

RGS4 Is Arginylated and Degraded by the N-end Rule Pathway *in Vitro**

Received for publication, February 28, 2000, and in revised form, April 22, 2000
Published, JBC Papers in Press, April 26, 2000, DOI 10.1074/jbc.M001605200

Iliia V. Davydov and Alexander Varshavsky‡

From the Division of Biology, California Institute of Technology, Pasadena, California 91125

The N-end rule relates the *in vivo* half-life of a protein to the identity of its N-terminal residue. We used an expression-cloning screen to search for mouse proteins that are degraded by the ubiquitin/proteasome-dependent N-end rule pathway in a reticulocyte lysate. One substrate thus identified was RGS4, a member of the RGS family of GTPase-activating proteins that down-regulate specific G proteins. A determinant of the RGS4 degradation signal (degron) was located at the N terminus of RGS4, because converting cysteine 2 to either glycine, alanine, or valine completely stabilized RGS4. Radiochemical sequencing indicated that the N-terminal methionine of the lysate-produced RGS4 was replaced with arginine. Since N-terminal arginine is a destabilizing residue not encoded by RGS4 mRNA, we conclude that the degron of RGS4 is generated through the removal of N-terminal methionine and enzymatic arginylation of the resulting N-terminal cysteine. RGS16, another member of the RGS family, was also found to be an N-end rule substrate. RGS4 that was transiently expressed in mouse L cells was short-lived in these cells. However, the targeting of RGS4 for degradation in this *in vivo* setting involved primarily another degron, because N-terminal variants of RGS4 that were stable in reticulocyte lysate remained unstable in L cells.

A multitude of regulatory circuits, including those that control the cell cycle, cell differentiation, and responses to stress, involve metabolically unstable proteins (1–5). A short *in vivo* half-life of a regulator provides a way to generate its spatial gradients and allows for rapid adjustments of its concentration, or subunit composition, through changes in the rate of its synthesis or degradation. Damaged or otherwise abnormal proteins tend to be short-lived as well (6). Features of proteins that confer metabolic instability are called degradation signals or degrons (7, 8). The essential component of one degradation signal, called the N-degron, is a destabilizing N-terminal residue of a protein (9). A set of amino acid residues that are destabilizing in a given cell yields a rule, called the N-end rule, which relates the *in vivo* half-life of a protein to the identity of its N-terminal residue. The N-end rule pathway is present in all organisms examined, from mammals and plants to fungi and prokaryotes (10, 11).

* This work was supported by Grants GM31530 and DK39520 (to A. V.) from the National Institutes of Health. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed: Division of Biology, 147–75, California Institute of Technology, 1200 E. California Blvd., Pasadena, CA 91125. Tel.: 626-395-3785; Fax: 626-440-9821; E-mail: avarsh@caltech.edu.

In eukaryotes, an N-degron consists of two determinants, a destabilizing N-terminal residue and an internal lysine or lysines (12–15). The Lys residue is the site of formation of a multiubiquitin chain (16). The N-end rule pathway is thus one of the pathways of the ubiquitin (Ub)¹ system. Ub is a 76-residue protein whose covalent conjugation to other proteins plays a role in a vast range of biological processes (4, 5, 17). In most of them, Ub acts through routes that involve the degradation of ubiquitylated² proteins by the 26 S proteasome, an ATP-dependent multisubunit protease (18, 19).

The N-end rule has a hierarchic structure. In the yeast *Saccharomyces cerevisiae*, Asn and Gln are tertiary destabilizing N-terminal residues in that they function through their deamidation, by the NTA1-encoded³ N-terminal amidase (Nt-amidase), to yield the secondary destabilizing residues Asp and Glu (20). The destabilizing activity of N-terminal Asp and Glu requires their conjugation, by the ATE1-encoded Arg-tRNA-protein transferase (R-transferase), to Arg, one of the primary destabilizing residues (21, 22). The primary destabilizing N-terminal residues are bound directly by UBR1, also called N-recogin, the E3 (recognition) component of the N-end rule pathway (10, 11).

In mammals, the deamidation step is mediated by two Nt-amidases, Nt^N-amidase and Nt^Q-amidase, which are specific, respectively, for N-terminal Asn and Gln (23, 24). In vertebrates, the set of secondary destabilizing residues contains not only Asp and Glu but also Cys, which is a stabilizing residue in yeast (25, 26). The mammalian counterpart of the yeast R-transferase Ate1p exists as two distinct species, ATE1–1 and ATE1–2, that are produced through alternative splicing of *Ate1* pre-mRNA (27). Both ATE1–1 and ATE1–2 are similar in specificity to the ATE1-encoded yeast R-transferase, in that these R-transferases can arginylate N-terminal Asp and Glu, but cannot arginylate N-terminal Cys (27), suggesting the existence of a distinct R-transferase specific for N-terminal Cys.

¹ The abbreviations used are: Ub, ubiquitin; Nt-amidase, N-terminal amidase; R-transferase, Arg-tRNA-protein transferase; Ub-X-nsP4*, Ub-X-nsP4_{1–254}; GST, glutathione transferase; RGS, regulator of G-protein signaling; MetAP, methionine aminopeptidase; GAP, GTPase-activating protein; DUB, deubiquitylating enzyme; AMP-PNP, 5'-adenylylimidodiphosphate; CAPS, 3-(cyclohexylamino)propanesulfonic acid; PAGE, polyacrylamide gel electrophoresis; PCR, polymerase chain reaction; ORF, open reading frame.

² 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 (5), which is also recommended by the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (74), brings ubiquitin-related terms in line with the standard chemical terminology.

³ Throughout the text, the names of genes are italicized and all uppercase. The names of proteins are roman and all uppercase. This usage, a modification of the existing conventions (73), provides a uniform nomenclature in a text that refers to both fungal and metazoan genes and proteins.

UBR1 (N-recogin) of yeast and mammals has two binding sites for the primary destabilizing N-terminal residues of either proteins or short peptides. The type 1 site is specific for the basic N-terminal residues Arg, Lys, and His. The type 2 site is specific for the bulky hydrophobic N-terminal residues Phe, Leu, Trp, Tyr, and Ile (25, 28, 29). UBR1 contains yet another substrate-binding site, which targets proteins bearing internal (non-N-terminal) degrons. These proteins include CUP9 and GPA1 in yeast (30–32) and the encephalomyocarditis virus 3C protease in metazoans (33).

The known functions of the N-end rule pathway include the control of peptide import in *S. cerevisiae*, through the degradation of CUP9, a transcriptional repressor of the peptide transporter PTR2 (30) (this control includes a positive feedback mediated by the type 1 and type 2 sites of UBR1 (31)); the degradation of GPA1, one of two $G\alpha$ proteins in *S. cerevisiae* (32); and the degradation of alphaviral RNA polymerases and other viral proteins in infected metazoan cells (33, 34). Physiological N-end rule substrates were also identified among the proteins secreted into the cytosol of the mammalian cell by intracellular parasites such as the bacterium *Listeria monocytogenes* (35). Selective perturbation of the N-end rule pathway was reported to interfere with mammalian cell differentiation (36, 37) and with limb regeneration in amphibians (38). Studies of the Ub-dependent proteolysis of endogenous proteins in muscle extracts suggested that the N-end rule pathway plays a role in catabolic states that result in muscle atrophy (39).

Until the present work, physiological substrates of Nt-amidases and R-transferases were unknown in either yeast or larger eukaryotes. Engineered N-end rule substrates, including the substrates of Nt-amidases and R-transferases, can be produced *in vivo* through the Ub fusion technique, in which a Ub-X-reporter fusion is cleaved, cotranslationally, after the last residue of Ub by deubiquitylating enzymes (DUBs) (40), yielding a reporter protein bearing the predetermined N-terminal residue X (9, 10, 41).

In the present work, we employed a modification of the cDNA-based sib-selection strategy in a transcription-translation lysate from rabbit reticulocytes (42, 43) to identify putative physiological substrates of the N-end rule pathway. Specifically, we used dipeptides bearing destabilizing N-terminal residues as selective inhibitors of the N-end rule pathway, and we screened for mouse cDNAs that expressed proteins whose relative abundance in the lysate was altered in the presence of relevant dipeptides.

Among the putative N-end rule substrates identified through the use of this approach was mouse RGS4, a GTPase-activating protein (GAP) for specific $G\alpha$ subunits of heterotrimeric G proteins, and a member of the family of RGS (regulator of G protein signaling) proteins (44–50). We discovered that in addition to the expected removal of N-terminal Met from the newly formed RGS4, the resulting N-terminal Cys of RGS4 was arginylated, presumably by a distinct R-transferase whose Cys specificity is different from that of the known R-transferases. Thus modified RGS4 bore N-terminal Arg, a primary destabilizing residue, and was degraded by the N-end rule pathway in reticulocyte lysate.

EXPERIMENTAL PROCEDURES

Plasmids—The plasmids pcDNA3-Ub-X-nsP4_{1–254} (where X is Met, Arg, or Tyr) were used to express Ub-X-nsP4_{1–254} fusions, denoted below as Ub-X-nsP4*, from the phage T7 promoter in the transcription-translation reticulocyte lysate system by Promega (Madison, WI). The notation nsP4_{1–254} refers to the 254-residue N-terminal fragment of the 69-kDa nsP4, the Sindbis virus RNA polymerase (34). The pcDNA3-Ub-X-nsP4_{1–254} plasmids were constructed using a set of open reading frames (ORFs) encoding fusions between Ub and the full-length nsP4 (51). The Ub-X-nsP4 ORFs in pJCEX1 (a gift from Dr. T. Rü-

menapf, Federal Research Center for Virus Diseases of Animals, Tübingen, Germany) were used as templates for PCR with two primers 5'-TTCGGATCCGCCACCATGCAGATCTTCGTGAAGACCCTG-3' and 5'-CCTTCTAGACTATGCGGTGACAACTCAGTGGTAAT-3', to produce DNA fragments encoding Ub-X-nsP4_{1–254} fusions (the underlined sequences corresponded to the start and stop codons, respectively). These fragments were digested with BamHI and XbaI and cloned into the BamHI/XbaI-cut pcDNA3 vector (Invitrogen, Carlsbad, CA).

The plasmid pcDNA3-Ub-Lys-mCL expressed Ub-Lys-mCL, where Lys-mCL was an N-terminally truncated large subunit of m-calpain starting with Lys-10. The primers 5'-ATTCGCGGTGGCAAAGACC-GCGAGGCGGCCGAGGGGCTG-3' and 5'-CCTACTAGTCTATCATAG-GACTGAAAACTCAGCCACGAGAT-3' were used to amplify a DNA fragment from pET24–80k, which contained cDNA encoding the large subunit of rat m-calpain (52) (a gift from Dr. J. S. Elce, Queen's University, Kingston, Ontario, Canada). The resulting fragment was digested with SacII and SpeI and cloned into SacII/SpeI-cut pcDNA3-Ub-Met-nsP4_{1–254}. The plasmids pcDNA3-p94 and pcDNA3-p94_{C129A} were used to express the mouse-specific calpain p94 and its C129A mutant. To produce these constructs, plasmids containing the rat p94 ORF and the p94_{C129A} ORF (53) (gifts from Dr. H. Sorimachi and Dr. K. Suzuki, University of Tokyo, Japan) were PCR-amplified and cloned into the pcDNA3 vector.

The plasmid pcDNA-RGS4 was used for expression of the mouse RGS4 in reticulocyte lysate and for transfection-mediated expression of RGS4 in mouse L cells. This plasmid contained the T7 and CMV promoters, and the mouse RGS4 ORF followed by two stop codons. The RGS4 ORF was PCR-amplified from the plasmid pcDNA3-26-16-11 (see below), using the primers 5'-TTCGGATCCGCCACCATGTGCAAAGG-ACCTGCAGGTCTG-3' and 5'-CCTTCTAGATCATATTAGCCACACTGG-GAGACCAGGGA-3', followed by digestion with BamHI and XbaI and cloning into BamHI/XbaI-cut pcDNA3 (the underlined sequences corresponded to the start codon and two stop codons, respectively). The plasmids pcDNA3-RGS4_{M19A}, pcDNA3-RGS4_{C2G}, pcDNA3-RGS4_{C2V}, pcDNA3_{C2A}, pcDNA3-RGS4_{K3R}, pcDNA3-RGS4_{K3S}, and pcDNA3-RGS4_{R19–205}, which expressed specific RGS4 mutants, were constructed from pcDNA3-RGS4 using PCR-based site-directed mutagenesis (54).

The plasmid pcDNA3-RGS4_{M19A}-GST expressed RGS4_{M19A}-GST, a fusion of RGS4_{M19A} and glutathione transferase (GST). It was constructed through a PCR-mediated fusion of DNA fragments encoding RGS4 and GST, followed by cloning into BamHI/XbaI-cut pcDNA3. PCR-mediated site-directed mutagenesis was then used to introduce the M19A mutation. The plasmid pcDNA3-RGS16, expressing mouse RGS16 from the T7 promoter (55), and a PCR-produced fragment encoding mouse $G\beta_{5L}$ were gifts from Dr. C. K. Chen and Dr. M. I. Simon (California Institute of Technology, Pasadena, CA). The entire coding regions of the final plasmid constructs were verified by DNA sequencing.

In Vitro Transcription-Translation-Degradation System—The TNT Quick-coupled Transcription-Translation System (Promega) contained a rabbit reticulocyte lysate pre-mixed with most of the reaction components necessary to carry out transcription/translation in the lysate, including all of the amino acids except methionine. [³⁵S]Methionine (>1,000 Ci/mmol, Amersham Pharmacia Biotech) was used to label newly formed proteins in the lysate. Proteins labeled with ³H-amino acids were produced in the TNT T7-coupled Reticulocyte Lysate (Promega), a version of the system where the main components of the reaction were supplied separately. The reactions were set up according to the manufacturer's instructions. The reaction mixtures containing dipeptides also contained 0.15 mM bestatin (Sigma), an inhibitor of some aminopeptidases, to reduce degradation of the added dipeptides (28). Stock samples of dipeptides were 0.5 M solutions in 10 mM K-HEPES, pH 7.5. The reactions were incubated at 30 °C, unless stated otherwise; they were terminated by the addition of SDS-sample buffer, heated at 95 °C for 5 min, and fractionated by SDS-PAGE, followed by autoradiography or fluorography. ³⁵S in protein bands was determined using PhosphorImager (Molecular Dynamics, Sunnyvale, CA).

Screening of Mouse cDNA Library of Small Pools in Rabbit Reticulocyte Lysate—Total RNA was isolated from the brains of female mice with the guanidine thiocyanate method (54). Poly(A)⁺ RNA was purified using Oligotex mRNA kit (Qiagen, Valencia, CA) and was used to synthesize oligo(dT)-primed double-stranded cDNA with the SuperScript cDNA synthesis kit (Life Technologies, Inc.). After ligating the DNA fragments to a BstXI adapter (Invitrogen, San Diego) and digesting with NotI, the cDNAs were cloned into BstXI-NotI-cut pcDNA3. The resulting cDNA library was introduced, by electroporation, into *Escherichia coli* SURE cells (Stratagene, La Jolla, CA), and the transformants were frozen in a number of samples. Titration of the library

showed that it contained $3\text{--}4 \times 10^6$ independent clones. Pools of the library cDNAs containing each about 50 clones were prepared as described (42). Individual pools were added to the T7 TNT Quick-coupled Transcription-Translation System (Promega) together with [^{35}S]methionine, in either the presence or the absence of the mixture of 5 mM Arg- β -Ala, 5 mM Trp-Ala and 0.15 mM bestatin. All components of a reaction except the lysate were gently mixed together, followed by the addition of lysate. The reactions were performed in the total volume of 6.25 μl in a 96-well microtiter plate at 30 °C for 3.5 h. These conditions were chosen after preliminary optimization, using Tyr-nsP4* as a test protein. Once a cDNA pool was found to express a protein whose relative abundance was increased in the presence of the dipeptides, the pool was progressively subdivided and retested, until the isolation of a single positive cDNA clone. One of the clones thus isolated, termed pcDNA3-26-16-11, contained as ~ 3 -kilobase pair mouse RGS4 cDNA (see "Results").

N-terminal Radiochemical Sequencing—RGS4_{M19A}-GST was expressed in the TNT T7-coupled Reticulocyte Lysate System (Promega) in the presence of either [^3H]arginine (49 Ci/mmol), [^3H]leucine (155 Ci/mmol), or [^3H]lysine (87 Ci/mmol) (Amersham Pharmacia Biotech), 1 mM Arg- β -Ala, and 0.15 mM bestatin for 30 min at 30 °C, in the total volume of 0.4 ml. The reaction was stopped by the addition of 20 mM AMP-PNP (Sigma), a non-hydrolyzable ATP analog, and the labeled RGS4_{M19A}-GST was partially purified by affinity chromatography on glutathione-Sepharose 4B (Amersham Pharmacia Biotech), followed by SDS-10% PAGE and the electrophoretic transfer of separated proteins in 10% methanol, 10 mM CAPS, pH 11, to Immobilon-P membrane (Millipore, Burlington, MA). The transferred band of RGS4_{M19A}-GST was cut out (its position was determined with respect to stained protein markers) and subjected to multiple cycles of Edman degradation, using the 476A Applied Biosystems sequencer (Perkin-Elmer). The fractions from each cycle were collected, and ^3H in the fractions was determined using a scintillation counter.

Transfections of Mouse L Cells and Pulse-Chase Assay—Transient transfections of mouse L cells and pulse-chase analysis were performed as described previously (56). The affinity-purified polyclonal antibody raised against a peptide near the C terminus of RGS4 was purchased from Santa Cruz Biotechnology (Santa Cruz, CA). In the experiments with proteasome inhibitor MG132 (Calbiochem), it was added to cells, at the final concentration of 50 μM , 45 min before the addition of [^{35}S]methionine and was present throughout the pulse-chase assay.

RESULTS

Degradation of N-end Rule Substrates in a Transcription-Translation System Derived from Reticulocytes—The N-end rule pathway can be selectively inhibited in reticulocyte lysates through the addition of dipeptides bearing either type 1 or type 2 destabilizing N-terminal residues (25, 28). A commercially produced transcription-translation system derived from reticulocyte lysate and containing the phage T7 RNA polymerase was used to express a protein of interest from DNA template in one reaction mixture. A putative N-end rule substrate in the lysate was detected through its increased concentration in the presence of a cognate dipeptide inhibitor of the N-end rule pathway. We began by examining the degradation of model N-end rule substrates, which were derived from nsP4, the 69-kDa Sindbis virus RNA polymerase, a physiological substrate of the mammalian N-end rule pathway that bears N-terminal Tyr (34). The nsP4 protein is produced during Sindbis virus infection through site-specific cleavage of the viral polyprotein precursor. In the present work, the Ub fusion technique (9, 10, 51) was used to synthesize, in reticulocyte lysate, a set of X-nsP4 derivatives that contained the first 254 residues of nsP4 and in addition bore different residues at the Ub-nsP4 junction. These test proteins, Ub-Arg-nsP4₁₋₂₅₄, Ub-Tyr-nsP4₁₋₂₅₄, and Ub-Met-nsP4₁₋₂₅₄, are denoted below as Ub-Arg-nsP4*, Ub-Tyr-nsP4*, and Ub-Met-nsP4*, respectively. Their N-terminal Ub moiety was cotranslationally cleaved off by DUBs in the lysate yielding Tyr-nsP4*, Arg-nsP4*, and Met-nsP4*, respectively. Earlier experiments by T. Rümnapf have shown that these

constructs were N-end rule substrates in the transcription-translation reticulocyte lysate.⁴

When a Ub-X-nsP4* fusion was expressed in reticulocyte lysate and monitored as a function of time, a major band corresponding to Ub-lacking X-nsP4* was observed (Fig. 1, A and C). Although the electrophoretic mobility of X-nsP4* proteins was slightly higher than expected from their predicted molecular mass of 29 kDa, the observed proteins were clearly X-nsP4*, rather than, for example, Ub-X-nsP4*, because the removal of Ub occurs cotranslationally or nearly so (57), and also because SDS-PAGE of the same samples using a more concentrated polyacrylamide gel revealed the ~ 8 -kDa band of labeled free Ub (data not shown). In addition, the degradation patterns of presumed X-nsP4* derivatives of Ub-X-nsP4* fusions conformed to the N-end rule, as shown below.

The metabolically stable Met-nsP4* accumulated rapidly during the first 30 min of the transcription-translation reaction and reached a plateau around 30 min, because the protein synthesis (but not the activity of the N-end rule pathway; see below) ceased in the lysate by that time (Fig. 1, C and D). Thus, in this setting, the incubation times in excess of ~ 30 min were operationally equivalent to the "chase" part of a pulse-chase experiment. In contrast to Met-nsP4*, the relative levels of Arg-nsP4* (Fig. 1, A and B, lanes labeled c (controls)) and Tyr-nsP4* (data not shown) began to decrease after 30 min, reflecting their continuing degradation by the N-end rule pathway. Arg and Tyr are, respectively, a type 1 (basic) and a type 2 (bulky hydrophobic) destabilizing residue in the N-end rule (10). A ladder of higher molecular mass bands, apparently of multiubiquitylated Arg-nsP4*, was observed with Arg-nsP4* (Fig. 1A) but not with the metabolically stable Met-nsP4* (data not shown).

The addition of Arg- β -Ala, a type 1 inhibitor of the N-end rule pathway (25), strongly inhibited the ubiquitylation and degradation of Arg-nsP4* (Fig. 1A, lanes labeled 1). The same dipeptide had no effect on the degradation of Tyr-nsP4*, which bore a type 2 destabilizing N-terminal residue (data not shown). Conversely, Trp-Ala, a type 2 inhibitor, strongly decreased the degradation of Tyr-nsP4* (Fig. 2B) but had little effect on the degradation of Arg-nsP4* (Fig. 1A, lanes labeled 2). At 1 mM, none of the several dipeptides tested, including those bearing type 1, type 2, or type 3 destabilizing N-terminal residues (25), affected the relative band intensity of Met-nsP4* (Fig. 1, C and D), indicating that these dipeptides did not perturb transcription and translation in this system. However, at 10 mM, some of the dipeptides significantly delayed the synthesis of X-nsP4* (data not shown). Therefore the experiments were carried out with dipeptides at the initial concentration of 1 mM. At this concentration, the inhibition of the N-end rule pathway was strong but incomplete. The inhibition of degradation of, for example, a type 1 substrate such as Arg-nsP4* by a cognate dipeptide could be observed through large differences in the amount of a test protein produced by 30 min, the time of cessation of protein synthesis in the lysate, and through even larger differences at 60 min (Fig. 1B).

m-Calpain Bearing N-terminal Lysine Is Not a Substrate of the N-End Rule Pathway—m-Calpain is a calcium-activated cysteine protease composed of an 80-kDa large subunit (mCL) and a 30-kDa small subunit (mCS) (58). The m- and μ -calpains are ubiquitously expressed in metazoan cells. In the presence of Ca^{2+} , calpain subunits undergo autoproteolytic cleavages in their N-terminal regions. The processed calpains have increased activity. The *in vitro* autoproteolytic cleavage of mCL was shown to remove the first 9 residues of mCL, yielding a

⁴ T. Rümnapf and A. Varshavsky, unpublished data.

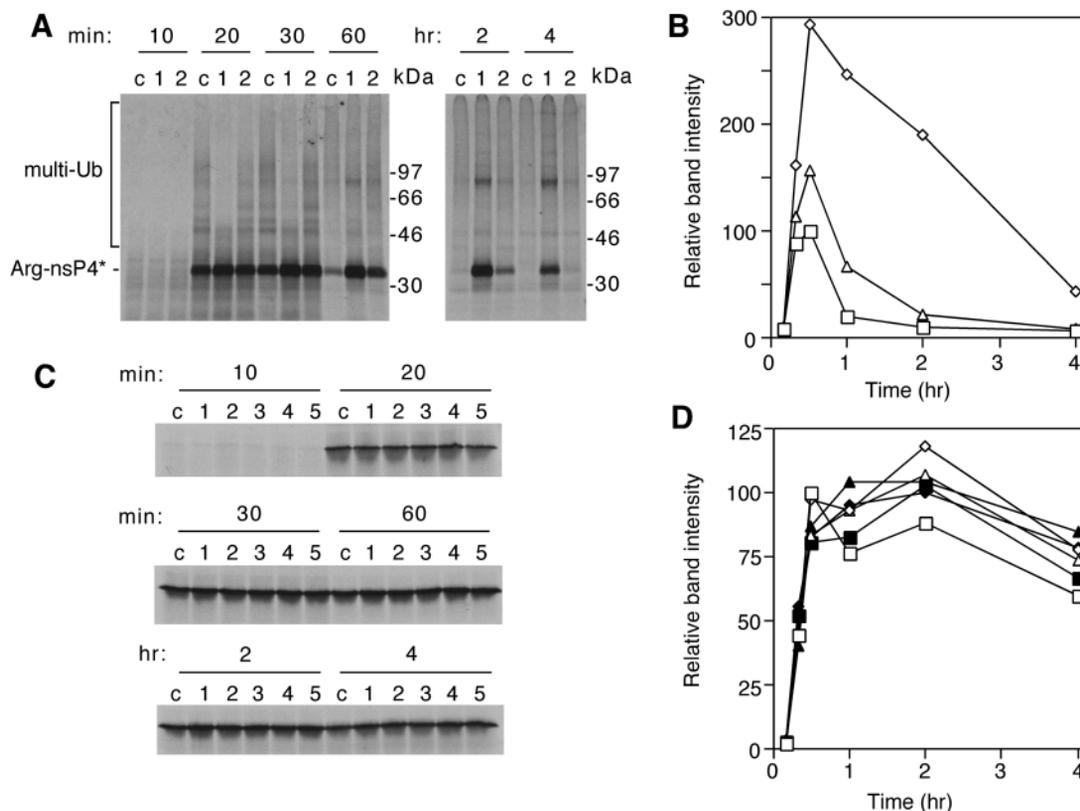


FIG. 1. Effects of dipeptides on the accumulation of Arg-nsP4* and Met-nsP4* in the transcription-translation reticulocyte lysate. Plasmids expressing Ub-Arg-nsP4* (A) or Ub-Met-nsP4* (C) were added to reticulocyte lysate in the presence of [³⁵S]methionine and either in the absence of N-end rule inhibitors (lane c (controls)) or in the presence of either 1 mM Arg-β-Ala (lanes 1), or 1 mM Trp-Ala (lanes 2), or 1 mM Ala-Lys (lanes 3), or 1 mM Lys-Ala (lanes 4), or 1 mM Ala-Arg (lanes 5). The samples in lanes 1–5 also contained 0.15 mM bestatin. The reactions were carried out for 10, 20, 30, 60, 120, and 240 min, followed by SDS-PAGE, autoradiography, and quantitation. The position of R-nsP4* band and the multi-Ub-containing species are indicated on the left. The positions of molecular mass markers are indicated on the right. B, relative amounts of ³⁵S-labeled Arg-nsP4* in A were determined using PhosphorImager and were plotted as a function of incubation time. The amount of Arg-nsP4* in the reaction without dipeptides at 30 min was assigned the value of 100. Control reaction are as follows: in the absence of dipeptides (□); in the presence of Arg-β-Ala (◇); in the presence of Trp-Ala (Δ); in the presence of Ala-Lys (■); in the presence of Lys-Ala (◆); and in the presence of Ala-Arg (▲). D, relative amounts of [³⁵S]Met-nsP4* in C were determined and plotted as in B.

modified subunit, denoted below as Lys-mCL, which bears N-terminal Lys (59), a type 1 destabilizing residue (10). We asked whether Lys-mCL was unstable in reticulocyte lysate and, if so, whether its instability required N-terminal Lys residue. A plasmid encoding the Ub-Lys-mCL fusion was constructed, and the fusion was expressed in the transcription-translation lysate. It was found that under standard conditions (in the absence of added Ca²⁺), the fusion-derived Lys-mCL was metabolically stable in the lysate (Fig. 2A). In the presence of added Ca²⁺ (2 mM), the newly formed Lys-mCL underwent rapid (apparently autolytic) degradation in the lysate. Significantly, this Ca²⁺-induced degradation of Lys-mCL was not inhibited by Arg-β-Ala, a type 1 inhibitor of the N-end rule pathway (Fig. 2). Thus, the Lys-mCL derivative of the large subunit of m-calpain is not a substrate of the N-end rule pathway in reticulocyte lysate, despite the presence of a destabilizing N-terminal residue (see “Discussion”).

The muscle-specific calpain p94 is homologous to the ubiquitous calpains and was shown to be metabolically unstable (53, 60). We expressed rat p94 and its proteolytically inactive mutant p94_{C129A} in reticulocyte lysate. The full-length p94 was rapidly cleaved in the lysate, yielding several fragments, whereas proteolytically inactive p94_{C129A} remained stable (data not shown), in agreement with the earlier evidence (53). Dipeptides Arg-β-Ala and Trp-Ala had no effect on the relative yields of bands corresponding to either the full-length p94 or the products of its autolysis (data not shown). Thus, neither p94 nor its autolytically produced fragments are the substrates of the N-end rule pathway.

Degradation of RGS4 in Reticulocyte Lysate Is Decreased by Type 1 Inhibitors of the N-end Rule Pathway and by a Proteasome Inhibitor—To search for putative N-end rule substrates in a systematic way, we employed a modification of the previously developed *in vitro* screening method that involves the use of small pools of cDNA clones (42, 43). A mouse brain cDNA library was constructed, and about 500 mouse cDNA pools, each containing about 50 cDNAs, were tested, sequentially, in the transcription-translation reticulocyte lysate. Each reaction was carried out in either the absence or the presence of a mixture of two dipeptides, Arg-β-Ala and Trp-Ala, which have been shown to inhibit selectively the degradation of engineered N-end rule substrates in the lysate (Figs. 1, A and B, and 2B). The resulting ³⁵S-labeled proteins were visualized by SDS-PAGE and autoradiography. We searched for protein bands that were selectively enhanced in the presence, but not in the absence, of the dipeptide inhibitors. A cDNA pool containing a putative substrate of the N-end rule pathway was subjected to subcloning, followed by re-analysis in the lysate. This procedure yielded specific cDNAs encoding putative N-end rule substrates.

Thus far, we found 7 cDNAs that encoded proteins whose expression patterns identified them as putative substrates of the N-end rule pathway. Most of these cDNAs encoded N-terminally truncated polypeptides, produced by translation from Met codons that were internal in the corresponding wild type ORFs (data not shown). Although some of these truncated proteins may prove to be physiologically relevant substrates of the N-end rule pathway, we focused at first on a putative N-end

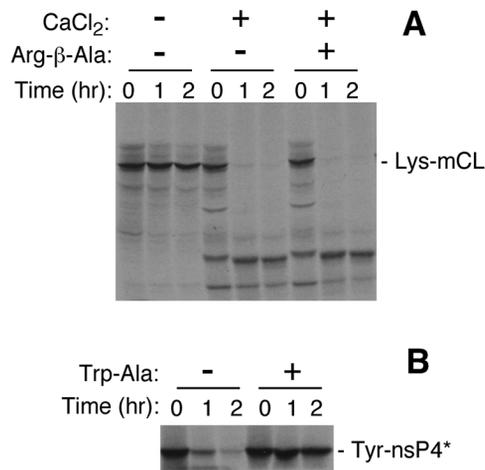


FIG. 2. Lys-mCL is not an N-end rule substrate in reticulocyte lysate. *A*, plasmid DNA encoding a Ub-Lys-mCL, a Ub fusion to a derivative of the large m-calpain subunit bearing N-terminal Lys (see “Results”), was expressed in the transcription-translation reticulocyte lysate in the presence of [³⁵S]methionine. The reaction was allowed to proceed for 20 min at 37 °C (“pulse”). Thereafter, cycloheximide and unlabeled L-methionine were added to all samples, to the final concentrations of 0.1 mg/ml and 0.5 mM, respectively. CaCl₂ and/or Arg-β-Ala were added to some of the samples, to the final concentrations of 2 and 2.5 mM, respectively. 0.15 mM bestatin was present in these samples as well. The reactions mixtures were incubated for the indicated chase times of 0, 1, or 2 h, followed by SDS-PAGE and autoradiography. *B*, plasmid DNA encoding Ub-Tyr-nsP4* was expressed in the lysate, and the reactions were allowed to proceed as in *A* except that Trp-Ala was added, instead of Arg-β-Ala, after the pulse. The positions of (deubiquitylated) Lys-mCL and Tyr-nsP4* bands are as indicated at the right.

rule substrate encoded by a full-length cDNA, which was found to encode the mouse RGS4, a member of the RGS family of GAPs for specific α subunits of G proteins (44, 45). The ~3-kilobase pair *RGS4* mRNA contained the ~2-kilobase pair long 3'-untranslated region. To optimize the expression of RGS4 in reticulocyte lysate, we constructed the plasmid pcDNA3-RGS4, which contained the full-length *RGS4* ORF, followed by two stop codons, and lacked the wild type 3'-untranslated region of *RGS4* mRNA. As with the original *RGS4* cDNA isolate (data not shown), the transcription-translation of pcDNA3-RGS4 yielded two protein bands, of 25 and 22 kDa (Fig. 3, *A* and *D*). The larger protein was the full-length RGS4, whereas the smaller one was produced by translation that started from a downstream Met (AUG) codon at amino acid position 19, because pcDNA3-RGS4_{M19A}, in which the Met-19 codon was converted to an Ala codon, did not express the lower protein band (Fig. 3*D*).

The RGS4 cDNA was initially picked by this screen because upon the incubation times of 2–4 h, the lower (22-kDa) band increased in intensity in the presence, but not in the absence, of dipeptide inhibitors. However, this effect was later found to be unrelated to the N-end rule pathway, as it was observed in the presence of all dipeptides tested, including those that did not inhibit the pathway (Fig. 3*A*). In contrast, at the incubation times from 20 min to 1 h, the relative amount of the full-length RGS4 increased up to 5-fold in the presence of either Arg-β-Ala or Lys-Ala, the type 1 inhibitors of the N-end rule pathway (Fig. 3, *A* and *C*), but not in the presence of Trp-Ala (type 2 inhibitor of the same pathway), or Ala-Lys and Ala-Arg (type 3 inhibitors) RGS4 (Fig. 4). A ladder of high molecular mass bands, apparently of multiubiquitylated RGS4, was observed above the RGS4 and RGS4_{M19A} bands at 10, 20, and 30 min of incubation (Fig. 3, *A* and *B* and data not shown). Arg-β-Ala and Lys-Ala, but not the other tested dipeptides, delayed the appearance of multiubiquitylated RGS4 (Fig. 3, *A* and *B*). The

temporal pattern of inhibition of degradation of RGS4 by dipeptide inhibitors was similar to that described (and explained) above for the engineered N-end rule substrate Arg-nsP4* (Fig. 1, *A* and *B*). Taken together, these data suggested that RGS4 was a type 1 substrate of the N-end rule pathway in reticulocyte lysate.

At longer reaction times (either 2 or 4 h) the effect of type 1 N-end rule inhibitors on the relative level of RGS4 was not observed (Fig. 3, *A* and *C*, and Fig. 4). By contrast, the band of a model N-end rule substrate such as Arg-nsP4* was stronger in the presence of Arg-β-Ala inhibitor at all reaction times, including 2 and 4 h (Fig. 1*B* and Fig. 4). This difference in the patterns of inhibition of RGS4 and Arg-nsP4* degradation could be explained if, in contrast to Arg-nsP4*, there were two pools of the newly formed RGS4 molecules, one of which underwent rapid degradation, while the other remained long-lived and accumulated during the chase, after ~30 min of incubation. The fraction of short-lived RGS4 would be strongly increased by the type 1 inhibitors Arg-β-Ala or Lys-Ala at the early times of incubation (before ~1 h), but over longer times (from 1 to 4 h) this degradation-susceptible form of RGS4 would still be degraded in the lysate, presumably because of incomplete inhibition of the N-end rule pathway by dipeptides, in addition to gradual destruction of the dipeptides by peptidases in the lysate. Thus, at reaction times of 2 h or longer, only the subset of long-lived RGS4 molecules remained.

The fraction of short-lived RGS4 comprised 75–80% of the total RGS4; the remainder of RGS4 molecules (15–25%) was long-lived. This estimate stemmed from the fact that at 30 min, the amount of RGS4 was 4 to 5 times higher in the presence of Arg-β-Ala than in its absence (Fig. 3*C*, Fig. 4, and data not shown). (That the effect of Arg-β-Ala did not result from its influence on the rates of transcription/translation in the lysate was indicated by control experiments with the long-lived Met-nsP4* (Fig. 1, *C* and *D*).) In contrast to RGS4, most Arg-nsP4* molecules were short-lived in the lysate. As a result, the relative amount of Arg-nsP4* was higher in the presence of Arg-β-Ala than in its absence throughout the 4-h incubation (Fig. 4), although the amount of Arg-nsP4* continued to decline after 1 h even in the presence of Arg-β-Ala (Fig. 1), for the reasons described above.

We also examined the effect of a proteasome inhibitor, MG115 (61), on the degradation of RGS4 and Arg-nsP4* in the lysate. MG115 at 10 μM markedly increased the levels of both RGS4 and Arg-nsP4* at the incubation times shorter than 1 h (Fig. 5, *A–D*). MG115 at 0.1 mM further increased the amounts of RGS4 and Arg-nsP4* at these incubation times (Fig. 5, *A–D*). In addition, the ladders of apparently multiubiquitylated RGS4 and Arg-nsP4* were enhanced in the presence of MG115 (Fig. 5, *A* and *C*), in contrast to the effect of N-end rule inhibitors, which decreased the relative level of multi-Ub ladders (Fig. 3*B*). These patterns were consistent with MG115 inhibiting post-ubiquitylation steps of the proteasome-mediated degradation of RGS4 and Arg-nsP4*, in contrast to the action of dipeptides, which inhibited pre-ubiquitylation steps (25, 61).

Similarly to the effects of dipeptide inhibitors (Fig. 1), the proteasome level inhibition by MG115 was also incomplete and resulted in a gradual decrease of the RGS4 and Arg-nsP4* levels after 30 min of incubation (Fig. 5). Note that this decrease led to the nearly complete disappearance of Arg-nsP4* by 4 h of incubation, whereas in the case of RGS4 the decrease stopped when 20–25% of RGS4 still remained in the lysate (Fig. 5, *B* and *D*). These data provided independent evidence for the existence of a degradation-resistant fraction of RGS4. In summary, we identified mouse RGS4 as a type 1 substrate of the N-end rule pathway in the reticulocyte lysate.

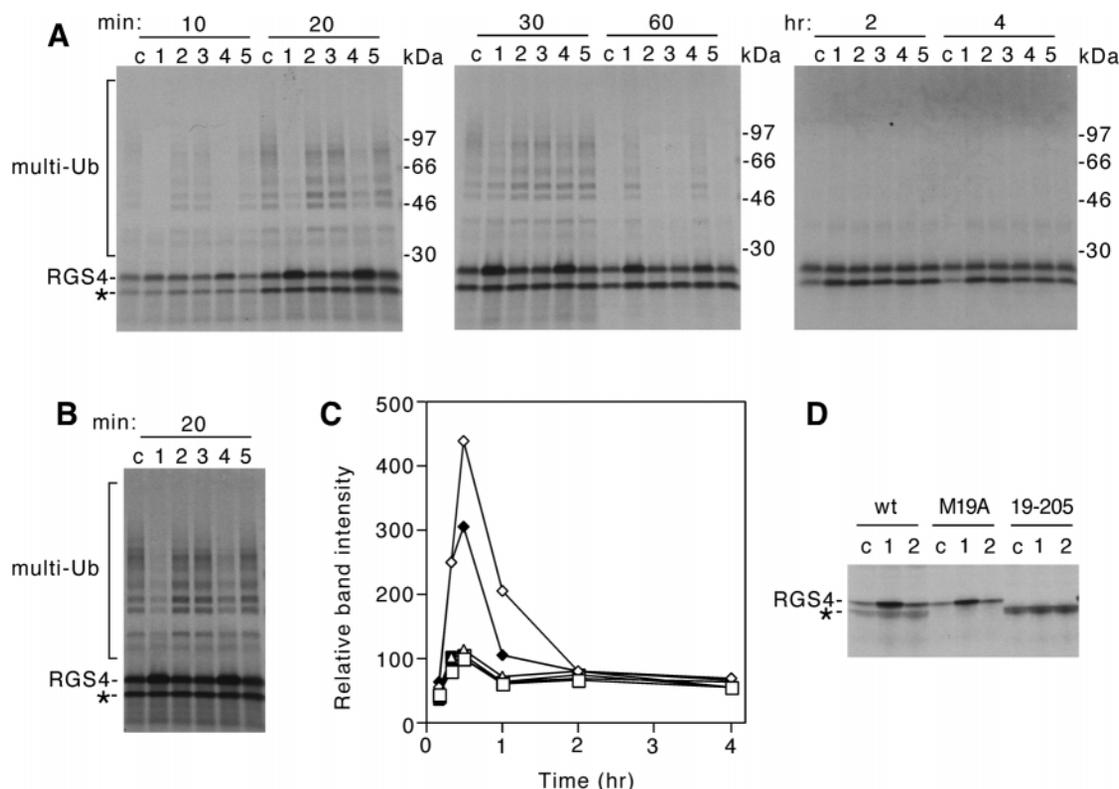


FIG. 3. Type 1 dipeptides Arg- β -Ala and Lys-Ala inhibit degradation and ubiquitylation of RGS4 in reticulocyte lysate. *A*, the plasmid pcDNA3-RGS4 (see “Experimental Procedures”) was expressed in the transcription-translation reticulocyte lysate in the presence of [35 S]methionine and either in the absence of dipeptides (lanes *c*) or in the presence of 1 mM Arg- β -Ala (lane 1), 1 mM Trp-Ala (lane 2), 1 mM Ala-Lys (lane 3), 1 mM Lys-Ala (lane 4), and 1 mM Ala-Arg (lane 5). Dipeptide-containing samples also contained 0.15 mM bestatin. The reactions were carried out for 10, 20, 30, 60, 120, and 240 min, followed by SDS-PAGE, autoradiography, and quantitation. The positions of mouse RGS4 and the multi-Ub-containing species are indicated on the left. An asterisk marks the N-terminally truncated RGS4 derivatives, which were produced through alternative translational initiation from the internal start codon at the Met-19 position of RGS4. The positions of molecular mass markers are indicated on the right. *B*, autoradiographic overexposure of the 20-min part of *B* to visualize the ladders of multi-Ub-containing species. *C*, relative amounts of 35 S-RGS4 in *A* were determined using PhosphorImager and were plotted as a function of incubation time. The amount of RGS4 in the reaction without dipeptides at 30 min was assigned the value of 100. Control reactions were as follows: in the absence of dipeptides (\square); in the presence of Arg- β -Ala (\diamond); in the presence of Trp-Ala (Δ); in the presence of Ala-Lys (\blacksquare); in the presence of Lys-Ala (\blacklozenge); and in the presence of Ala-Arg (\blacktriangle). *D*, plasmids encoding either wild type RGS4 (*wt*), or RGS4_{M19A}, or RGS4₁₉₋₂₀₅ were expressed in the lysate either in the absence of dipeptides (lane *c*), or in the presence of 1 mM Arg- β -Ala and 0.15 mM bestatin (lane 1), or in the presence of 1 mM Trp-Ala and 0.15 mM bestatin (lane 2).

The Degradation Signal of RGS4 Is a Cysteine-based N-degron—The N-terminally truncated RGS4, produced through translation initiation at the second (Met-19) start codon of the RGS4 ORF (see above), was not a substrate of the N-end rule pathway, because degradation of this RGS4₁₉₋₂₀₅ derivative in the lysate was not selectively decreased by dipeptide inhibitors of the N-end rule pathway (Fig. 3*A* and data not shown). This finding strongly suggested that the UBR1-dependent degron of the full-length RGS4 was located near its N terminus. The deduced N-terminal sequence of RGS4 is Met-Cys-Lys-Gly-(62). Extensive evidence indicates that cytosolic methionine aminopeptidases (MetAPs), a class of proteases present in all organisms, cleave off N-terminal Met from proteins and short peptides depending largely on the identity of the next residue, which becomes N-terminal after the cleavage. In particular, MetAPs rapidly (cotranslationally) cleave Met off N-terminal sequences beginning with Met-Cys (63–65). Cys is a stabilizing residue in prokaryotes and in the yeast *S. cerevisiae* but a secondary destabilizing residue in the metazoan N-end rule pathway (25, 26). Specifically, the destabilizing activity of N-terminal Cys is tRNA-dependent, strongly suggesting that an Arg-tRNA-protein transferase (R-transferase) conjugates Arg, a primary destabilizing residue, to the N-terminal Cys of an N-end rule substrate (25). This, and the fact that the degradation of RGS4 was inhibited by type 1 but not by type 2 dipeptides (Fig. 3), suggested the following model: the N-terminal

Met of newly formed RGS4 is removed by MetAPs; the resulting N-terminal Cys is arginylated by R-transferase, and the arginylated RGS4 is targeted for degradation by the UBR1-encoded E3 α and the rest of the N-end rule pathway.

One prediction of this model was that a significant fraction of RGS4 in the reticulocyte lysate should contain the N-terminal sequence Arg-Cys-Lys-Gly-, as distinguished from either Cys-Lys-Gly- or the encoded sequence Met-Cys-Lys-Gly-. Another prediction was that a mutational replacement of Cys-2 with a stabilizing residue that still allows the efficient removal of N-terminal Met should make the resulting RGS4 variant long-lived in the lysate. We tested both predictions.

Since the conventional N-terminal sequencing of RGS4 produced in the lysate would have required a major scale up of the reaction protocol, we employed radiochemical sequencing. A plasmid encoding RGS4_{M19A} fused to glutathione transferase (GST) was constructed. The RGS4_{M19A}-GST protein was expressed in the lysate and was found to be degraded by the N-end rule pathway indistinguishably from the unmodified RGS4 (data not shown). The M19A mutation was introduced to preclude the formation of N-terminally truncated RGS4 variant (Fig. 3, *A* and *D*). RGS4_{M19A}-GST was synthesized in the lysate in the presence of Arg- β -Ala (type 1 inhibitor of the N-end rule pathway) and either [3 H]arginine, [3 H]lysine, or [3 H]leucine. The resulting 3 H-labeled RGS4_{M19A}-GST proteins were purified using glutathione-Sepharose affinity chromatog-

raphy and SDS-PAGE. The band of RGS4_{M19A}-GST was subjected to automated Edman degradation, and ³H released in each cycle was determined.

RGS4_{M19A}-GST labeled with [³H]arginine yielded a major peak of ³H in the first Edman cycle, followed by a second peak in the cycle 14, consistent with Arg at positions 1 and 14 (Fig. 6A). Whereas Arg at position 1 would have to be conjugated to RGS4 posttranslationally, as predicted by the above model, the first encoded Arg residue of RGS4 was, in fact, located at the encoded position 14 of arginylated RGS4 (in the sequence frame that begins with Arg-Cys-Lys-Gly-) (Fig. 6A). This experiment was carried out twice, with two independent preparations of [³H]RGS4_{M19A}-GST, and gave the same result.

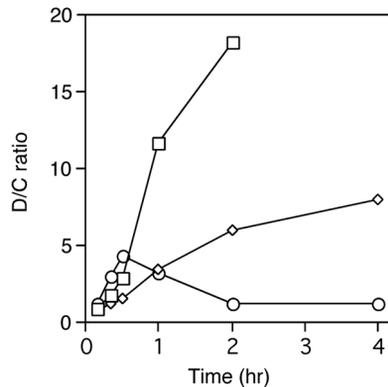


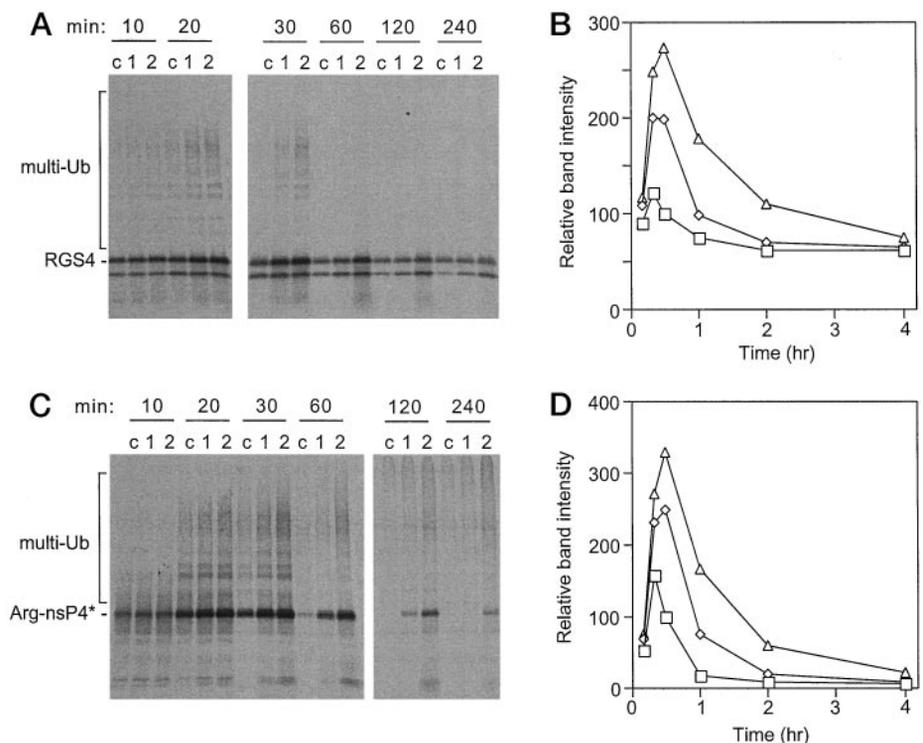
FIG. 4. The relative levels of RGS4 and engineered N-end rule substrates in the presence of cognate dipeptide inhibitors in reticulocyte lysate. ○, the ratios, denoted as dipeptides/controls (*D/C*), of the relative amounts of RGS4 in the presence *versus* the absence of 1 mM Arg-β-Ala, as a function of time in the lysate. These ratios were calculated from the data in Fig. 3C. ◇, the analogous ratio curve but for Tyr-nsP4* in the presence of 1 mM Trp-Ala (primary data not shown). □, the analogous ratio curve but for Arg-nsP4* in the presence of 1 mM Arg-β-Ala (calculated from the data in Fig. 1B; the 4-h ratio could not be determined reliably because of too low ³⁵S in the band of Arg-nsP4* in the absence of dipeptides). Note the monotonous increase of the dipeptides/controls ratio for engineered N-end rule substrate but not for RGS4.

RGS4_{M19A}-GST labeled with [³H]lysine yielded virtually no at position 1 but produced a major peak at position 3 and elevations at positions 17 and 20, consistent with the known positions of three Lys residues in the N-terminal sequence of arginylated RGS4 (Fig. 6B). RGS4_{M19A}-GST labeled with [³H]leucine yielded no ³H in the cycle 1 but produced peaks in the cycles 5 and 8, once again consistent with the known positions of two Leu residues in the N-terminal sequence of arginylated RGS4 (Fig. 6C). Taken together, these independent sets of radiochemical sequencing data indicated that the bulk of RGS4 produced in reticulocyte lysate indeed bore the N-terminal sequence Arg-Cys-Lys-Gly-, with the Arg residue of this sequence not encoded by *RGS4* mRNA.

To verify the second prediction of the model, we constructed several mutant alleles of RGS4 and expressed them in the lysate in either the absence or presence of N-end rule inhibitors (Fig. 7). It was found that the replacement of Cys-2 with either Gly, Val, or Ala residues completely stabilized the resulting RGS4_{C2G}, RGS_{C2V}, and RGS4_{C2A} in the lysate (Fig. 7), in agreement with the model's prediction. We also converted Lys-3 of RGS4 into either Ser or Arg. Interestingly, whereas the Lys → Ser replacement completely stabilized the resulting RGS4_{K3S} variant, the Lys → Arg replacement had no effect on the degradation of RGS4_{K3R} (Fig. 7). Thus, although Lys-3 is not required for ubiquitylation of RGS4 (since it could be replaced with the non-ubiquitylatable Arg), the presence of a basic residue (either Lys or Arg) at position 3 (position 2 after the removal of N-terminal Met) is required for the RGS4 degradation by the N-end rule pathway.

The fact that RGS4_{K3S}, which bears N-terminal Cys and differs from RGS4 exclusively at position 3 (Lys → Ser replacement), was stable in the lysate (Fig. 7) indicated that the presence of N-terminal Cys was not sufficient for making a protein an N-end rule substrate. To explore this issue with other natural proteins, we expressed in the lysate the mouse protein Gβ_{5L}, a member of the family of β subunits of heterotrimeric G proteins (66). The predicted N-terminal sequence of Gβ_{5L} is (Met)-Cys-Asp-Gln-Thr-. We found that Arg-β-Ala, the type 1 inhibitor of the N-end rule pathway, had no effect on the

FIG. 5. Proteasome inhibitor MG115 decreases degradation of RGS4 and Arg-nsP4* in reticulocyte lysate. RGS4 (A) and Arg-nsP4* (C) were expressed in the transcription-translation reticulocyte lysate in the presence of [³⁵S]methionine and either in the absence of MG115 (Me₂SO solvent alone; lanes *c*) or in the presence of 10 μM (lane 1) and 100 μM (lane 2) MG115. The reactions were carried out for 10, 20, 30, 60, 120, and 240 min, followed by SDS-PAGE, autoradiography, and quantitation. The positions of RGS4 and Arg-nsP4* bands and of the ladder of multi-Ub-containing species are indicated on the left. B, relative amounts of ³⁵S-RGS4 in A were determined using PhosphorImager and were plotted as a function of incubation time. The amount of RGS4 in the reaction without MG115 at 30 min was assigned the value of 100. □, control reaction, in the absence of MG115; ◇, in the presence of 10 μM MG115; Δ, in the presence of 100 μM MG115. D, the same as in B but for Arg-nsP4*.



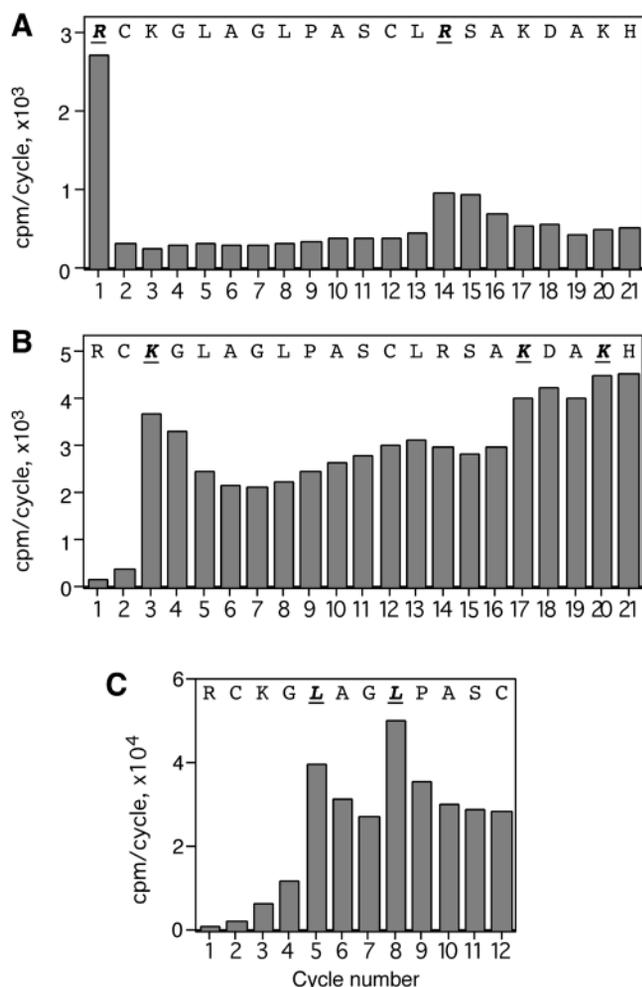


FIG. 6. Radiochemical N-terminal sequencing of RGS4 produced in the reticulocyte lysate. RGS4_{M19A}-GST was synthesized in the transcription-translation reticulocyte lysate for 30 min in the presence of 1 mM Arg- β -Ala and one of the ³H-containing amino acids: [³H]arginine (A), [³H]lysine (B), or [³H]leucine (C). The resulting ³H-labeled RGS4_{M19A}-GST was purified by affinity chromatography and SDS-PAGE, followed by radiochemical N-terminal sequencing through Edman degradation, as described under "Experimental Procedures." Plotted are the amounts of ³H (cpm) recovered in each Edman cycle. The inferred N-terminal sequence of RGS4_{M19A}-GST (see "Results") is shown at the top of each panel, with the expected ³H residues highlighted in *bold*.

pattern of accumulation of G β _{5L} in reticulocyte lysate (data not shown).

The encoded N-terminal sequence of RGS4 is similar to those of RGS5 and RGS16; the similarities include Cys-2 and a basic residue at position 3 (55, 67). To test whether RGS16 was an N-end rule substrate, we expressed RGS16 in reticulocyte lysate. As shown in Fig. 8A, the relative amount of RGS16 was indeed significantly higher in the presence of Arg- β -Ala (type 1 inhibitor) than either in the presence of Trp-Ala (type 2 inhibitor) or in the absence of dipeptides. The temporal pattern of the inhibitor-produced increase in the RGS16 concentration was similar to that described for RGS4 in Fig. 3B (note the 30- and 60-min points in Fig. 8A). The effect of Arg- β -Ala on the accumulation of RGS16 (Fig. 8A), although significant, was smaller than its effect on the accumulation of RGS4 (Fig. 3A). Longer autoradiographic exposures showed the ladders of multi-Ub chains, apparently conjugated to RGS16 (Fig. 8B). Consistent with the weaker effect of Arg- β -Ala, the relative level of RGS16-specific multi-Ub chains (in the absence of inhibitors) was lower than that of RGS4-specific multi-Ub chains

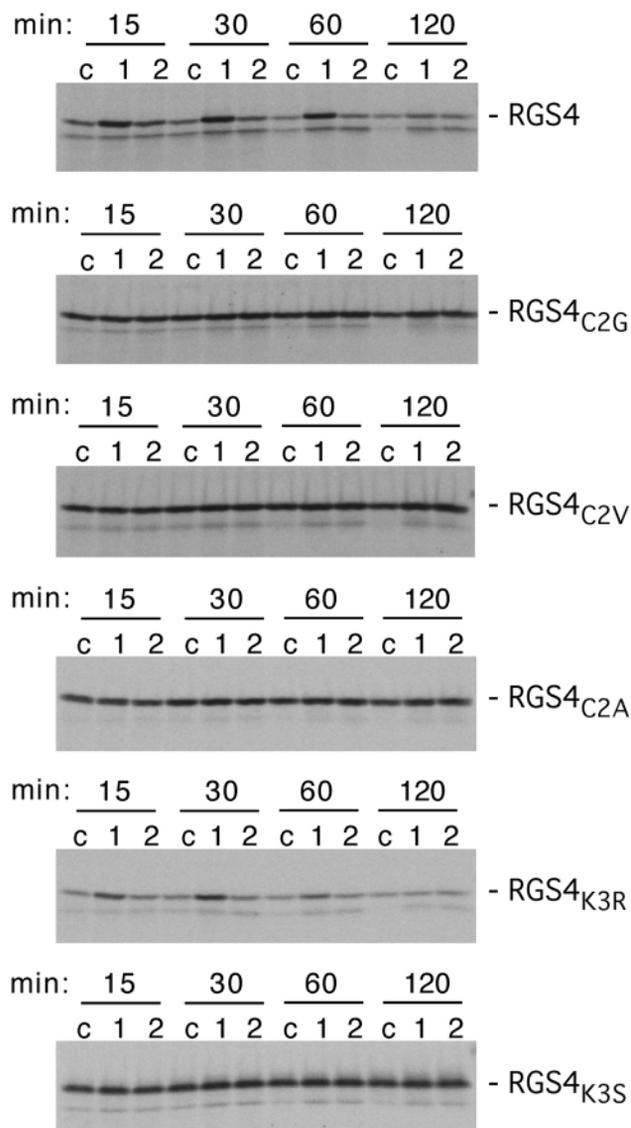


FIG. 7. Single residue substitutions at the N terminus of RGS4 abolish its degradation in reticulocyte lysate. Wild type mouse RGS4 and the indicated single residue derivatives of RGS4p were expressed in the transcription-translation reticulocyte lysate in the presence of [³⁵S]methionine and either in the absence of N-end rule inhibitors (lane c), or in the presence of 1 mM Arg- β -Ala (lane 2), or in the presence of 1 mM Trp-Ala (lane 2). Dipeptide-containing samples also contained 0.15 mM bestatin. The reactions were carried out for 15, 30, 60, and 120 min, followed by SDS-PAGE and autoradiography.

(Figs. 3A and 8B and data not shown). Finally, the ubiquitylation of RGS16 was significantly delayed in the presence of Arg- β -Ala (type 1 inhibitor) but not in the presence of Trp-Ala (type 2 inhibitor) (Fig. 8B). Taken together, these data (Fig. 8) indicated that RGS16 was also a substrate of the N-end rule pathway in reticulocyte lysate.

RGS4 Is Metabolically Unstable in Vivo—A mouse RGS4-expressing plasmid was transiently transfected into mouse L cells, and the metabolic stability of RGS4 was measured in a pulse-chase assay, using an antibody to RGS4. The transiently expressed RGS4 was found to be short-lived in L cells, with the half-life of 40–50 min (Fig. 9, A and B). When L cells were treated with the proteasome inhibitor MG132, both before and during pulse-chase, RGS4 was significantly stabilized (Fig. 9, A and B). Thus, similarly to the results with reticulocyte lysate, the *in vivo* degradation of RGS4 was proteasome-dependent.

To verify whether the degradation of RGS4 *in vivo* required its Cys-based N-degron, L cells were transfected with the plas-

mids expressing the RGS4_{C2G} and RGS4_{C2V} mutants, and the *in vivo* decay curves of these proteins were compared with that of wild type RGS4. It was found that these variants of RGS4, which were completely stable in reticulocyte lysate (Fig. 7), were degraded in L cells similarly to wild type RGS4 (data not shown).

DISCUSSION

We searched for substrates of the mammalian N-end rule pathway by testing some of the previously known candidates and also by employing a modification of the sib selection-based *in vitro* screen (42, 43) in a transcription-translation reticulocyte lysate. Two N-end rule substrates thus identified were mouse RGS4 and RGS16. These proteins are members of the RGS family of GTPase-activating proteins (GAPs) that down-regulate specific G proteins (44–50, 67). We report the following results.

1) The *in vitro* activated, proteolytically processed larger (80 kDa) subunit of the cysteine protease m-calpain was previously shown to bear N-terminal Lys (59), a type 1-destabilizing residue in the N-end rule (10). We expressed this subunit, termed

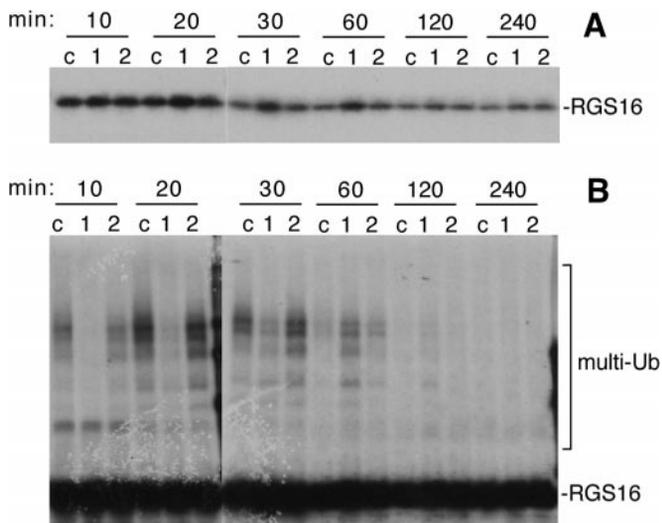
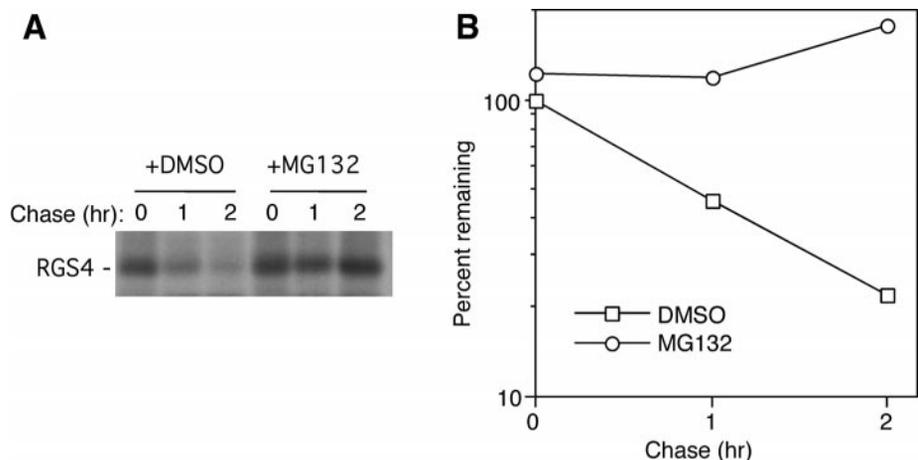


FIG. 8. Type 1 N-end rule inhibitor decreases degradation and ubiquitylation of RGS16 in reticulocyte lysate. A, RGS16 was expressed in the transcription-translation reticulocyte lysate in the presence of [³H]leucine and either in the absence of N-end rule inhibitors (lane c) or in the presence of 1 mM Arg-β-Ala (lane 1) or 1 mM Trp-Ala (lane 2). Dipeptide-containing samples also contained 0.15 mM bestatin. The reactions were carried out for 10, 20, 30, 60, 120, and 240 min, followed by SDS-PAGE and fluorography. B, the same as in A, but with autoradiographic overexposure to visualize the ladders of multi-Ub-containing species.

FIG. 9. *In vivo* degradation of RGS4 in transiently transfected mouse L cells. A, L cells were transfected with an RGS4-expressing plasmid (see “Experimental Procedures”). Cells were preincubated in methionine-free medium for 45 min and then labeled for 12 min with [³⁵S]methionine/cysteine and chased for 0, 1, or 2 h. Either 50 μM MG132 (from a stock solution in Me₂SO) or the equivalent amount of Me₂SO were present during both preincubation in methionine-free medium and the pulse-chase. Cell extracts were immunoprecipitated with anti-RGS4 antibody, followed by SDS-PAGE and autoradiography. B, the pattern in A was quantitated using PhosphorImager. The amount of ³⁵S in RGS4 at time 0 (end of the pulse) was taken as 100%. □, no proteasome inhibitor; circle, in the presence of 50 μM MG132.



Lys-mCL, in the reticulocyte lysate as a Ub-Lys-mCL fusion. The resulting Lys-mCL was not a substrate of the N-end rule pathway, despite the presence of a destabilizing N-terminal residue. Since an N-degron is a bipartite signal (12, 13, 15), the absence of active N-degron from Lys-mCL could be due to the absence of a targetable internal Lys residue, the second determinant of an N-degron. Another possibility is a steric hindrance in the binding of UBR1 to the N-terminal lysine of Lys-mCL. Thus, a destabilizing N-terminal residue is not the sole essential determinant of an active N-degron, as demonstrated previously with engineered N-end rule substrates (12).

2) A major fraction (75–85%) of mouse RGS4 synthesized in reticulocyte lysate was rapidly degraded by a pathway that was proteasome-dependent and apparently also Ub-dependent. The other 15–25% of the newly made RGS4 was found to be stable in the lysate.

3) The degradation of RGS4 was inhibited by dipeptides bearing type 1-destabilizing N-terminal residues (Arg-β-Ala or Lys-Ala) but was unaffected by dipeptides bearing a type 2-destabilizing N-terminal residue (Trp-Ala) or a type 3 residue (Ala-Lys and Ala-Arg).

4) Radiochemical sequencing of RGS4 produced in reticulocyte lysate and labeled with either [³H]arginine, [³H]lysine, or [³H]leucine revealed the presence of posttranslationally conjugated Arg at the N terminus of RGS4. The deduced N-terminal sequence of RGS4 was Met-Cys-Lys-Gly-. The observed/inferred sequence was Arg-Cys-Lys-Gly-. This result and, independently, the fact that degradation of RGS4 was inhibited by type 1 but not by type 2 dipeptides strongly suggested the following model: the N-terminal Met of newly formed RGS4 is removed by MetAPs; the resulting N-terminal Cys is arginylated by R-transferase; and the arginylated RGS4 is targeted for processive degradation by the UBR1-encoded E3α and the rest of the N-end rule pathway.

5) In agreement with a prediction of this model, the degradation of RGS4 in reticulocyte lysate was found to require Cys-2 residue, which becomes N-terminal after the MetAP-mediated cotranslational removal of Met. Specifically, the RGS4 mutants C2G, C2V, and C2A, in which Cys-2 was replaced with Gly, Val or Ala, were completely stable in the lysate. The former two residues are stabilizing in the N-end rule, and neither of the 3 residues is expected to interfere with the removal of N-terminal Met by MetAPs (64).

6) A Lys → Ser replacement at the encoded position 3 of RGS4 also completely stabilized the resulting RGS4_{K3S} protein. However, a Lys → Arg replacement at this position had no effect on the degradation of RGS4_{K3R}. Thus, Lys-3 is not required for ubiquitylation of RGS4. However, the presence of a basic residue (either Lys or Arg) immediately after Cys is

essential for the RGS4 degradation by the N-end rule pathway. One possibility is that the requirement for a basic residue at this position reflects the substrate specificity of an uncharacterized R-transferase that arginylates the N-terminal Cys.

7) The mouse protein G β_{5L} , a member of the family of β subunits of heterotrimeric G proteins (66) that bore the (initial) N-terminal Met-Cys, was tested and found not to be an N-end rule substrate, similarly to the RGS4_{K3S} mutant, which also bore the (initial) N-terminal sequence Met-Cys but was not degraded by the N-end rule pathway. Thus, the N-terminal Cys of wild type RGS4 is an essential determinant of its N-degron, but other N terminus-proximal residues are relevant as well.

8) Similarities among the N-terminal sequences of RGS4, RGS5, and RGS16 (67) suggested that RGS5 and RGS16 may also be N-end rule substrates. This prediction was tested, thus far, with RGS16, and was confirmed.

9) RGS4 was transiently expressed in mouse L cells and found to be an unstable protein, with the half-life of 40–50 min. The *in vivo* degradation of RGS4 was proteasome-dependent, as indicated by the nearly complete stabilization of RGS4 in the presence of proteasome inhibitor MG132. However, the targeting of RGS4 in this *in vivo* setting involved primarily a degron distinct from the Cys-based N-degron, because N-terminal variants of RGS4 such as RGS4_{C2V} and RGS4_{C2G}, which were completely stable in reticulocyte lysate, remained unstable in L cells.

We consider the latter result first. 15–25% of the reticulocyte lysate-produced RGS4 was resistant to degradation by the N-end rule pathway (paragraph 2 above). An analogous protection of RGS4 against targeting by the N-end rule pathway in L cells might involve a much higher fraction of the newly formed RGS4. The mechanism of protection may be a modification of the N-terminal Cys, for example, its palmitoylation (45, 47, 62, 67) or acetylation (68) that would be expected to preclude the arginylation and degradation of RGS4 by the N-end rule pathway. Thus it is possible, indeed likely, that there is a kinetic competition among these reactions at the N-terminal Cys of a nascent RGS4 protein. In experiments to address this model, we added varying amounts of acetyl-CoA or palmitoyl-CoA (the substrates of N-terminal acetylases and palmitoyltransferases) to reticulocyte lysate; no effect of the added compounds on the degradation of newly formed RGS4 in the lysate was observed.⁵ A second possibility is a spatial localization of RGS4 in L cells that protects it from degradation by the N-end rule pathway but leaves RGS4 still targetable by another proteasome-dependent pathway(s) that is inactive in reticulocytes. Yet another possible reason for the difference between the results with reticulocyte lysate *versus* L cells is that the cysteine branch of the N-end rule pathway may be cell type-specific; for example, a Cys-specific R-transferase is present in reticulocytes but might be expressed at a lower level in L cells. In recent experiments, RGS4 was expressed in *Xenopus laevis* oocytes, through microinjection of *RGS4* mRNA. It was found that, similarly to the results with reticulocyte extracts, RGS4 was degraded in oocytes by the cysteine branch of the N-end rule pathway.⁶

One difficulty in identifying physiological N-end rule substrates (as distinguished from those produced through the Ub fusion technique (10)) is caused by the fact that MetAPs remove N-terminal Met from newly formed proteins if, and only if, the second residue in a polypeptide chain is stabilizing in the yeast-type N-end rule. Specifically, the known MetAPs remove N-terminal Met if the second position residues are Gly, Val,

Ala, Ser, Thr, Cys, or Pro (63–65). All of these residues are stabilizing in the yeast-type N-end rule (10). Therefore, a natural substrate of an N-end rule pathway that bears an N-degron in, for instance, *S. cerevisiae* cells can be produced exclusively through cleavages (anywhere along the polypeptide chain) by proteases distinct from MetAPs. An example of N-end rule substrate of this class is the *S. cerevisiae* protein SCC1, a subunit of the cohesin complex that holds together sister chromatids. In a reaction that requires ESP1, the 566-residue SCC1 is cleaved (at the time of sister chromatid separation) at positions 180 and/or 268, resulting in fragments that bear N-terminal Arg, a type 1-destabilizing residue (69). These fragments were found to be degraded by the N-end rule pathway *in vivo*.⁷

The situation in metazoans such as mammals is similar to the one in yeast, in that the yeast-type destabilizing residues cannot be exposed at the N termini of mammalian proteins through cleavages by the known MetAPs (10, 64). The difference here is that Cys, Ala, Ser, Thr, and Pro are stabilizing residues in yeast but destabilizing in mammalian cells (10, 25, 70). In contrast to the other destabilizing residues, the second position Cys, Ala, Ser, Thr, and Pro can be efficiently exposed at the N termini of newly formed proteins through the action of MetAPs. Of these five residues, N-terminal Cys is a special case; its destabilizing activity requires its posttranslational conjugation to Arg, a primary destabilizing residue recognized by the type 1 site of UBR1 (E3 α) (10).

The finding of the dependence of destabilizing activity of N-terminal Cys on the presence of tRNA (25, 26), and the identification, in the present work, of Arg as the posttranslationally linked N-terminal residue of mouse RGS4, indicated the existence of an Arg-tRNA-protein transferase (R-transferase) that conjugates Arg to N-terminal Cys. This R-transferase is distinct from the two previously characterized mammalian R-transferases, both of which are encoded by the alternatively spliced *ATE1* gene (27). The two forms of mouse ATE1 R-transferase have different specific activities but apparently the same substrate specificity; similarly to their *S. cerevisiae* homolog, they arginylate N-terminal Asp and Glu but cannot arginylate N-terminal Cys (27).

The ~20 known proteins of the mammalian RGS family have in common the ~130-residue RGS domain that binds to G α subunits and is responsible for the GAP function of RGS proteins (44, 45). As would be expected of pleiotropic regulators at key junctions of the cellular metabolism, the activity of RGS proteins is controlled at several levels, including regulation of their synthesis and localization (44). It is becoming increasingly clear that the (conditional) degradation of RGSs is yet another way in which the activity of these proteins is modulated in cells. Besides RGS4 and RGS16 of the present work, one other member of the family, RGS7, was recently shown to be unstable *in vivo* (71, 72). The normally short-lived RGS7 is stabilized through its interaction with polycystin, a *PKDI*-encoded protein involved in polycystic kidney disease (71). The degradation of RGS7 is also decreased upon exposure of cells to tumor necrosis factor α ; the resulting increase in the concentration of RGS7 may contribute to sepsis-induced changes in the nervous system (72). The degradation signal(s) of RGS7 remains to be characterized.

Our major results are the findings that RGS4 and RGS16 (and possibly also RGS5) are substrates of the cysteine branch of the N-end rule pathway and that the Cys-based N-degron of these proteins functions through the posttranslational arginylation of N-terminal Cys by an unidentified, apparently Cys-specific R-transferase. RGS4 and RGS16 are the first physio-

⁵ I. Davydov and A. Varshavsky, unpublished data.

⁶ J. Sheng, I. Davydov, and A. Varshavsky, unpublished data.

⁷ H. Rao and A. Varshavsky, unpublished data.

logical substrates of the N-end rule pathway that bear a secondary destabilizing N-terminal residue, which functions through its conjugation to Arg, a primary destabilizing residue.

Specific functions of the metabolic instability of RGS4 and RGS16 remain to be understood. One route to these functions is through the cloning of Cys-specific R-transferase and construction of mouse strains that lack this enzyme and (therefore) the cysteine branch of the N-end rule pathway. It would also be important to define, in detail, Cys-proximal N-terminal sequences in proteins that promote the arginylation of N-terminal Cys. This information will facilitate the identification of Cys-specific N-end rule substrates in data bases of protein sequences.

Acknowledgments—We thank T. Rümnapf for providing Ub-XnsP4-encoding plasmids, which were used as PCR templates in the present work, and for sharing unpublished data on the N-end rule-dependent degradation of truncated Tyr-nsP4 in reticulocyte lysate. We are grateful to G. Hathaway for expert assistance with the radiochemical sequencing of RGS4_{M19A}-GST and to Y. T. Kwon and N. Barteneva for the mouse brains used to prepare the cDNA library. We thank C. K. Chen and M. I. Simon for the gifts of pcDNA3-RGS16 and a Gβ_{5L}-coding DNA fragment; H. Sorimachi and K. Suzuki for the gifts of p94- and p94_{C129A}-encoding plasmids; and J. S. Elce for the gift of pET-80k. We gratefully acknowledge discussions of the paper and comments on the manuscript by E. Smirnova and members of the Varshavsky laboratory.

REFERENCES

- Maniatis, T. (1999) *Genes Dev.* **13**, 505–510
- Nasmyth, K. (1999) *Trends Biochem. Sci.* **24**, 98–104
- Peters, J.-M., King, R. W., and Deshaies, R. J. (1998) in *Ubiquitin and the Biology of the Cell* (Peters, J.-M., Harris, J. R., and Finley, D., eds), pp. 345–387, Plenum Publishing Corp., New York
- Hershko, A., and Ciechanover, A. (1998) *Annu. Rev. Biochem.* **76**, 425–479
- Varshavsky, A. (1997) *Trends Biochem. Sci.* **22**, 383–387
- Wickner, S., Maurizi, M. R., and Gottesman, S. (1999) *Science* **286**, 1888–1893
- Varshavsky, A. (1991) *Cell* **64**, 13–15
- Laney, J. D., and Hochstrasser, M. (1999) *Cell* **97**, 427–430
- Bachmair, A., Finley, D., and Varshavsky, A. (1986) *Science* **234**, 179–186
- Varshavsky, A. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93**, 12142–12149
- Kwon, Y. T., Reiss, Y., Fried, V. A., Hershko, A., Yoon, J. K., Gonda, D. K., Sangan, P., Copeland, N. G., Jenkins, N. A., and Varshavsky, A. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95**, 7898–7903
- Bachmair, A., and Varshavsky, A. (1989) *Cell* **56**, 1019–1032
- Johnson, E. S., Gonda, D. K., and Varshavsky, A. (1990) *Nature* **346**, 287–291
- Pickart, C. M. (1997) *FASEB J.* **11**, 1055–1066
- Suzuki, T., and Varshavsky, A. (1999) *EMBO J.* **18**, 6017–6026
- Chau, V., Tobias, J. W., Bachmair, A., Marriotti, D., Ecker, D. J., Gonda, D. K., and Varshavsky, A. (1989) *Science* **243**, 1576–1583
- Hochstrasser, M. (1996) *Annu. Rev. Genet.* **30**, 405–439
- Baumeister, W., Walz, J., Zühl, F., and Seemüller, E. (1998) *Cell* **92**, 367–380
- DeMartino, G. N., and Slaughter, C. A. (1999) *J. Biol. Chem.* **274**, 22123–22126
- Baker, R. T., and Varshavsky, A. (1995) *J. Biol. Chem.* **270**, 12065–12074
- Balzi, E., Choder, M., Chen, W., Varshavsky, A., and Goffeau, A. (1990) *J. Biol. Chem.* **265**, 7464–7471
- Li, J., and Pickart, C. M. (1995) *Biochemistry* **34**, 15829–15837
- Stewart, A. E., Arfin, S. M., and Bradshaw, R. A. (1995) *J. Biol. Chem.* **270**, 25–28
- Grigoryev, S., Stewart, A. E., Kwon, Y. T., Arfin, S. M., Bradshaw, R. A., Jenkins, N. A., Copeland, N. G., and Varshavsky, A. (1996) *J. Biol. Chem.* **271**, 28521–28532
- Gonda, D. K., Bachmair, A., Wüning, I., Tobias, J. W., Lane, W. S., and Varshavsky, A. (1989) *J. Biol. Chem.* **264**, 16700–16712
- Davydov, I. V., Patra, D., and Varshavsky, A. (1998) *Arch. Biochem. Biophys.* **357**, 317–325
- Kwon, Y. T., Kashina, A. S., and Varshavsky, A. (1999) *Mol. Cell. Biol.* **19**, 182–193
- Reiss, Y., Kaim, D., and Hershko, A. (1988) *J. Biol. Chem.* **263**, 2693–2698
- Baker, R. T., and Varshavsky, A. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **87**, 2374–2378
- Byrd, C., Turner, G. C., and Varshavsky, A. (1998) *EMBO J.* **17**, 269–277
- Turner, G. C., Du, F., and Varshavsky, A. (2000) *Nature* **405**, 579–583
- Schauber, C., Chen, L., Tongaonkar, P., Vega, I., and Madura, K. (1998) *Genes Cells* **3**, 307–319
- Lawson, T. G., Gronros, D. L., Evans, P. E., Bastien, M. C., Michalewich, K. M., Clark, J. K., Edmonds, J. H., Graber, K. H., Werner, J. A., Lurvey, B. A., and Cate, J. M. (1999) *J. Biol. Chem.* **274**, 9871–9880
- deGroot, R. J., Rümnapf, T., Kuhn, R. J., and Strauss, J. H. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88**, 8967–8971
- Sijts, A. J., Pilip, I., and Pamer, E. G. (1997) *J. Biol. Chem.* **272**, 19261–19268
- Obin, M., Mesco, E., Gong, X., Haas, A. L., Joseph, J., and Taylor, A. (1999) *J. Biol. Chem.* **274**, 11789–11795
- Hondermarck, H., Sy, J., Bradshaw, R. A., and Arfin, S. M. (1992) *Biochem. Biophys. Res. Commun.* **30**, 280–288
- Taban, C. H., Hondermarck, H., Bradshaw, R. A., and Boilly, B. (1996) *Experientia (Basel)* **52**, 865–870
- Solomon, V., Baracos, V., Sarraf, P., and Goldberg, A. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95**, 12602–12607
- Wilkinson, K., and Hochstrasser, M. (1998) in *Ubiquitin and the Biology of the Cell* (Peters, J.-M., Harris, J. R., and Finley, D., eds) pp. 99–125, Plenum Publishing Corp., New York
- Baker, R. T. (1996) *Curr. Opin. Biotechnol.* **7**, 541–546
- Lustig, K. D., Stukenberg, P. T., McGarry, T. J., King, R. W., Cryns, V. L., Mead, P. E., Zon, L. I., Yuan, J., and Kirschner, M. W. (1997) *Methods Enzymol.* **283**, 83–99
- Stukenberg, P. T., Lustig, K. D., McGarry, T. J., King, R. W., Kuang, J., and Kirschner, M. W. (1997) *Curr. Biol.* **7**, 338–348
- De Vries, L., and Farquhar, M. G. (1999) *Trends Cell Biol.* **9**, 138–144
- Berman, D. M., and Gilman, A. G. (1998) *J. Biol. Chem.* **273**, 1269–1272
- Heximer, S. P., Srinivasa, S. P., Bernstein, L. S., Bernard, J. L., Linder, M. E., Hepler, J. R., and Blumer, K. J. (1999) *J. Biol. Chem.* **274**, 34253–34259
- Tu, Y., Popov, S., Slaughter, C., and Ross, E. M. (1999) *J. Biol. Chem.* **274**, 38260–38267
- DiBello, P. R., Garrison, T. R., Apanovitch, D. M., Hoffman, G., Shuey, D. J., Mason, K., Cockett, M. I., and Dohlman, H. G. (1998) *J. Biol. Chem.* **273**, 5780–5784
- Yan, Y., Chi, P. P., and Bourne, H. R. (1997) *J. Biol. Chem.* **272**, 11924–11927
- Wieland, T., Chen, C. K., and Simon, M. I. (1997) *J. Biol. Chem.* **272**, 8853–8856
- Lemm, J. A., Rümnapf, T., Strauss, E. G., Strauss, J. H., and Rice, C. M. (1994) *EMBO J.* **13**, 2295–2934
- Elce, J., Hegadorn, C., Simon, J., and Arthur, J. S. C. (1997) *J. Biol. Chem.* **272**, 11268–11275
- Sorimachi, H., Toyamasorimachi, N., Saido, T. C., Kawasaki, H., Sugita, H., Miyasaka, M., Arahata, K., Ishimura, S., and Suzuki, K. (1993) *J. Biol. Chem.* **268**, 10593–10605
- Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Smith, J. A., Seidman, J. G., and Struhl, K. (eds) (1998) *Current Protocols in Molecular Biology*, Wiley-Interscience, New York
- Chen, C.-K., Wieland, T., and Simon, M. I. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93**, 12885–12889
- Lévy, F., Johnsson, N., Rümnapf, T., and Varshavsky, A. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93**, 4907–4912
- Johnsson, N., and Varshavsky, A. (1994) *EMBO J.* **13**, 2686–2698
- Sorimachi, H., Ishiura, S., and Suzuki, K. (1997) *Biochem. J.* **328**, 721–732
- Brown, N., and Crawford, C. (1993) *FEBS Lett.* **322**, 65–68
- Kinbara, K., Ishiura, S., Tomioka, S., Sorimachi, H., Jeong, S. Y., Amano, S., Kawasaki, H., Kolmerer, B., Kimura, S., Labeit, S., and Suzuki, K. (1998) *Biochem. J.* **335**, 589–596
- Rock, K. L., Gramm, C., Rothstein, L., Clark, K., Stein, R., Dick, L., Hwang, D., and Goldberg, A. L. (1994) *Cell* **78**, 761–771
- Druey, K. M., Blumer, K. J., Kang, V. H., and Kehrl, J. H. (1996) *Nature* **379**, 742–746
- Moerschell, R. P., Hosokawa, Y., Tsunasawa, S., and Sherman, F. (1990) *J. Biol. Chem.* **265**, 19638–19643
- Walker, K. W., and Bradshaw, R. A. (1999) *J. Biol. Chem.* **274**, 13403–13409
- Turk, B. E., Griffith, E. C., Wolf, S., Biemann, C., Chang, Y.-H., and Liu, J. O. (1999) *Chem. Biol.* **6**, 823–833
- Watson, A. J., Agaray, A. M., Slepak, V. Z., and Simon, M. I. (1996) *J. Biol. Chem.* **271**, 28154–28160
- Srinivasa, S. P., Bernstein, L. S., Blumer, K. J., and Linder, M. E. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95**, 5584–5589
- Cook, R. K., Sheff, D. R., and Rubenstein, P. A. (1991) *J. Biol. Chem.* **266**, 16825–16833
- Uhlmann, F., Lottspeich, F., and Nasmyth, K. (1999) *Nature* **400**, 37–42
- Nishizawa, M., Furuno, N., Okazaki, K., Tanaka, H., Ogawa, Y., and Sagata, N. (1993) *EMBO J.* **12**, 4021–4027
- Kim, E., Arnould, T., Sellin, L., Benzing, T., Comella, N., Kocher, O., Tsiokas, L., Sukhatme, V. P., and Walz, G. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96**, 6371–6376
- Benzing, T., Brandes, R., Sellin, L., Schermer, B., Lecker, S., Walz, G., and Kim, E. (1999) *Nat. Med.* **5**, 913–918
- Stewart, A. (1995) *Trends in Genetics Nomenclature Guide*, Elsevier Science, Ltd., Cambridge, U.K.
- Webb, E. C. (ed) (1992) *Enzyme Nomenclature 1992*, p.527, Academic Press, New York