

Characterization of Novel Substrates of N-end Rule-Dependent Ubiquitylation Identified through a Proteome-Wide *in vitro* Screen

Ilia V. Davydov, John H. Kenten,
Yassamin J. Safiran, Pankaj Oberoi,
Hans A. Biebuyck and Jacob N. Wohlstadter



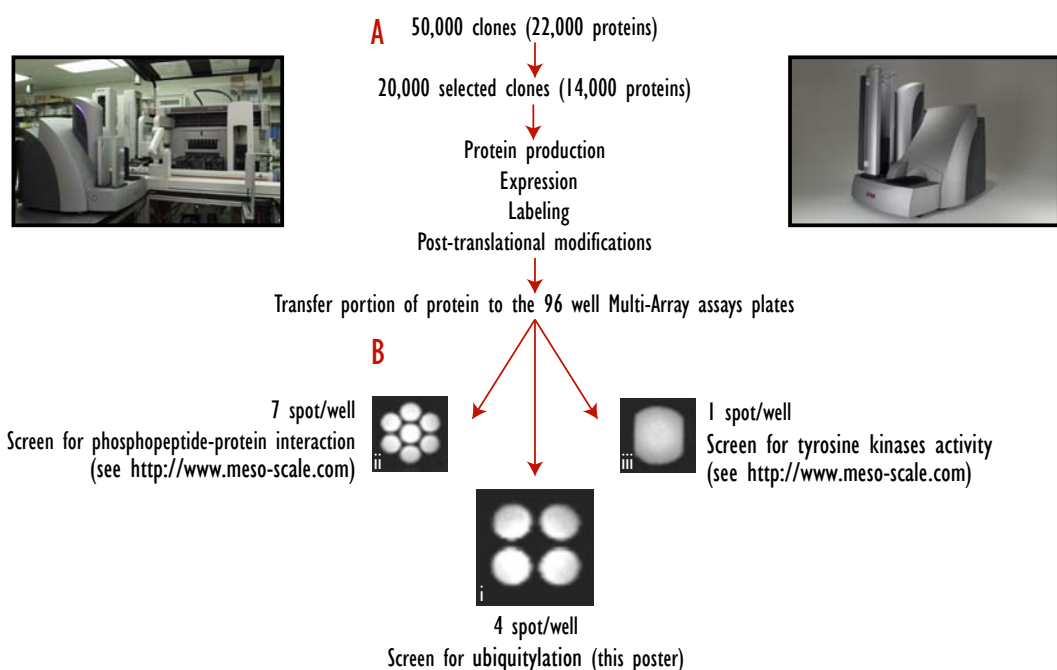
Meso Scale Discovery[®]
A division of Meso Scale Diagnostics, LLC.

Characterization of Novel Substrates of N-end Rule-Dependent Ubiquitylation Identified through a Proteome-Wide *in vitro* Screen

1 Abstract

Ubiquitin is a highly conserved eukaryotic protein. Ubiquitylation of intracellular proteins, i.e. their modification by the covalent addition of ubiquitin, is an important regulatory process in eukaryotic cells. Attachment of multiple ubiquitin moieties to a specific protein, for example, can mark it for degradation by the 26S proteasome. Ubiquitin ligases (E3s) mediate ubiquitylation by their recognition of specific structural elements, called degradation signals, on protein substrates. One well-characterized degradation signal, recognized by the ubiquitin ligase UBR1, is the presence of a destabilizing amino acid residue like arginine at the N-terminus of a protein. This pathway of protein ubiquitylation is called the N-end rule pathway. Here, we demonstrate a systematic, high throughput screen for potential substrates of N-end rule-dependent ubiquitylation from a library of 14,000 individual proteins expressed from ~18,000 cDNA clones. We developed an assay for the detection of protein ubiquitylation that used electrochemiluminescence and Multi-Array™ technology from Meso Scale Discovery. Each protein from our library was exposed to the ubiquitylation machinery of reticulocyte lysates, a system known to have an active N-end rule pathway. We used our ubiquitylation assay in high throughput to select 600 proteins that had high levels of ubiquitylation in reticulocyte lysates. We further studied ubiquitylation of these proteins in the absence and in the presence of dipeptide inhibitors of the N-end rule pathway. We selected and confirmed 60 proteins whose ubiquitylation showed inhibition by the dipeptides. These proteins included RGS4, RGS5, RGS16, known N-end rule substrates (Davydov and Varshavsky, 2000, J. Biol. Chem., 275, 22931-22941). We also discovered several new N-end rule substrates. Measurements by electrochemiluminescence and conventional immunoprecipitation-Western blots confirmed the ubiquitylation of these proteins *in vitro* and their inhibition by the dipeptides. We further characterize and contrast the mechanisms of ubiquitylation for several proteins among this group.

2 Multi-Array Technology



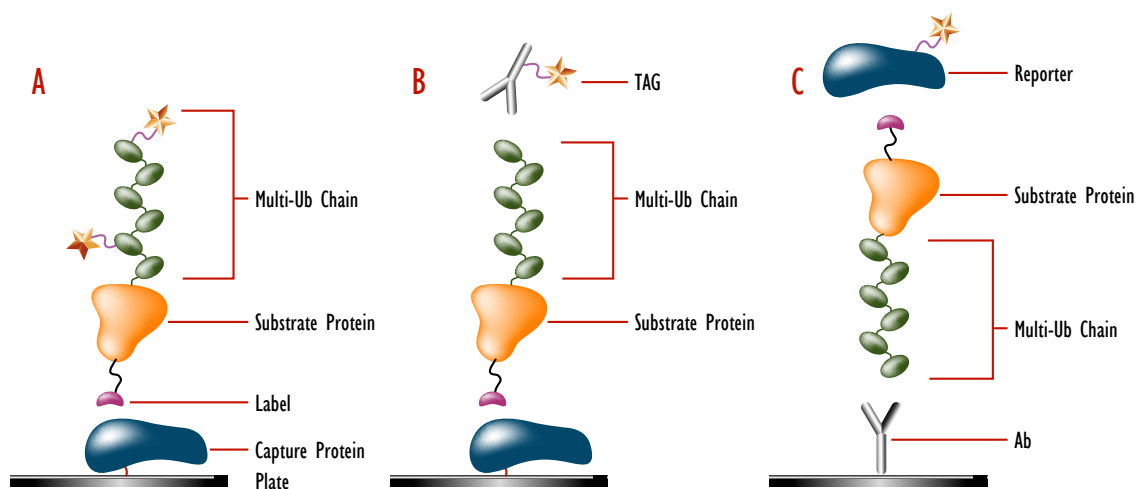
(A) We carried out a proteome wide screen using measurements of 20 different functional assays to survey a number of functional characteristics of each protein of the library including their (i) ubiquitylation, phosphorylation, (ii) ability to act as binding partners to specific phosphopeptides or to act as (iii) tyrosine kinases. Each assay used a common set of modified and labeled proteins that we distributed using automation (Figure at left) into Multi-Spot™ plates for subsequent measurement on the Sector HTS™ reader (Figure at right) using electrochemiluminescence detection. (B) The Multi-Spot plates feature one or more spots in each well of a 96-well plate that allowed us to carry out experiments of many different types and in many different formats (see examples below). This poster focuses on the results from one part of the screen, the analysis of proteins ubiquitylated by reticulocytes lysates.



Meso Scale Discovery™
A division of Meso Scale Diagnostics, LLC.

Characterization of Novel Substrates of N-end Rule-Dependent Ubiquitylation Identified through a Proteome-Wide *in vitro* Screen

3 Ubiquitylation assay formats



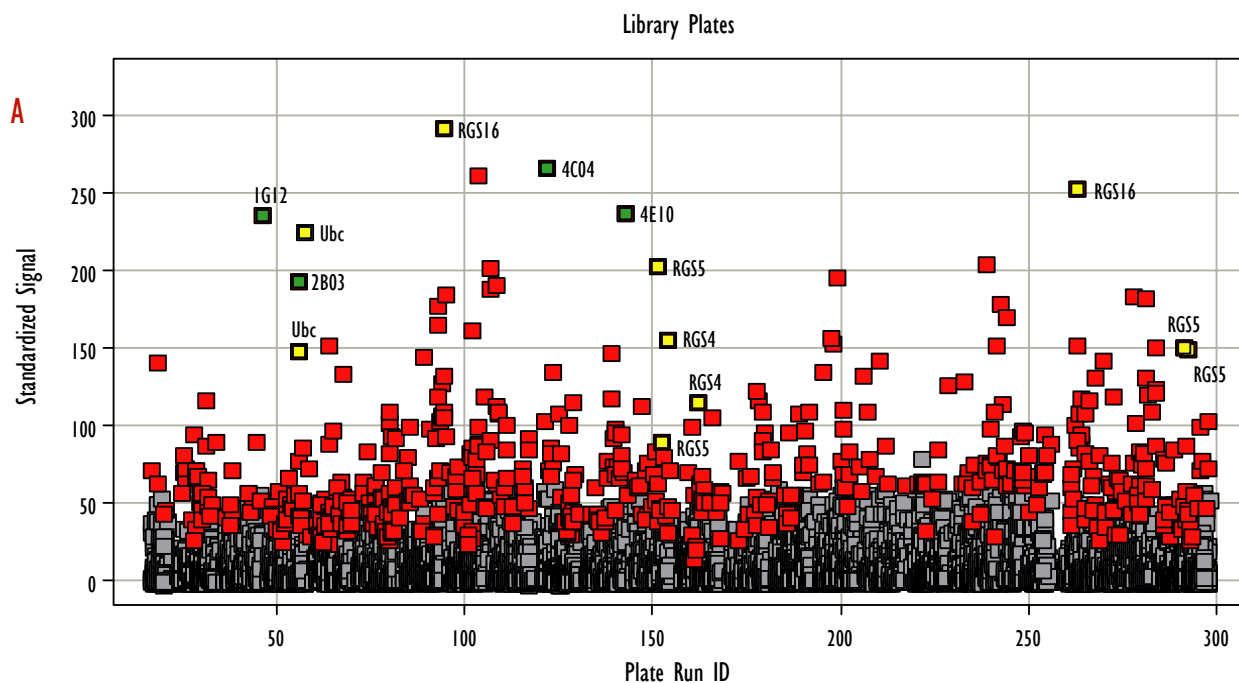
Substrate proteins resulted from their expression and random labeling from the corresponding full length cDNA clone. The proteins were exposed to ubiquitylating enzymes endogenous to reticulocyte lysates. Measurement of the degree of modification of the protein followed using binding assays specific for ubiquitylated products. We developed three assay formats that employed Multi-Array assay plates and the Sector HTS reader for detection. In (A), MSD TAG™ labeled ubiquitin incorporates into multi-Ub chains; the ubiquitylated products are captured at the surface of Sector HTS plates by capture protein-label interactions. Alternatively (B), the ubiquitylated product is detected using a MSD TAG labeled antibody against ubiquitin-protein conjugates. This format can also be inverted (C) using MSD TAG labeled capture protein as a reporter. In this study we used format C.



Meso Scale Discovery™
A division of Meso Scale Diagnostics, LLC.

Characterization of Novel Substrates of N-end Rule-Dependent Ubiquitylation Identified through a Proteome-Wide *in vitro* Screen

4 Primary Screen of 18,000 cDNA clones



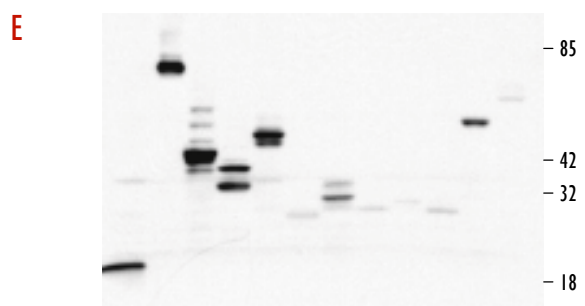
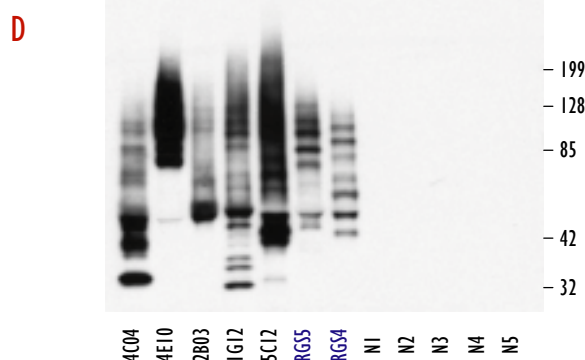
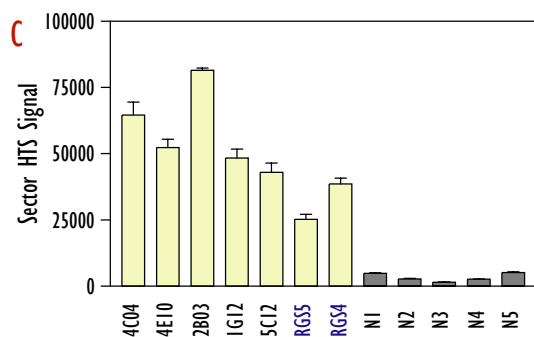
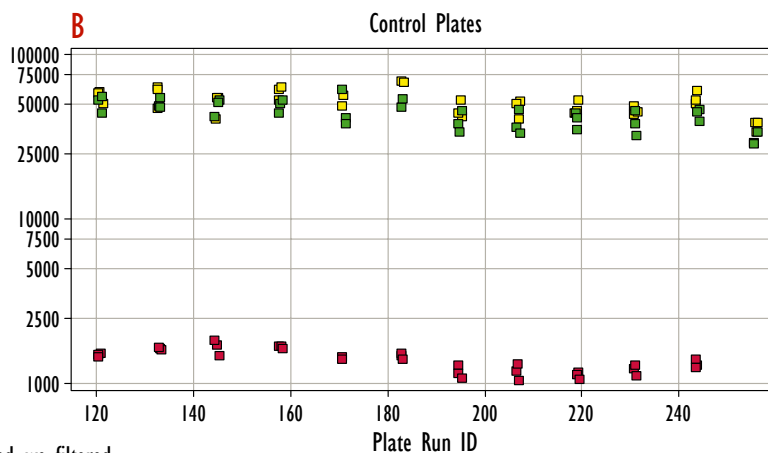
Meso Scale Discovery[®]
A division of Meso Scale Diagnostics, LLC.

Characterization of Novel Substrates of N-end Rule-Dependent Ubiquitylation Identified through a Proteome-Wide *in vitro* Screen

Plasmid DNA from 18,000 clones encoding human and mouse proteins was purified and stored in a library of 190 96-well plates. We used standard laboratory automation in high throughput to produce, label, ubiquitylate, and measure ubiquitylated protein from each member of this library. (A) The primary screen involved 270 96-well plates, including 35 control plates, 45 library plates run in duplicate and the remainder of the library plates. We picked 528 ubiquitylated proteins in the primary screen (yellow, red and green squares) for secondary screening using several criteria to sort the database of measurements.

First, signals on a given plate were standardized to the lowest 50% of signals on a plate (background). Second, we filtered this set for clones having standard signals (as defined above) greater than 30 and a signal to background of greater than 10. Finally, we used an estimate of protein expression to cull the set into proteins having the highest relative ubiquitylation. Using these criteria, we identified, among others, clones from our library that are known substrates for *in vitro* ubiquitylation in reticulocyte lysates: RGS4, RGS5, and RGS16 (Davydov and Varshavsky, 2000, *J. BioChem.*, 275, 22931). Our library contained multiple clones for these substrates, each recovered in our primary screen at high levels of signal. This recovery, and data from gels for other clones (see below) validated our methodology and confirmed its general appropriateness for the discovery of proteins targeted and controlled by ubiquitylation pathways.

(B) Control plates having wells with either one of two RGS4 cDNAs (positive control for ubiquitylation) or empty vector (negative control for ubiquitylation) demonstrated the absence of drift in the expression and measurement systems across weeks of the experiment and the robustness (high signal to background, high reproducibility, and high specificity) of our approach.



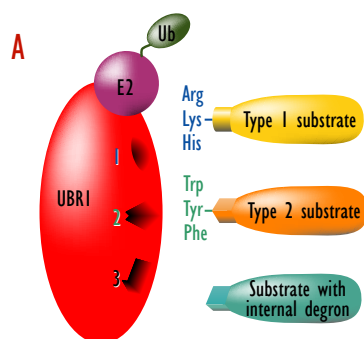
5 Specificity of the Ubiquitylation assay

Comparison of signals from selected clones measured by the Multi-Array assay (C), immunoprecipitation with antibodies to ubiquitylated proteins followed by a western blot with peroxidase-coupled reporter specific for labeled proteins (D), and a western blot for total labeled proteins produced in our system (E). These data demonstrate the correlation between levels of ubiquitylation measured by conventional gels and our high throughput method, further confirming that proteins selected as hits in our screen were truly ubiquitylated.



Characterization of Novel Substrates of N-end Rule-Dependent Ubiquitylation Identified through a Proteome-Wide *in vitro* Screen

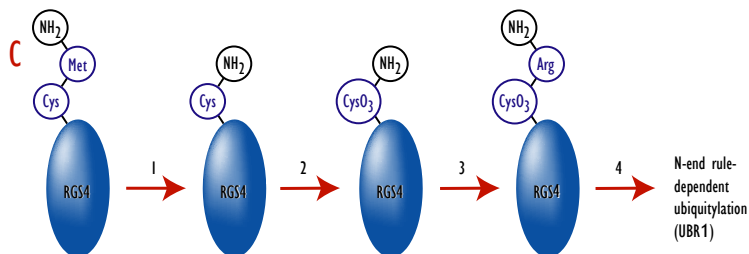
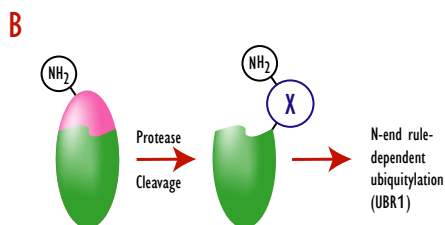
6 Recognition of N-end rule substrates for ubiquitylation



A. Ubiquitin ligase UBR1 has three substrate binding sites (A. Varshavsky, *Nat. Medicine*, 6, 1077 (2000)). The type 1 site is specific for N-terminal residues Arg, Lys, and His. The type 2 site is specific for hydrophobic N-terminal residues Trp, Tyr, Phe, Ile, and Leu. The type 3 site, described for the *S. cerevisiae* UBR1, targets ubiquitylation of CUP9. The dipeptide Arg-Ala competitively inhibits ubiquitylation of type 1 substrates. The dipeptide Trp-Ala inhibits ubiquitylation of type 2 substrates. Interestingly, both Arg-Ala and Trp-Ala enhance ubiquitylation of CUP9, demonstrating allosteric control of this E3.

B. In general, Type 1 and Type 2 N-end rule substrates can result from proteolytic cuts to a protein that reveal destabilizing residues at the new N-terminus. Known examples of this type of processing include yeast SCC1, a subunit of the cohesin complex (Rao et al. *Nature*, 410, 955-959 (2001)) and nSP4, an RNA polymerase of Sindbis virus (deGroot et al, *Proc. Nat. Acad. Sci.*, 88, 8967-8971 (1991)).

C. RGS4 becomes a type 1 N-end rule substrate in reticulocyte lysates after sequential N-terminal modifications: (1) removal of Met-1; (2) oxidation of Cys-2 into cysteic acid CysO₃; (3) conjugation of Arg to the N-terminus (Davydov and Varshavsky, *J.Biol.Chem.*, 275, 22931 (2000); Kwon et al., *Science*, 297, 96 (2002))

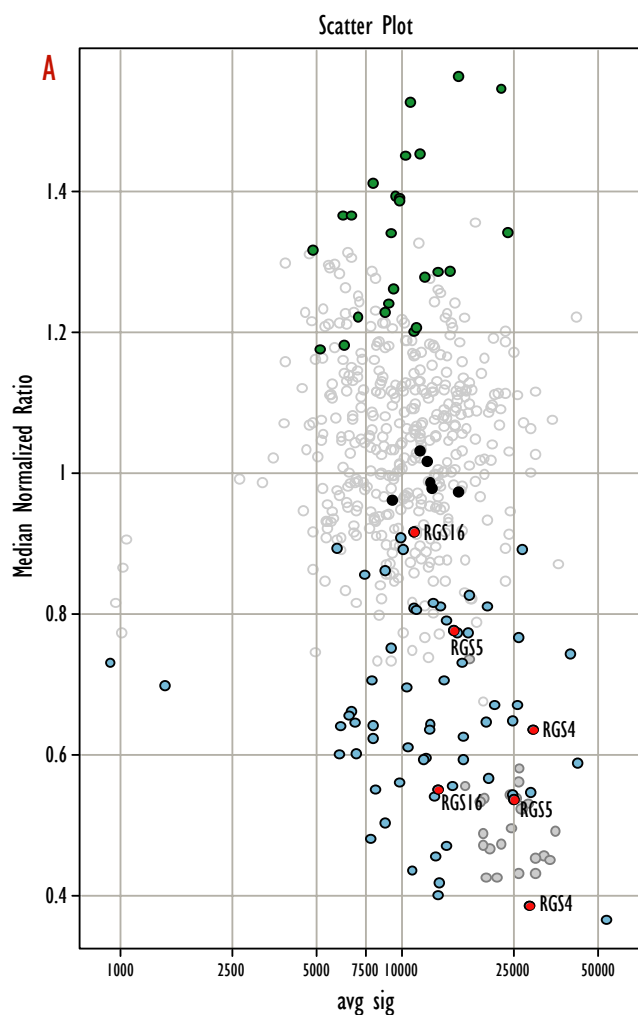
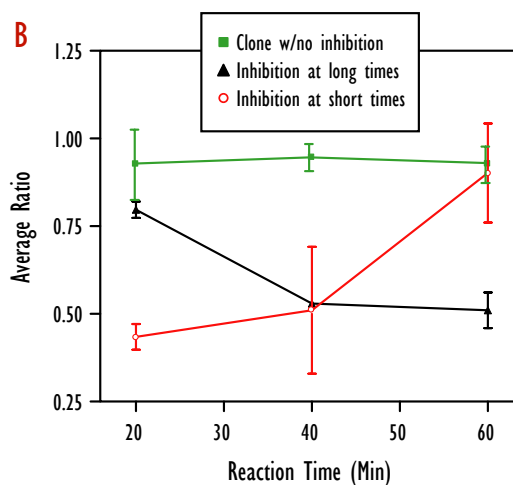


Meso Scale Discovery[®]
A division of Meso Scale Diagnostics, LLC.

Characterization of Novel Substrates of N-end Rule-Dependent Ubiquitylation Identified through a Proteome-Wide *in vitro* Screen

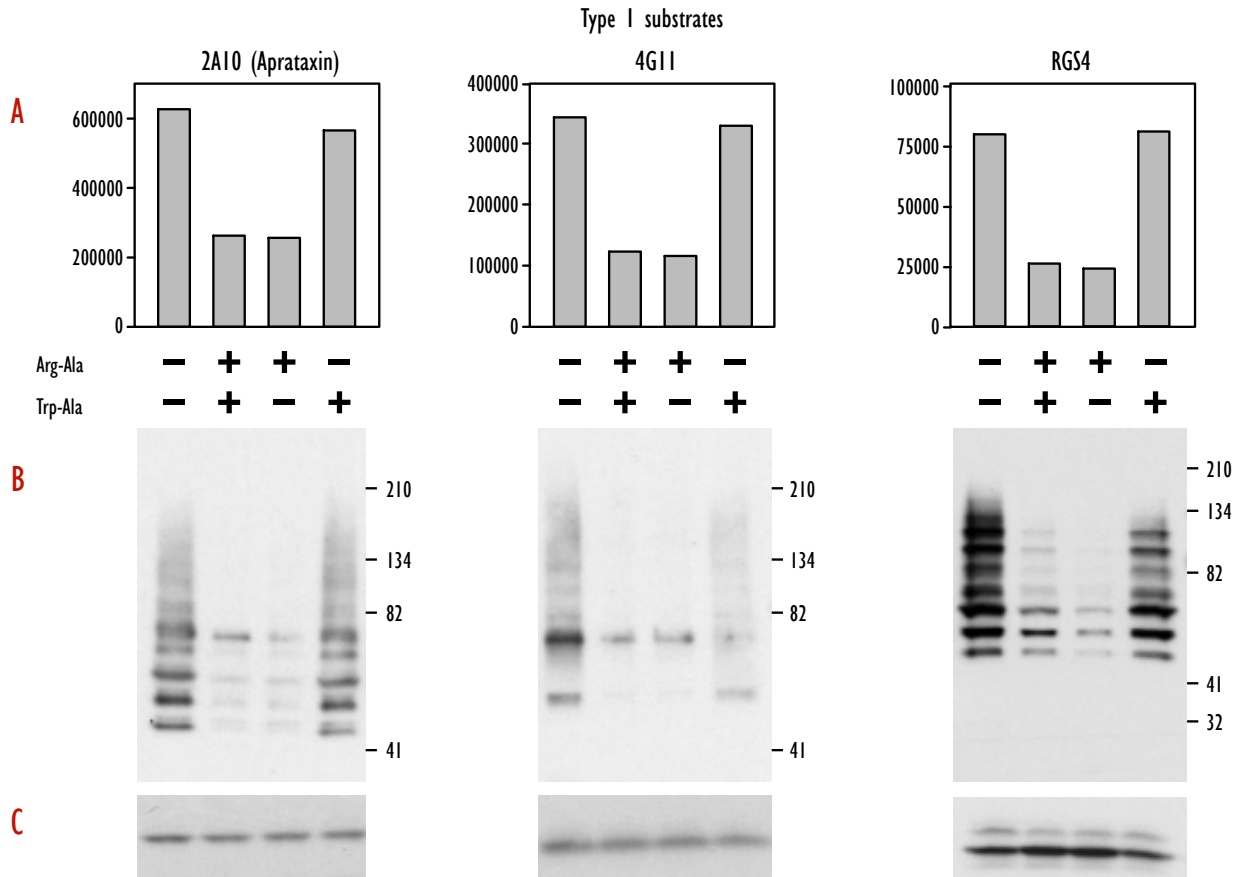
7 Screen for N-end rule substrates

We further refined the ubiquitylated hits from the primary screen into a set of putative N-end rule substrates using a mixture of Arg-Ala and Trp-Ala to modulate the ubiquitylation reaction by the UBR family of E3 ligases present in the reticulocyte lysate. We expected that Type 1 and Type 2 substrates would have lower levels of ubiquitylation in the presence of dipeptides. In contrast, we expected Type 3 substrates to have increased levels of ubiquitylation in the presence of the dipeptides. Hits having no N-end rule participation should have no significant changes in levels of ubiquitylation. The assay was run in duplicate at 20, 40, and 60 minute time points. The ratio of the signal in the presence of dipeptides to the signal in the absence of dipeptides (median normalized ratio; where a value of 1 indicates no effect) was used to identify clones selected for further investigation. (A) We selected 60 clones (blue circles) where the presence of the dipeptides caused significant suppression of the ubiquitylation signals (median normalized ratio < 0.75). (B) We further selected additional clones based on kinetic profiles that showed significant signal suppression at some time point. The figure shows three types of representative time courses for the effect of dipeptides on ubiquitylation of the target protein. These data demonstrate how differently putative N-end rule substrates respond to inhibition of their ubiquitylation and provided additional criteria for the selection of clones (i.e., those having mean normalized ratios of > 0.75). We used 8 (black circles) clones as negative controls in subsequent work and 24 clones (green circles) that showed enhancement of ubiquitylation with the addition of the dipeptides.



Characterization of Novel Substrates of N-end Rule-Dependent Ubiquitylation Identified through a Proteome-Wide *in vitro* Screen

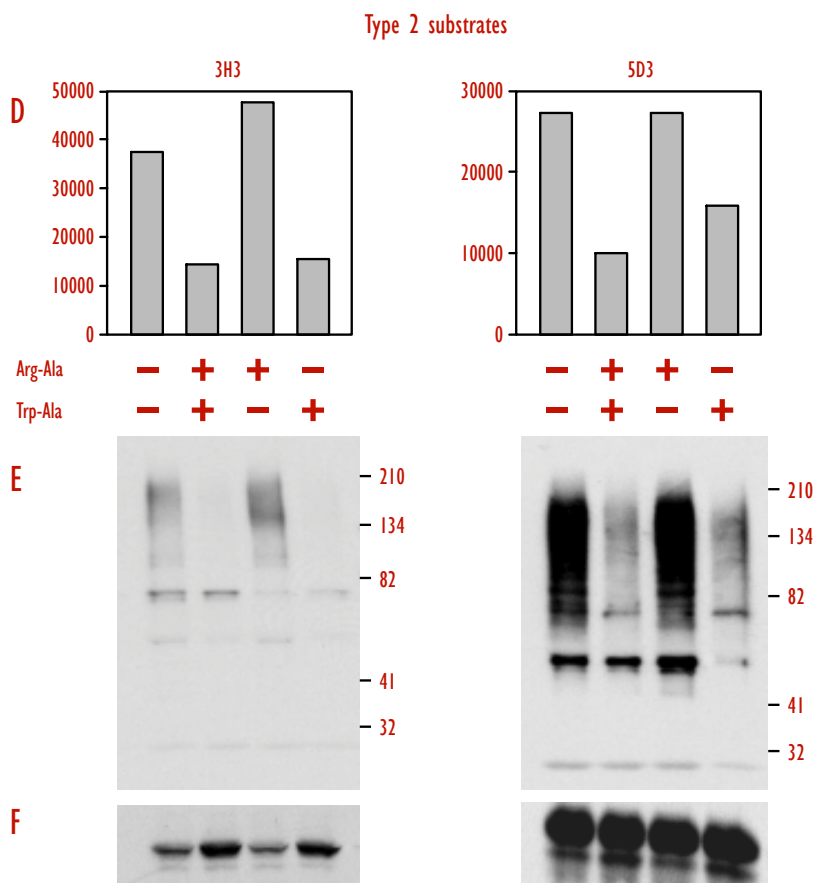
8 Characterization of N-end rule substrates from the screen



Meso Scale Discovery[®]
A division of Meso Scale Diagnostics, LLC.

Characterization of Novel Substrates of N-end Rule-Dependent Ubiquitylation Identified through a Proteome-Wide *in vitro* Screen

Potential N-end rule substrates from the secondary screen were grouped into type 1 or type 2 substrates based on their response to Arg-Ala (type 1 inhibitor) and Trp-Ala (type 2 inhibitor) dipeptides. Comparison of (A,D) signals from selected clones measured by the Sector HTS reader; (B,E) immunoprecipitation with antibodies to ubiquitylated proteins followed by a western blot with a peroxidase coupled reporter specific for labeled proteins and (C,F) a western blot for total labeled proteins produced in our system. The correlation between these methods was excellent.

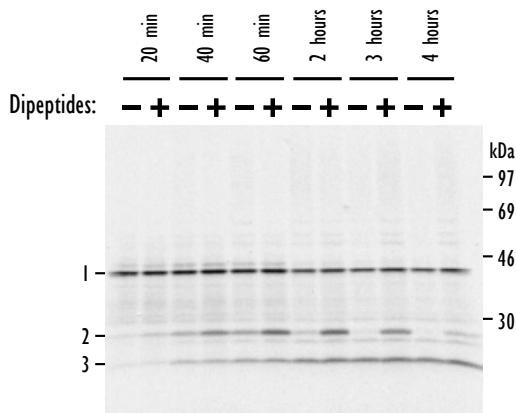


Meso Scale Discovery[®]
A division of Meso Scale Diagnostics, LLC.

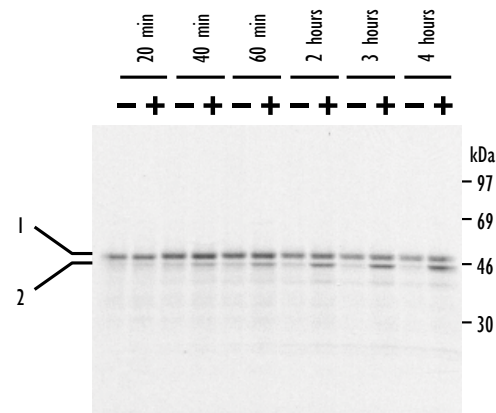
Characterization of Novel Substrates of N-end Rule-Dependent Ubiquitylation Identified through a Proteome-Wide *in vitro* Screen

9 N-end rule substrates may be generated through proteolytic cuts

2A10 (Aprataxin)



5D3



Generation of N-end rule substrates can proceed proteolytic processing of a protein. Potential N-end rule substrates were labeled and ubiquitylated in reticulocyte lysates either in the absence, or in the presence of a mixture of Arg-Ala and Trp-Ala dipeptides. Reactions were stopped at different reaction times and separated on SDS-PAGE followed by autoradiography.

2A10 (aprataxin, Date et al., Nat. Genet., 29, 184-188 (2001), Moreira et al., Nat. Genet., 29, 189-193 (2001)) – The original protein (band 1) is proteolyzed into two fragments (band 2 and band 3). Band 2 represents a C-terminal fragment of the original protein with type 1 destabilizing residue at the N-terminus. Aprataxin has multi-domain structure and is possibly involved in DNA repair (Caldecott, Cell, 112, 7-10 (2003)). The observed proteolytic cleavage separates N-terminal PNK-homology domain of aprataxin from the C-terminal HIT domain.

5D3 – The original protein (band 1) is likely proteolytically trimmed from the N-terminus to generate a shorter polypeptide (band 2) with type 2 residue at the N-terminus that is metabolically stabilized by the dipeptides.



Characterization of Novel Substrates of N-end Rule-Dependent Ubiquitylation Identified through a Proteome-Wide *in vitro* Screen

10 Conclusions

We developed a facile proteome wide method for the investigation of protein post-translational modifications based on a unique gene library, high throughput protein production and Multi-Array technology. Our means of making, labeling and detecting proteins, along with its speed, sensitivity, specificity and stability make our methodology a powerful tool for the deterministic discovery of functional aspects of a proteome using a mixture of hypothesis and non-hypothesis driven experimentation.

Here we used our methodology to screen for protein substrates of N-end rule dependent ubiquitylation. We showed the blinded recovery of clones from known substrates of this pathway in our library and, importantly, greatly expanded the ranks of putative substrates of the N-end rule pathway.

- First, our screen of 18,000 full length clones demonstrated, not wholly unexpectedly, the ubiquitylation of many, but certainly not all, proteins in reticulocyte lysates indicating that this pathway, like phosphorylation, is an omnipresent source of signaling and regulation within cells across all known families of proteins.
- Second, from a selected set of ~500 ubiquitylated proteins we identified ~60 potential substrates of the N-end rule-dependent ubiquitylation pathway based on modulation of the ubiquitylation of these proteins by Arg-Ala and Trp-Ala dipeptides in an automated Multi-Array format.
- Third, we grouped the new substrates of the N-end rule pathway into three sets based on their response, inhibitory or stimulatory, to specific dipeptides. Our observation and *in vitro* validation of a large number of new substrates suggests that the N-end rule pathway may be much more important in regulating proteins than currently appreciated.
- Fourth, we demonstrate that proteolytic cleavage of a protein can precede its degradation by the N-end rule pathway for two mammalian proteins. In one such case, proteolytic cleavage of aprataxin resulted in separation of its N-terminal PNK-homology domain from the C-terminal HIT domain. The C-terminal fragment became ubiquitylated and degraded by the N-end rule pathway. Our results extend the observation of this type of processing of proteins from yeast [Rao, H. et al. *Nature* 955-959 (2001)] to humans and underscores its potential importance as a regulatory pathway in cells.

