the residue *X* of a Ub–*X*–polypeptide is proline, in which case the cleavage still takes place but at a much lower rate.[1,33] More recently, a specific DUB was identified that can efficiently cleave at the Ub–proline junction.[7]

The Ub fusions can be deubiquitylated *in vitro* as well.[25,34,35] The high activity and specificity of DUBs should make them the reagents of choice for applications that involve, for example, the removal of affinity tags from overexpressed and purified proteins. Unfortunately, there are no commercially available DUBs at present, in part because of difficulties encountered in purifying and stabilizing large DUBs such as *S. cerevisiae* Ubp1p, and also because Proteinix (Rockville, MD), a company that has held the licenses for Ub fusion patents over the last decade, has not commercialized this technology.

Another major application of the Ub fusion technique resulted from the observations that expression of a protein as a Ub fusion can dramatically

[23] C. Byrd, G. C. Turner, and A. Varshavsky, *EMBO J.* **17**, 269 (1998).
[24] Y. T. Kwon, Y. Reiss, V. A. Fried, A. Hershko, J. K. Yoon, D. K. Gonda, P. Sangan, N. G. Copeland, N. A. Jenkins, and A. Varshavsky, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 7898 (1998).
[25] I. V. Davydov, D. Patra, and A. Varshavsky, *Arch. Biochem. Biophys.* **357**, 317 (1998).
[26] Y. T. Kwon, A. S. Kashina, and A. Varshavsky, *Mol. Cell. Biol.* **19**, 182 (1999).
[27] Y. T. Kwon, F. Lévy, and A. Varshavsky, *J. Biol. Chem.* **274**, 18135 (1999).
[28] F. Lévy, J. A. Johnston, and A. Varshavsky, *Eur. J. Biochem.* **259**, 244 (1999).
[29] T. Suzuki and A. Varshavsky, *EMBO J.* **18**, 101 (1999).
[30] Y. Xie and A. Varshavsky, *Curr. Genet.* **36**, 113 (1999).
[31] P. O. Falnes and S. Olsnes, *EMBO J.* **17**, 615 (1998).
[32] R. A. Bradshaw, W. W. Brickey, and K. W. Walker, *Trends Biochem. Sci.* **23**, 263 (1998).
[33] E. S. Johnson, B. W. Bartel, and A. Varshavsky, *EMBO J.* **11**, 497 (1992).
[34] D. K. Gonda, A. Bachmair, I. Wünning, J. W. Tobias, W. S. Lane, and A. Varshavsky, *J. Biol. Chem.* **264**, 16700 (1989).
[35] R. T. Baker, *Curr. Opin. Biotechnol.* **7**, 541 (1996).

augment the yield of the protein.[36–39] The yield enhancement effect of Ub was observed with short peptides as well.[40,41] This and other applications of Ub fusions are described below, with references to the original articles and specific constructs.

Production and Uses of N-Degrons

An N-degron comprises the destabilizing N-terminal residue of a protein and an internal lysine residue.[2,29,42,43] A set of N-degrons containing different N-terminal residues that are destabilizing in a given cell defines the N-end rule of the cell.[2] The lysine determinant of an N-degron is the site of formation of a substrate-linked multi-Ub chain.[13,18,44] A way to produce an N-degron in a protein of interest is to express the protein as a Ub fusion in which the junctional residue (which becomes N-terminal on removal of the Ub moiety) is destabilizing (Fig. 2). An appropriately positioned internal lysine residue (or residues) is the second essential determinant of N-degron. Many natural proteins lack such "targetable" lysines, and therefore would remain long-lived even if their N-terminal residue were replaced by a destabilizing residue. One way to bypass this difficulty is to link a protein of interest to a relatively short (<50 residues) portable N-degron that contains both an N-terminal destabilizing residue (produced through a Ub fusion) and a requisite lysine residue(s). The earliest portable N-degron of this kind is still among the strongest known (Fig. 3B).[1,29,42] It was found, using the new strategy of a screen in the sequence space of just two amino acids, lysine and asparagine, that certain sequences containing exclusively lysines and asparagines can function *in vivo* as highly effective N-degrons.[29] The portability and modular organization of N-degrons make possible a variety of applications whose common feature is the conferring of a constitutive or conditional metabolic instability on a protein of interest.

[36] T. R. Butt, S. Jonnalagadda, B. P. Monia, E. J. Sternberg, J. A. Marsh, J. M. Stadel, D. J. Ecker, and S. T. Crooke, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 2540 (1989).

[37] D. J. Ecker, J. M. Stadel, T. R. Butt, J. A. Marsh, B. P. Monia, D. A. Powers, J. A. Gorman, P. E. Clark, F. Warren, and A. Shatzman, *J. Biol. Chem.* **264**, 7715 (1989).

[38] P. Mak, D. P. McDonnell, N. L. Weigel, W. T. Schrader, and B. W. O'Malley, *J. Biol. Chem.* **264**, 21613 (1989).

[39] R. T. Baker, S. A. Smith, R. Marano, J. McKee, and P. G. Board, *J. Biol. Chem.* **269**, 25381 (1994).

[40] Y. Yoo, K. Rote, and M. Rechsteiner, *J. Biol. Chem.* **264**, 17078 (1989).

[41] A. Pilon, P. Yost, T. E. Chase, G. Lohnas, T. Burkett, S. Roberts, and W. E. Bentley, *Biotechnol. Prog.* **13**, 374 (1997).

[42] A. Bachmair and A. Varshavsky, *Cell* **56**, 1019 (1989).

[43] C. P. Hill, N. L. Johnston, and R. E. Cohen, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 4136 (1993).

[44] V. Chau, J. W. Tobias, A. Bachmair, D. Marriott, D. J. Ecker, D. K. Gonda, and A. Varshavsky, *Science* **243**, 1576 (1989).

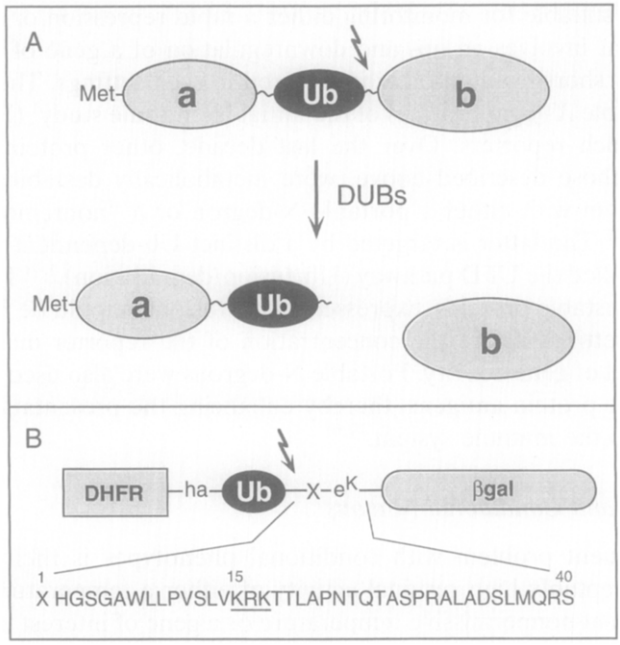

FIG. 3. The UPR (ubiquitin/protein/reference) technique.[79] (A) A tripartite fusion containing **a**, the reference protein moiety whose C terminus is linked, via a spacer peptide, to the Ub moiety. The C terminus of Ub is linked to **b**, a protein of interest. *In vivo*, this tripartite fusion is cotranslationally cleaved[9] by deubiquitylating enzymes (DUBs) at the Ub–**b** junction, yielding equimolar amounts of the unmodified protein **b** and **a**–Ub, the reference protein **a** bearing a C-terminal Ub moiety. If **a**–Ub is long-lived, a measurement of the ratio of **a**-Ub to **b** as a function of time or at steady state yields, respectively, the *in vivo* decay curve or the relative metabolic stability of protein **b**.[29,79] (B) Example of a specific UPR-type Ub fusion.[29] This fusion contains the following elements: DHFRha, a mouse dihydrofolate reductase (DHFR) moiety extended at the C terminus by a sequence containing the hemagglutinin-derived ha epitope; the Ub moiety (more specifically, the Ub^{R48} moiety bearing the Lys → Arg alteration at position 48); a 40-residue, *E. coli* Lac repressor-derived sequence, termed e^K [extension (*e*) containing lysines (*K*)] and shown below in single-letter abbreviations for amino acids; a variable residue *X* between Ub and e^K; the *E. coli* βgal moiety lacking the first 24 residues of wild-type β-Gal. The lightning arrow indicates the site of *in vivo* cleavage by DUBs.[29]

N-Degron and Reporter Proteins

A change in the physiological state of a cell that is preceded or followed by the induction or repression of specific genes can be monitored through the use of promoter fusions to a variety of protein reporters, such as, for example, βgal, β-glucuronidase, luciferase, and green fluorescent protein (GFP). A long-lived reporter is useful for detecting the induction of genes,

but is less suitable for monitoring either a rapid repression or a temporal pattern that involves an up- and downregulation of a gene of interest. A sufficiently short-lived reporter is required in such settings. The metabolically unstable *X*–βgal proteins of the initial N-end rule study[1] (Fig. 2) were the first such reporters. Over the last decade, other protein reporters, including those described above, were metabolically destabilized by extending them with either a portable N-degron or a "nonremovable" Ub moiety.[45–47] The latter is targeted by a distinct Ub-dependent proteolytic pathway called the UFD pathway (Ub/fusion/degradation).[1,48] These metabolically unstable proteins, expressed as Ub fusions, should be particularly useful in settings where the concentration of the reporter must reflect a recent level of gene activity. Portable N-degrons were also used to destabilize specific protein antigens, thereby enhancing the presentation of their peptides to the immune system.[49,50]

N-Degron and Conditional Mutants

A frequent problem with conditional phenotypes is their leakiness, i. e., unacceptably high residual activity of either a temperature-sensitive (*ts*) protein at nonpermissive temperature or a gene of interest in the "off" state of its promoter. Another problem is "phenotypic lag," which often occurs between the imposition of nonpermissive conditions and the emergence of a relevant null phenotype. Phenotypic lag tends to be longer with proteins that are required in catalytic rather than stoichiometric amounts.

In one application of Ub fusions and the N-end rule pathway to the problem of phenotypic lag, a constitutive N-degron (produced as a Ub fusion) was linked to a protein expressed from an inducible promoter.[51] This method is constrained by the necessity of using a heterologous promoter and by the constitutively short half-life of a target protein, whose levels may therefore be suboptimal under permissive conditions. An alternative approach is to link the N-degron to a normally long-lived protein in a strain in which the N-end rule pathway can be induced or repressed. Such strains have been constructed with *S. cerevisiae*,[52,53] but can also be designed in

[45] C. K. Worley, R. Ling, and J. Callis, *Plant Mol. Biol.* **37**, 337 (1998).
[46] H. Deichsel, S. Friedel, A. Detterbeck, C. Coyne, U. Hamker, and H. K. MacWilliams, *Dev. Genes Evol.* **209**, 63 (1999).
[47] I. Paz, J.-R. Meunier, and M. Choder, *Gene* **236**, 33 (1999).
[48] E. S. Johnson, P. C. Ma, I. M. Ota, and A. Varshavsky, *J. Biol. Chem.* **270**, 17442 (1995).
[49] A. Townsend, J. Bastin, K. Gould, G. Brownlee, M. Andrew, B. Coupar, D. Boyle, S. Chan, and G. Smith, *J. Exp. Med.* **168**, 1211 (1988).
[50] T. Tobery and R. F. Siliciano, *J. Immunol.* **162**, 639 (1999).
[51] E. C. Park, D. Finley, and J. W. Szostak, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 1249 (1992).
[52] Z. Moqtaderi, Y. Bai, D. Poon, P. A. Weil, and K. Struhl, *Nature* (*London*) **383**, 188 (1996).
[53] M. Ghislain, R. J. Dohmen, F. Levy, and A. Varshavsky, *EMBO J.* **15**, 4884 (1996).

other species, including mammalian cells. The metabolic stabilities, and hence also the levels of N-degron-bearing proteins, in a cell with an inducible N-end rule pathway are either normal or low, depending on whether Ubr1p, the recognition (E3) component of the N-end rule pathway, is absent or present.[52,53] These conditional mutants can be constructed with any cytosolic or nuclear protein whose function tolerates an N-terminal extension.

Yet another design is a portable N-degron that is inactive at a low (permissive) temperature but becomes active at a high (nonpermissive) temperature. Such an N-degron was constructed, using the Ub fusion technique, from a specific *ts* allele of the 20-kDa mouse dihydrofolate reductase (DHFR) bearing the N-terminal arginine, a strongly destabilizing residue.[54] Linking this DHFR-based, heat-inducible N-degron to proteins of interest yielded a new class of *ts* mutants, called *td* (temperature-activated degron). The *td* method does not require an often unsuccessful search for a *ts* mutation in a gene of interest. If the corresponding protein can tolerate N-terminal extensions, the corresponding *td* fusion is functionally unperturbed at permissive temperature. In contrast, low activity of a *ts* protein at permissive temperature is a frequent problem with conventional *ts* mutants. The *td* method eliminates or reduces the phenotypic lag, because the activation of N-degron results in rapid disappearance of a *td* protein. Another advantage of the *td* technique is the possibility of employing two sets of conditions: a *td* protein-expressing strain at permissive versus nonpermissive temperature or, alternatively, the same strain versus a congenic strain lacking the N-end rule pathway, with both strains at nonpermissive temperature.[54] This powerful internal control, provided in the *td* technique by two alternative sets of permissive/nonpermissive conditions, is unavailable with conventional *ts* mutants. Since 1994, a few laboratories described successful uses of the *td* method to construct *ts* alleles of specific proteins (e.g., Refs. 55 and 56). A recent modification of the *td* technique combines the galactose-inducible overexulsion of Ubr1p and the temperature-sensitive (*td*) N-degron.[56a]

N-Degron and Conditional Toxins

A major limitation of the current pharmacological strategies stems from the absence of drugs that are specific for two or more independent molecular

[54] R. J. Dohmen, P. Wu, and A. Varshavsky, *Science* **263**, 1273 (1994).
[55] G. Caponigro and R. Parker, *Genes Dev.* **9**, 2421 (1995).
[56] J. Wolf, M. Nicks, S. Deitz, E. van Tuinen, and A. Franzusoff, *Biochem. Biophys. Res. Commun.* **243**, 191 (1998).
[56a] K. Labib, J. A. Tercero, and J. F. X. Diffley, *Science* **288**, 1643 (2000).

targets. For the reasons discussed in detail elsewhere,[57,58] it is desirable to have a therapeutic agent that possesses a multitarget, combinatorial selectivity, which requires the presence of two or more predetermined targets in a cell and simultaneously the absence of one or more targets for the drug to exert its effect. Note that simply combining two or more "conventional" drugs against different targets in a multidrug regimen would not yield the multitarget selectivity, because the two drugs together would perturb not only cells containing both targets but also cells containing either one of the targets.

A strategy for designing protein-based reagents that are sensitive to the presence or absence of more than one target at the same time was proposed in 1995.[57] A key feature of these reagents is their ability to utilize codominance, the property characteristic of many signals in proteins, including degrons and nuclear localization signals (NLSs). Codominance, in this context, refers to the ability of two or more signals in the same molecule to function independently and not to interfere with each other. The critical property of a degron-based multitarget reagent is that its intrinsic toxicity is the same in all cells, whereas its half-life (and, consequently, its steady state level and overall toxicity) in a cell depends on the protein composition of the cell, specifically on the presence of "target" proteins that have been chosen to define the profile of a cell to be eliminated.[57] A related but different design involves a toxic protein made short-lived (and therefore relatively nontoxic) by the presence of a degradation signal such as an N-degron. (The latter is produced by the Ub fusion technique.) If a cleavage site for a specific viral processing protease is placed between the toxic moiety of the fusion and the N-degron, the fusion would be cleaved in virus-infected cells but not in uninfected cells. As a result, the toxic moiety of the fusion would become long-lived (and therefore more toxic) only in virus-infected cells.[59] The codominance concept and the ideas about protein-size multitarget reagents have been extended to small (<1-kDa) multitarget compounds.[58]

Overproduction of Proteins as Ubiquitin Fusions

A major application of the Ub fusion technique is its use to augment the yields of recombinant proteins.[36–39] This approach increases the yield of short peptides as well.[40,41,60] The yield-enhancing effect of Ub was ob-

[57] A. Varshavsky, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 3663 (1995).
[58] A. Varshavsky, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 2094 (1998).
[59] A. Varshavsky, *Cold Spring Harbor Symp. Quant. Biol.* **60**, 461 (1996).
[60] T. H. LaBean, S. A. Kauffman, and T. R. Butt, *Mol. Divers.* **1**, 29 (1995).

served not only with eukaryotic cells (where the Ub moiety is present in a nascent fusion but not in its mature counterpart) but also in prokaryotes, which lack the Ub system, including DUBs, and therefore retain the Ub moiety in a translated fusion.[35-37] (*Escherichia coli* transformed with a plasmid expressing the *S. cerevisiae* DUB Ubp1p acquires the ability to deubiquitylate Ub fusions.[61])

The yield-enhancing effect of Ub stems at least in part from rapid folding of the nascent Ub moiety, whose presence at the N terminus of an emerging polypeptide chain may thereby partially protect the still unfolded chain from attacks by proteolytic pathways of the cytosol (most of these pathways are a part of the Ub system). The remarkably strong increases in protein yield even in eukaryotic cells, where the Ub moiety of the fusion is retained transiently (for it is rapidly removed by DUBs), suggest that this protection by Ub is particularly critical during translation, when an emerging, partially unfolded polypeptide chain may present degrons that are buried in the folded version of the same polypeptide. The chaperone role of Ub in this setting reflects one of its physiological functions. Specifically, the experiments with natural Ub fusions containing ribosomal proteins have shown that the transient presence of Ub in front of a ribosomal protein moiety is required for the efficient incorporation of that moiety into the nascent ribosomes,[12] most likely because of the transient protection effect described above.

The Ub-mediated increase in total yield is often accompanied by an even greater increase in the solubility of overexpressed protein. In this regard, the effect of Ub is analogous to that of several other proteins, such as thioredoxin[62] and maltose-binding protein (MBP).[63] When these moieties are cotranslationally linked to a protein of interest, they often increase its yield and solubility. A model of the underlying mechanism suggested for MBP[63] may also be relevant to the effect of Ub moiety. Specifically, a partially unfolded nascent protein is presumed to weakly interact with the nearby (upstream) MBP moiety, thereby transiently precluding intermolecular self-interactions that could result in irreversible aggregation before the protein has had the time to attain its mature conformation.[63]

The first engineered Ub fusions utilized pUB23-X, a family of high copy plasmids that expressed Ub–*X*–βgal proteins containing different junctional residues (*X*) in *S. cerevisiae* from a galactose-inducible, glucose-repressible promoter.[1,42] Subsequent designs facilitated the construction of ORFs encoding Ub–*X*–polypeptide fusions by introducing a *Sac*II (*Sst*II)

[61] J. W. Tobias, T. E. Shrader, G. Rocap, and A. Varshavsky, *Science* **254**, 1374 (1991).

[62] E. R. LaVallie, E. A. DiBlasio, S. Kovacic, K. L. Grant, P. F. Schendel, and J. M. McCoy, *BioTechnology* **11**, 187 (1993).

[63] R. B. Kapust and D. S. Waugh, *Protein Sci.* **8**, 1668 (1999).

site within the codons for the last three residues of the Ub moiety.[39] In this cloning scheme, an ORF of interest is amplified by polymerase chain reaction (PCR) and a primer in which the 5′ extension encodes the last three residues of Ub. Another cloning route employs double-stranded oligonucleotides with *Sac*II cohesive ends that are used to join the DNA fragments.[39] The expression of a resulting Ub–*X*–polypeptide fusion in a eukaryotic cell (or in a prokaryotic cell that contains the *S. cerevisiae* Ubp1p DUB) yields an *X*–polypeptide bearing a predetermined N-terminal residue *X* (Figs. 1 and 2).

In their natural milieu, proteins of biotechnological or pharmacological interest are often products of the secretory pathway, and therefore are cleaved by signal peptidase on their entrance into the endoplasmic reticulum (ER). This cleavage frequently yields destabilizing residues at the N-termini of these proteins. When the same proteins are overexpressed in the cytosol of a heterologous bacterial or eukaryotic host, their N-terminal methionine tends to be retained, because MAPs cannot cleave off N-terminal methionine if it is followed by a destabilizing residue (see above). It is in these, quite frequent, cases that the expression of a protein as a Ub–*X*–protein fusion attains two aims at once: producing a protein of interest bearing the desired N-terminal residue (Fig. 2) and also, quite often, increasing the yield of the protein, in comparison with an otherwise identical expression of the Ub-lacking protein.[35]

There are numerous examples of Ub-mediated increases in the yield and solubility of overexpressed proteins. For instance, a conventional heterologous expression of the *Streptomyces* tyrosinase in *E. coli* yielded inactive enzyme, whereas expression of tyrosinase as a Ub fusion resulted in an abundant and active enzyme.[64] Another example of the use of Ub fusions in *E. coli* was an abundant expression of the soluble human collagenase catalytic domain. In contrast, the expression of the same protein in the absence of N-terminal Ub moiety resulted in low yield and insoluble product.[65] A 60-fold increase in the yield of the human pi class glutathione transferase GSTP1 was observed on the addition of a Ub-coding sequence to the *GSTP1* ORF.[39] A strong increase in protein yield in *E. coli* was reported with a combination of the T7 RNA polymerase promoter system and Ub fusions.[66] Several other examples of the Ub fusion approach to

[64] K. Han, J. Hong, H. C. Lim, C. H. Kim, Y. Park, and J. M. Cho, *Ann. N.Y. Acad. Sci.* **721**, 30 (1994).

[65] M. R. Gehring, B. Condon, S. A. Margosiak, and C. C. Kan, *J. Biol. Chem.* **270**, 22507 (1995).

[66] M. H. Koken, H. H. Odijk, M. Van Duin, M. Fornerod, and J. H. Hoeijmakers, *Biochem. Biophys. Res. Commun.* **195**, 643 (1993).

protein overexpression[38,67-69] are described in an earlier review by Baker.[35] More recently, Hondred and colleagues applied the Ub fusion technique to augment protein expression in transgenic plants.[70]

Ubiquitin-Assisted Dissection of Protein Translocation across Membranes

A 1994 method called UTA (ubiquitin translocation assay) employs Ub as a kinetic probe in the context of signal sequence-bearing Ub fusions.[8] After emerging from ribosomes in the cytosol, a protein may remain in the cytosol, or may be transferred to compartments separated from the cytosolic space by membranes. With a few exceptions, noncytosolic proteins begin journeys to their respective compartments by crossing membranes that enclose intracellular organelles such as the ER and mitochondria in eukaryotes or the periplasmic space in bacteria. Amino acid sequences that enable a protein to cross the membrane of a compartment are often located at the protein's N terminus. These "signal" sequences[71] are targeted by translocation pathways specific for each compartment. The translocation of a protein across a compartment membrane can start before the synthesis of the protein is completed, resulting in docking of the still translating ribosome at the transmembrane channel. The UTA technique takes advantage of the rapid (cotranslational) cleavage of a Ub fusion to examine temporal aspects of protein transport across the ER membrane in living cells.[8] Specifically, if a Ub fusion that has been engineered to bear an N-terminal signal sequence (SS) upstream of the Ub moiety is cleaved in the cytosol by DUBs, the fusion's reporter moiety would fail to be translocated into the ER. Conversely, if a nascent SS mediates the docking of a translating ribosome at the transmembrane channel rapidly enough, or if the fusion Ub moiety is located sufficiently far downstream of the SS, then by the time the Ub moiety emerges from the ribosome the latter is already docked, and the nascent Ub moiety enters the ER before it can fold and/or be targeted by DUBs. Thus, the cleavage at the Ub moiety of an SS-bearing Ub fusion in the cytosol can serve as an *in vivo* kinetic marker and a tool for analyzing targeting in protein translocation.[8] The temporal sensitivity of the UTA technique stems from rapid folding of the nascent Ub moiety

[67] E. A. Sabin, C. T. Lee-Ng, J. R. Shuster, and P. J. Barr, *Bio Technology* **7**, 705 (1989).
[68] E. Rian, R. Jemtland, O. K. Olstad, J. O. Gordeladze, and K. M. Gautvik, *Eur. J. Biochem.* **213**, 641 (1993).
[69] M. Coggan, R. Baker, K. Miloszewski, G. Woodfield, and P. Board, *Blood* **9**, 2455 (1995).
[70] D. Hondred, J. M. Walker, D. E. Mathews, and R. D. Vierstra, *Plant Physiol.* **119**, 713 (1999).
[71] G. Blobel, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 1496 (1980).

that precludes its translocation and makes it a substrate of DUBs in the cytosol shortly after the emergence of the fusion Ub moiety from the ribosome.

Split-Ubiquitin Sensor for Detection of Protein–Protein Interactions

Another Ub-based method, termed the split-Ub sensor or USPS (Ub/split/protein/sensor), makes it possible to detect and monitor a protein–protein interaction as a function of time, at the natural sites of this interaction in a living cell.[72] These capabilities of the split-Ub technique distinguish it from the two-hybrid assay.[73] The design of a split-Ub sensor is based on the following observations: when a C-terminal fragment of the 76-residue Ub (C_{ub}) was expressed as a fusion to a reporter protein, the fusion was cleaved by DUBs only if an N-terminal fragment of Ub (N_{ub}) was also expressed in the same cell. This reconstitution of native Ub from its fragments, detectable by the *in vivo* cleavage assay, was not observed with a mutationally altered N_{ub}. However, if C_{ub} and the altered N_{ub} were each linked to polypeptides that interact *in vivo*, the cleavage of the fusion containing C_{ub} was restored, yielding a generally applicable assay for kinetic and equilibrium aspects of the *in vivo* protein interactions.[72]

Enhancement of Ub reconstitution by interacting polypeptides linked to fragments of Ub stems from a local increase in concentration of one Ub fragment in the vicinity of the other. This in turn increases the probability that the two Ub fragments coalesce to form a quasinative Ub moiety, whose (at least) transient formation results in the irreversible cleavage of the fusion by DUBs. This cleavage can be detected readily, and can be followed as a function of time or at steady state.[72,74] Unlike the two-hybrid method, which is based on the apposition of two structurally independent protein domains whose folding and functions do not require direct interactions between the domains, the split-Ub assay involves reconstituting the conformation of a small, single-domain protein. Applications of the split-Ub sensor have shown that this assay is capable of detecting transient *in vivo* interactions such as the binding of a signal sequence of a translocated protein to Sec62p, a component of the ER channel.[74] Different reporter readouts and selection-based screens have been devised for the split-Ub assay, making it possible to use this method for identifying the *in vivo* ligands of a protein

[72] N. Johnsson and A. Varshavsky, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 10340 (1994).
[73] S. Fields and O. Song, *Nature (London)* **340**, 245 (1989).
[74] M. Dünnwald, A. Varshavsky, and N. Johnsson, *Mol. Biol. Cell* **10**, 329 (1999).

of interest, similar to the main application of the two-hybrid assay.[75,76] Split-protein sensors analogous to split-Ub but employing other proteins, such as DHFR, have been developed as well.[77,78]

UPR Technique

Direct measurements of the *in vivo* degradation of intracellular proteins require a pulse–chase assay. It involves the labeling of nascent proteins for a short time with a radioactive precursor ("pulse"), the termination of labeling through the removal of radiolabel and/or the addition of a translation inhibitor, and the analysis of a labeled protein of interest at various times afterward ("chase"), using immunoprecipitation and sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE), or analogous techniques. Its advantage of being direct notwithstanding, a conventional pulse–chase assay is fraught with sources of error. For example, the immunoprecipitation yields may vary from sample to sample; the volumes of samples loaded on a gel may vary as well. If the labeling for specific chase times is done with separate batches of cells (as is the case, e.g., with anchorage-dependent mammalian cell cultures), the efficiency of labeling is yet another unstable parameter of the assay. As a result, pulse–chase data tend to be semiquantitative at best, lacking the means to correct for these errors.

A robust and convenient "internal reference" strategy was described in 1996. This strategy, an extension of the original Ub fusion method, was termed the UPR (ubiquitin/protein/reference) technique.[79] UPR can compensate for several sources of data scatter in a pulse–chase assay (Fig. 3). UPR employs a linear fusion in which Ub is located between a protein of interest and a reference protein moiety (Fig. 3A). The fusion is cotranslationally cleaved by DUBs after the last residue of Ub, producing equimolar amounts of the protein of interest and the reference protein bearing the C-terminal Ub moiety. If both the reference protein and the protein of interest are immunoprecipitated in a pulse–chase assay, the relative amounts of the protein of interest can be normalized against the reference protein in the same sample.[28,29,79] The UPR technique (Fig. 3)

[75] I. Stagljar, C. Korostensky, N. Johnsson, and S. te Heesen, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 5187 (1998).

[76] S. Wittke, N. Lewke, S. Müller, and N. Johnsson, *Mol. Biol. Cell* **10**, 2519 (1999).

[77] I. Remy and S. W. Michnick, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 5394 (1999).

[78] J. N. Pelletier, F. X. Campbell-Valois, and S. W. Michnick, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 12141 (1998).

[79] F. Lévy, N. Johnsson, T. Rumenapf, and A. Varshavsky, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 4907 (1996).

can thus compensate for the scatter of immunoprecipitation yields, sample volumes, and other sources of sample-to-sample variation. The increased accuracy afforded by UPR underscored the insufficiency of the current "half-life" terminology, because the *in vivo* degradation of many proteins deviates from first-order kinetics. For a discussion of this problem and the terminology for describing nonexponential decay, see Refs. 29 and 79.

Ubiquitin Sandwich Technique

Nascent polypeptides emerging from the ribosome may, in the process of folding, present degradation signals similar to those recognized by the Ub system in misfolded or otherwise damaged proteins. It has been a long-standing question whether a significant fraction of nascent polypeptides is cotranslationally degraded. Determining whether nascent polypeptides are actually degraded *in vivo* has been difficult because at any given time the nascent chains of a particular protein species are of different sizes, and therefore would not form a band on electrophoresis in a conventional pulse–chase assay. The Ub sandwich technique[9] makes it possible to detect cotranslational protein degradation by measuring the steady state ratio of two reporter proteins whose relative abundance is established cotranslationally.

Operationally, the Ub sandwich technique[9] is a three-protein version of the UPR assay.[79] A polypeptide to be examined for cotranslational degradation, termed **B**, is sandwiched between two stable reporter domains **A** and **C** in a linear fusion protein. The three polypeptides are connected via Ub moieties to create a fusion protein of the form **A**Ub–**B**Ub–**C**Ub. The independent polypeptides **A**Ub, **B**Ub, and **C**Ub that result from the cotranslational cleavage of **A**Ub–**B**Ub–**C**Ub by DUBs are called modules. The DUB-mediated cleavage establishes a kinetic competition between two mutually exclusive events during the synthesis of the **A**Ub–**B**Ub–**C**Ub fusion: cotranslational UBP cleavage at the **B**Ub–**C**Ub junction to release the long-lived **C**Ub module or, alternatively, cotranslational degradation of the entire **B**Ub–**C**Ub nascent chain by the 26S proteasome. In the latter case, the processivity of proteasome-mediated degradation results in the destruction of the Ub moiety between **B** and **C** before it can be recognized by UBPs. The resulting drop in levels of the **C**Ub module relative to levels of **A**Ub, referred to as the C/A ratio, reflects the cotranslational degradation of domain **B**. This measurement provides a minimal estimate of the total amount of cotranslational degradation, because nonprocessive cotranslational degradation events that do not extend into the **C** domain are not detected. The Ub sandwich method was used to demonstrate that more than

50% of nascent protein molecules bearing an N-degron can be degraded cotranslationally in *S. cerevisiae*, never reaching their mature size before their destruction by processive proteolysis.[9]

If cotranslational protein degradation by the Ub system is found to be extensive for at least some wild-type proteins (surveys of natural proteins remain to be carried out by this new technique), it could be accounted for as an evolutionary trade-off between the necessity of identifying and destroying degron-bearing mature proteins and the mechanistic difficulty of distinguishing between posttranslationally and cotranslationally presented degrons. Cotranslational protein degradation may also represent a previously unrecognized form of protein quality control, which destroys nascent chains that fail to fold correctly. These and other questions about physiological aspects of the cotranslational protein degradation can now be addressed directly in living cells through the Ub sandwich technique.[9]

Concluding Remarks

The Ub fusion technique is made possible by the ability of DUBs to cleave a Ub fusion *in vivo* or *in vitro* after the last residue of Ub irrespective of the flanking sequence context. Since its development, the Ub fusion technique has given rise to a number of applications whose common feature is utilization of the rapid and highly specific cleavage of a Ub-containing fusion by DUBs. Among these applications is the UPR technique, which increases the accuracy of pulse–chase and analogous measurements. I hope that the use of UPR will spread, supplanting the conventional, far less accurate pulse–chase protocols that lack a reference protein. The Ub sandwich technique, a descendant of UPR, has made it possible to determine the extent of cotranslational protein degradation *in vivo* for any protein of interest. One important feature of the Ub moiety is its ability, as a part of linear fusions, to increase the yields and solubility of overexpressed proteins or short peptides in either eukaryotic or bacterial hosts. In yet another class of Ub-based applications, the demonstrated coalescence of peptide-size Ub fragments into a quasinative Ub fold has yielded the split-Ub sensor for detecting protein interactions *in vivo*. Ub fusions continue to be useful in a remarkable variety of ways.

Acknowledgments

I am most grateful to the former and current members of my laboratory, whose work made possible some of the advances described in this review. I thank Daniel Finley (Harvard Medical School) and Rohan Baker (Australian National University) for their comments on the manuscript. Our studies are supported by grants from the National Institutes of Health (GM31530 and DK39520).