

UCSF

UC San Francisco Electronic Theses and Dissertations

Title

Profiling the Translational Activity of Ribosome Subpopulations

Permalink

<https://escholarship.org/uc/item/87k373gb>

Author

Williams, Christopher Card

Publication Date

2014

Peer reviewed|Thesis/dissertation

Profiling the translational activity of ribosome subpopulations

by

Christopher Card Williams

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Chemistry and Chemical Biology

in the

Copyright 2014
By
Christopher Williams

ACKNOWLEDGEMENTS

I could not have executed on a transition from chemistry to cell biology and computational genomics without the mentorship, generosity, and encouragement of many people.

During my first year, my rotation mentors Chris Bohlen, Marcy Diaz, Philipp Kimmig, and Onn Brandman jump-started this process by impressing on me good habits and teaching me more molecular and cell biology than I learned in all of college. I want to thank every single member of the Weissman lab for their incessant generosity with reagents, time, and ideas, as well as their willingness to drink scotch at the drop of a hat. I also want to thank Manny DeVara and Christopher Reiger for keeping the Weissman lab operation functional.

My thesis work would not have been possible, or nearly as good, without my close collaboration, mentorship, and friendship with Calvin Jan. Calvin is hands down the smartest and most selfless person I have ever worked with, and it has been a pleasure spending more than 40 hours a week with him for several years. In addition to being instrumental in teaching me how to do science, he has provided excellent company on long hours sitting on a bike seat.

I want to thank Jonathan for taking a chance on a non-biologist joining his lab. His unique perspective on science and technology has been valuable in more ways than I can enumerate. I want to also thank him for his generosity in promoting the personal development of members of his lab: sponsoring software development and data visualization classes, consenting to the re-allocation of many hours of lab time to Stanford computer science classes, and sending me to scientific conferences even when I had nothing to present, have all significantly contributed to my growth and graduate experience.

Finally, I have to thank my roommate Dylan Sather, both for putting up with the complaints and lack of expendable funds that come with having a PhD student for a roommate, as well as for perpetual encouragement to push onward down the path of computation.

Profiling the translational activity of ribosome subpopulations

Christopher C. Williams

Abstract

Although recent advances in sequencing technology have enabled unprecedented study of gene expression, systematic investigation of its sub-cellular and tissue-level heterogeneity remains a major challenge. With regard to protein expression, translational activity is known to vary in space at different subcellular sites, with ribosome composition, and over a translational cycle through interactions with mRNA or ribosome-associated factors. Here, I describe the development, validation, and application of a genetically controllable proximity-specific ribosome profiling approach to monitor the translational activity of sub-pools of ribosomes *in vivo*. Importantly, it is both comprehensive and highly precise, enabling mechanistic study of the spatiotemporal dynamics of translation at a genome scale.

We apply the this approach to study local translation at two broadly conserved subcellular sites, the endoplasmic reticulum (ER) and mitochondria, and also demonstrate its feasibility for specifically isolating ribosomes which interact with soluble factors. Study of ER translation revealed that co-translational translocation is far more pervasive than was previously thought, and that the location of a signal sequence within a protein is the major determinant for undergoing this form of import. Additionally, dissecting the functional roles of distinct translocon complexes enabled a more sophisticated understanding of how the cell supports efficient co-translational import of a diverse set of substrates. Interrogation of local mRNA translation at the mitochondrial outer membrane highlights additional questions whose study is enabled by proximity-specific ribosome profiling. We obtained direct evidence for co-translational import of mitochondrial inner membrane proteins *in vivo*, resolving a decades-long debate and in the field regarding the mechanism of protein insertion. By exploiting the sensitivity of the approach, we confidently identified dozens of

unannotated candidate mitochondrial proteins. Furthermore, we demonstrate the utility of our method for systematic study of protein dual localization through synthesis of our ER and mitochondrial datasets. The specificity of these spatially resolved proteomic maps revealed a novel ER-form of fumarate reductase, suggesting a mechanism for oxidative folding in the ER under anaerobic conditions. At a unique interface of gene expression and cell biology, future applications of proximity-specific ribosome profiling will inform our understanding of translational heterogeneity within cells and tissues.

CONTRIBUTIONS

Excluding the contributions specified here, this dissertation constitutes the work of Christopher Williams. Portions of Chapter II, and all of Chapters III and IV, are reproduced from two publications under review to be published in Science at the time of this writing. These are entitled “Principles of ER Co-Translational Translocation Revealed by Proximity-Specific Ribosome Profiling” and “Targeting and plasticity of mitochondrial proteins revealed by proximity-specific ribosome profiling.” As these are under review, the text here will certainly differ from that of the final published versions. The entirety of the work represented in these publications was carried out in close collaboration with, and often under the guidance of, Calvin H. Jan. Additional acknowledgements are included in Chapters III and IV. Jonathan S. Weissman supervised this work.

TABLE OF CONTENTS

PREFACE	Acknowledgements	iii
	Abstract	iv
	Contributions	vi
	Table of contents	vii
	List of tables	viii
	List of figures	ix
CHAPTER ONE	Introduction and review	1
	Preface	2
	Translation at the endoplasmic reticulum	3
	Translation at the mitochondrial outer membrane	6
	References	9
CHAPTER TWO	Development of the proximity-specific ribosome profiling approach	13
	Introduction	14
	Components of proximity-specific ribosome profiling	15
	a. Tagging the ribosome	15
	b. BirA localization	16
	c. Inducing biotinylation	16
	d. Quenching post-lysis biotinylation	17
	e. Isolating biotinylated ribosomes	18
	Conclusion	18
	Figures	19
References	23	
CHAPTER THREE	Principles of co-translational translocation at the endoplasmic reticulum	25
	Introduction	26
	Validation of proximity-specific ribosome profiling	26
	Co-translational targeting of SRP-dependent and -independent substrates <i>in vivo</i>	28
	Comprehensive analysis of co- versus post-translational translocation <i>in vivo</i>	30
	Timing and specificity of co-translational targeting to the ER	31
	Ssh1	32

	Sec63	33
	Dynamics of ER-associated ribosomes	33
	Discussion	34
	Figures	38
	Materials and methods	43
	Supplemental figures	57
	Supplemental tables	60
	References	62
CHAPTER FOUR	Trafficking and plasticity of the mitochondrial proteome	65
	Introduction	66
	Results	66
	Perspective	73
	Figures	74
	Materials and methods	76
	Supplementary figures	80
	Supplementary tables	84
	References	86
CHAPTER FIVE	Discussion and future perspectives	89
	Summary	90
	Translation at the ER	91
	Translation at the mitochondrial outer membrane	92
	Isolation of ribosome subpopulations	93
	Figures	95
	References	99
LIST OF TABLES		
CHAPTER THREE		
Table 3-1	Yeast strains	60
Table 3-2	Cell lines	61
Table 3-3	Plasmids	61
Table 3-4	Primers	61
CHAPTER FOUR		
Table 4-1	Yeast strains	84
Table 4-2	Plasmids	84

Table 4-3	Primers	84
-----------	---------	----

LIST OF FIGURES

CHAPTER TWO

Figure 2-1	Components for proximity-specific ribosome profiling in 293T cells	19
Figure 2-2	Inducing biotinylation in yeast	20
Figure 2-3	Control of post-lysis biotinylation in yeast	21
Figure 2-4	Pulldown specificity	22

CHAPTER THREE

Figure 3-1	A system for <i>in vivo</i> proximity-dependent ribosome biotinylation to monitor local protein synthesis at the ER	38
Figure 3-2	Specificity of proximity-dependent ribosome profiling across multiple systems	39
Figure 3-3	Global characterization of co- vs post-translational translocation <i>in vivo</i>	40
Figure 3-4	Timing and specificity of co-translational targeting to the ER	41
Figure 3-5	Dynamics of ER-associated ribosomes <i>in vivo</i>	42
Figure 3-S1	ROC for enrichment thresholds	57
Figure 3-S2	TMD-containing peroxisomal proteins are synthesized at the ER	58
Figure 3-S3	Functionality of Ssh1 BirA fusion proteins	58
Figure 3-S4	Determination of Sec66-dependence by fold change	59
Figure 3-S5	Gene-level translational enrichments upon acute loss of SRP	59

CHAPTER FOUR

Figure 4-1	Co-translational mitochondrial protein targeting as revealed by proximity-specific ribosome profiling	72
Figure 4-2	Comprehensive characterization of RNC targeting to the mitochondria	73
Figure 4-3	Conservation of paralogous protein localization	73
Figure 4-4	Dual localization of proteins to the mitochondria and ER	74
Figure 4-S1	Gene enrichment analysis	78
Figure 4-S2	Alternative initiation removes the MTSs of Vas1 and Trm1	79
Figure 4-S3	Position-specific enrichment analysis of mitop2 genes	80
Figure 4-S4	Consistency of dual localization across ER BirA fusions	80

CHAPTER FIVE

Figure 5-1	Minimal effects of candidate ER RNA binding proteins on gene-level ER enrichment	95
Figure 5-2	Properties of CHX-dependent genes	96
Figure 5-3	Ribosome-associated factors help maintain ER and mitochondrial targeting specificity	97
Figure 5-4	Proximity-specific ribosome profiling with a soluble factor: Srp72-heh2-BirA	98

CHAPTER ONE

Introduction and review

PREFACE

A hallmark of eukaryotic cells is the presence of highly specialized subcellular environments including both membrane- and non-membrane-bound compartments. Increasingly, localized protein synthesis has been shown to play a critical role in contributing to these subcellular structures by allowing protein production at the site of action and in response to local cellular need (Jung et al., 2014). Local translation is involved in diverse processes including developmental patterning, cellular motility, synaptic plasticity, and protein trafficking through the secretory pathway (St Johnston, 2005). Indeed, numerous microscopy-based studies of individual mRNAs have demonstrated a breadth of subcellular mRNA localizations.

Our characterization and understanding of the translational control mechanisms inherent to localized protein synthesis have been limited by the methods available to study this process. Few high-resolution tools are available that faithfully preserve the cellular spatial relationships necessary for the investigation of translation at specific subcellular sites on a genome scale. More broadly, there are no approaches to isolate defined subpopulations of ribosomes within cells, which would enable systematic analysis of their translational activity or compositional makeup. Recently, ribosome profiling has emerged as a global, quantitative, sequencing-based technique for precisely characterizing whole-cell translation in exquisite detail (Ingolia et al., 2009). In this work, I describe the development, validation, and application of a generalizable approach for labeling specific pools of ribosomes *in vivo* and precisely characterizing their translational activity through a ribosome profiling readout.

Chapter II describes the development of the genetically controllable ribosome profiling-based methodology which enables global investigation of translational heterogeneity within cells. Chapters III and IV demonstrate the power and generality of the approach through application to the study of translation at two fundamental cellular locations, the endoplasmic reticulum (ER) and the mitochondrial outer membrane. The

remainder of Chapter I provides additional motivation by reviewing current models for protein targeting to these two organelles. Finally, Chapter V concludes with a discussion of future applications of proximity-specific ribosome profiling, which will inform our understanding of regulated translational heterogeneity and its functional consequences.

TRANSLATION AT THE ENDOPLASMIC RETICULUM

The ER represents an evolutionarily conserved site of localized protein synthesis where 20-40% of proteins in a given eukaryote enter the secretory pathway (Ast and Schuldiner, 2013). In order for secreted, transmembrane domain (TMD)-containing, and resident endomembrane proteins to enter this organelle, both regulated targeting from the cytosol and physical translocation across or into the ER membrane must occur. Intensive study of these processes over several decades has revealed multiple routes for targeting nascent proteins to the ER and has also elucidated the core translocational machinery necessary for import (Mandon et al., 2013). These in-depth biochemical and mechanistic studies using a limited number of substrates provide an extensive theoretical framework for interpretation of our systematic, genome-scale analysis in Chapter III.

ER targeting

Targeting of proteins to the ER is governed by signal sequences (SSs) or TMDs encoded within nascent polypeptides, and can occur either co- or post-translationally. The principal route for co-translational import is thought to be the signal recognition particle (SRP)-dependent pathway, in which translation is halted upon binding of SRP to nascent SSs and resumes when the ribosome engages the translocon (Ast and Schuldiner, 2013). Subsequent to the discovery of SRP in the 1980s (Walter and Blobel, 1980), it was observed that certain substrates did not form cytosolic aggregates upon disruption of the SRP-dependent pathway *in vivo* (Ng et al., 1996). Additionally, separate experiments demonstrated that some

substrates are capable of import into biochemically purified microsomes in the absence of ribosomes (Panzner et al., 1995; Randall, 1983). This process was shown to be dependent on a ratcheting action mediated by luminal chaperones, although it is unclear whether this ATP-dependent process is utilized as a driving force for all forms of ER import (Brodsky et al., 1995; Jermy et al., 2006; Willer et al., 2003; Young, 2001). Together, these observations led to models of post-translational import pathways all categorized under the imprecise umbrella term “SRP-independent.”

In general, partitioning between the SRP-dependent and -independent pathways remains poorly understood. It has been appreciated that some substrates fail to engage SRP due to physical constraints of signal recognition. Specifically, the C-terminal targeting sequences found in tail-anchored (TA) or very short proteins preclude co-translational recognition by SRP (Ast and Schuldiner, 2013). Indeed, dedicated factors have been identified which mediate the post-translational targeting of these substrates (Schuldiner et al., 2008; Stefanovic and Hegde, 2007). Studies focusing on a limited number of substrates suggest that SRP also fails to bind targeting sequences whose hydrophobicities fall below a certain threshold (Ng et al., 1996). This metric was used recently to computationally predict dependence on SRP genome-wide in yeast, revealing that an unexpectedly large fraction (40%) target independent of this factor (Ast et al., 2013). Whether the import of these substrates occurs concurrently with their synthesis is unknown.

While targeting of co-translationally imported substrates to the ER is thought to be mediated solely by signals in nascent peptides, localized translation in other contexts can be mediated by *trans*-acting RNA binding proteins (RBPs) that recognize specific *cis* sequences in transcripts themselves (Kraut-Cohen and Gerst, 2010). Such targeted mRNA delivery is highly prevalent across cell types, ranging from other compartments in yeast, to *Drosophila* embryos, plants, neurons, and oligodendrocytes (Jung et al., 2014; St Johnston, 2005).

Intriguingly, several ER-localized RBPs have been identified in yeast, including Bfr1 (Lang and Fridovich-Keil, 2000), Whi3 (Colomina et al., 2008), p180 (Cui et al., 2012), and Scp160 (Frey et al., 2001), and there are several reports of translation-independent localization of individual mRNAs to the organelle across organisms (Hermesh and Jansen, 2013; Kraut-Cohen and Gerst, 2010). While the contribution of dedicated mRNA localization to overall translation at the ER remains unclear, such a redundant or alternative localization mechanism would be particularly appealing as an explanation for how yeast are able to adapt to the loss of SRP (Mutka and Walter, 2001). In Chapter III we argue that protein- and mRNA-mediated localization are not mutually exclusive, but instead act in concert to promote efficient translation at the ER. In Chapter V the role of specific RBPs in ER translation is revisited.

ER Translocation

Following targeting to the ER, substrates traverse or integrate into the lipid membrane through a channel formed by the conserved trimeric Sec61 translocon. This core complex may associate with any of several accessory factors, including Sec62, 63, 66, and 72, as well as TRAM and TRAP in mammals (Mandon et al., 2013). Many peripheral proteins are not essential for viability, but in many cases have been shown to increase the efficiency of translocation of specific substrates *in vitro*. In yeast, the Sec proteins are thought to mediate post-translational import although some have also been implicated in co-translational translocation (Mandon et al., 2013). In contrast to ER targeting factors, most proteins involved in translocation have been identified, but their precise functional roles and how they promote efficient translocation of specific sets of substrates remain unclear.

In yeast, translocation can also occur through Ssh1, a non-essential paralog of Sec61. The contribution of this alternative complex to translocation has not been clearly defined, however several lines of evidence suggest that it is functionally distinct from Sec61. In

particular, it is thought that Ssh1 specifically mediates co-translational import due to ribosome association and failure to engage Sec proteins *in vitro* (Spiller and Stirling, 2011). Additionally, deletion of Ssh1 leads to growth impairment that cannot be rescued by ectopic over-expression of Sec61 (Wittke et al., 2002), and this deletion is synthetically lethal with mutations in either Sec61 or components of the SRP pathway (Finke et al., 1996). Although there is limited evidence for substrate specificity between the two translocons, this hypothesis has also been proposed.

Despite our in-depth mechanistic and structural understanding of ER targeting and translocation, the broader cellular organization of these trafficking routes *in vivo* has remained largely unexplored. Experimental limitations have prevented a systematic characterization of substrate flux through the various ER trafficking pathways in unperturbed cells. Similarly, our understanding of ER dynamics remains limited due to the difficulty in precisely monitoring the complex molecular events that occur at this cellular site. In Chapter III, we apply proximity-specific ribosome profiling to comprehensively interrogate the timing of ribosome-nascent chain (RNC) recruitment to the ER surface, the specificity of various translocon components for individual secretory substrates, and finally to monitor RNC fate following translation termination.

TRANSLATION AT THE MITOCHONDRIA OUTER-MEMBRANE

Mitochondria descended from a free-living α -proteobacterium that invaded or was acquired by eukaryotic cells during endosymbiotic evolution. Since this acquisition, the vast majority of mitochondrial genes (93-99%) were transferred to the nuclear genome for increased host control of the organelle (Woodson and Chory, 2008). While presumably advantageous, this migration necessitates the proper targeting and intra-organelle sorting of nuclear-encoded mitochondrial proteins. Additionally, active communication between the two subcellular sites

must occur to coordinate gene expression in response to metabolic changes, environmental stimuli, or fluctuations in mitochondrial copy-number. Systematic study of the holistic principles that govern these processes, and comparison to those which apply to other organelles or locations, would more broadly inform our understanding of subcellular protein sorting and signaling.

Mitochondrial targeting and import

A series of chaperones and channel complexes mediate mitochondrial import and direct protein sorting to the outer membrane (OM), inner membrane (IM), inter-membrane space (IMS), and matrix. The mitochondrial targeting signals (MTSs) of matrix and IM proteins are often encoded as N-terminal, degenerate amphipathic helices that depend on the IM electrochemical gradient to drive translocation. Proteins destined to other mitochondrial compartments contain internal targeting signals, which are more variable and can dictate which of several import routes is utilized (Schmidt et al., 2010). No dedicated insertion machinery has been identified for mitochondrial OM TA proteins. Instead, it has been hypothesized that mitochondrial lipid composition is sufficient to facilitate spontaneous insertion and that quality control pathways degrade those which mis-localize (Okreglak and Walter, 2014).

There is conflicting evidence for whether import into mitochondria occurs post-translationally on unfolded polypeptides or concurrent with protein synthesis. Import of fully synthesized proteins into mitochondria was observed *in vitro* (Harmey et al., 1977; Maccacchini et al., 1979) and full-length protein precursors have been detected in the cytosol (Schülke et al., 1997; Wienhues et al., 1991). However, ribosomes have been observed on the OM by electron microscopy (Kellems et al., 1975) and studies using FISH, live-cell mRNA localization, and microarrays have demonstrated that many mRNAs associate

closely or co-purify with mitochondria (Marc et al., 2002; Saint-Georges et al., 2008; Suissa and Schatz, 1982). For some messages, this localization is dependent on the RBP Puf3, which localizes to the mitochondria but whose molecular function remains unclear (García-Rodríguez et al., 2007). Although these observations support a role for co-translational translocation for at least some substrates, this interpretation is confounded by the use of translation elongation inhibitors in several of the studies; this caveat is discussed at length in Chapter IV. Moreover, no unifying determinants for this perplexing heterogeneity in localization have been established.

Dual localization and proteomics

One mechanism by which cells achieve functional diversification of gene products is through dual targeting of single proteins to distinct subcellular sites. Differentially localized proteins can perform the same or distinct functions at each locale, and the ratio of different subpopulations can be regulated in response to changing cellular conditions. Mechanistically, dual localization can be mediated by the presence or absence of a single targeting signal, or through the encoding of a single ambiguous signal or multiple specific signals within the same polypeptide. Differential encoding of localization signals can be achieved through any combination of (i) gene duplication and paralog-specific signals, (ii) inclusion or exclusion of targeting signals for a single gene through alternate transcription initiation or splicing, and (iii) via alternate translation initiation events which result in differential targeting signals (Yogev and Pines, 2011).

Several documented cases have been reported for mitochondrial proteins with subpopulations in the cytosol, nucleus, peroxisome, ER, chloroplasts, or even separate mitochondrial compartments. For example, the targeting signal of human Fis1 can direct localization to either peroxisomes or mitochondria, where it can regulate local membrane

fission at either site (Delille and Schrader, 2008). Intriguingly, many proteins found in both the mitochondria and nucleus appear to function in the control of gene expression, indicating that dual targeting might enable crosstalk between the two compartments (Duchêne and Giegé, 2012). Indeed, there are documented instances of this form of signaling, for example the LRPPRC protein in humans (Sasarman et al., 2010), and in Chapter IV we identify new candidate proteins for such communication.

Comprehensive characterization of dually localized proteins through biochemical fractionation remains challenging due to impurities which confound subcellular proteomic analysis, particularly when only a small fraction of the protein resides in a given compartment (Naamati et al., 2009). More generally, comprehensive cataloging of the mitochondrial proteome has relied on subcellular proteomics, phylogenomics, and fluorescent tagging due to the difficulty in predicting MTSs. Thus, although sensitive mass spectrometry, yeast genetics, and improvements in biochemical preparations have enabled the identification of singly and dually localized mitochondrial proteins over the last decade (Schmidt et al., 2010), there is an obvious need for alternative and sensitive approaches for such characterization. In Chapter IV, we demonstrate the utility of proximity-specific ribosome profiling for subcellular proteomics, as well as for elucidating fundamental principles of mitochondrial protein import unaddressable by other methods.

REFERENCES

Ast, T., and Schuldiner, M. (2013). All roads lead to Rome (but some may be harder to travel): SRP-independent translocation into the endoplasmic reticulum. *Crit. Rev. Biochem. Mol. Biol.* 48, 273–288.

Ast, T., Cohen, G., and Schuldiner, M. (2013). A Network of Cytosolic Factors Targets SRP-Independent Proteins to the Endoplasmic Reticulum. *Cell* 152, 1134–1145.

Brodsky, J.L., Goekeler, J., and Schekman, R. (1995). BiP and Sec63p are required for both co- and posttranslational protein translocation into the yeast endoplasmic reticulum. *Proc. Natl. Acad. Sci. U. S. A.* 92, 9643–9646.

Colomina, N., Ferrezuelo, F., Wang, H., Aldea, M., and Garí, E. (2008). Whi3, a developmental regulator of budding yeast, binds a large set of mRNAs functionally related to the endoplasmic reticulum. *J. Biol. Chem.* 283, 28670–28679.

Cui, X.A., Zhang, H., and Palazzo, A.F. (2012). p180 promotes the ribosome-independent localization of a subset of mRNA to the endoplasmic reticulum. *PLoS Biol.* 10, e1001336.

Delille, H.K., and Schrader, M. (2008). Targeting of hFis1 to peroxisomes is mediated by Pex19p. *J. Biol. Chem.* 283, 31107–31115.

Duchêne, A.-M., and Giegé, P. (2012). Dual localized mitochondrial and nuclear proteins as gene expression regulators in plants? *Front. Plant Sci.* 3, 221.

Finke, K., Plath, K., Panzner, S., Prehn, S., Rapoport, T.A., Hartmann, E., and Sommer, T. (1996). A second trimeric complex containing homologs of the Sec61p complex functions in protein transport across the ER membrane of *S. cerevisiae*. *EMBO J.* 15, 1482–1494.

Frey, S., Pool, M., and Seedorf, M. (2001). Scp160p, an RNA-binding, polysome-associated protein, localizes to the endoplasmic reticulum of *Saccharomyces cerevisiae* in a microtubule-dependent manner. *J. Biol. Chem.* 276, 15905–15912.

García-Rodríguez, L.J., Gay, A.C., and Pon, L.A. (2007). Puf3p, a Pumilio family RNA binding protein, localizes to mitochondria and regulates mitochondrial biogenesis and motility in budding yeast. *J. Cell Biol.* 176, 197–207.

Harmey, M.A., Hallermayer, G., Korb, H., and Neupert, W. (1977). Transport of cytoplasmically synthesized proteins into the mitochondria in a cell free system from *Neurospora crassa*. *Eur. J. Biochem. FEBS* 81, 533–544.

Hermesh, O., and Jansen, R.-P. (2013). Take the (RN)A-train: localization of mRNA to the endoplasmic reticulum. *Biochim. Biophys. Acta* 1833, 2519–2525.

Ingolia, N.T., Ghaemmaghami, S., Newman, J.R.S., and Weissman, J.S. (2009). Genome-Wide Analysis in Vivo of Translation with Nucleotide Resolution Using Ribosome Profiling. *Science* 324, 218–223.

Jermy, A.J., Willer, M., Davis, E., Wilkinson, B.M., and Stirling, C.J. (2006). The Brl domain in Sec63p is required for assembly of functional endoplasmic reticulum translocons. *J. Biol. Chem.* 281, 7899–7906.

Jung, H., Gkogkas, C.G., Sonenberg, N., and Holt, C.E. (2014). Remote Control of Gene Function by Local Translation. *Cell* 157, 26–40.

Kellems, R.E., Allison, V.F., and Butow, R.A. (1975). Cytoplasmic type 80S ribosomes associated with yeast mitochondria. IV. Attachment of ribosomes to the outer membrane of isolated mitochondria. *J. Cell Biol.* 65, 1–14.

Kraut-Cohen, J., and Gerst, J.E. (2010). Addressing mRNAs to the ER: cis sequences act up! *Trends Biochem. Sci.* 35, 459–469.

- Lang, B.D., and Fridovich-Keil, J.L. (2000). Scp160p, a multiple KH-domain protein, is a component of mRNP complexes in yeast. *Nucleic Acids Res.* 28, 1576–1584.
- Maccacchini, M.L., Rudin, Y., Blobel, G., and Schatz, G. (1979). Import of proteins into mitochondria: precursor forms of the extramitochondrially made F1-ATPase subunits in yeast. *Proc. Natl. Acad. Sci. U. S. A.* 76, 343–347.
- Mandon, E.C., Trueman, S.F., and Gilmore, R. (2013). Protein translocation across the rough endoplasmic reticulum. *Cold Spring Harb. Perspect. Biol.* 5.
- Marc, P., Margeot, A., Devaux, F., Blugeon, C., Corral-Debrinski, M., and Jacq, C. (2002). Genome-wide analysis of mRNAs targeted to yeast mitochondria. *EMBO Rep.* 3, 159–164.
- Mutka, S.C., and Walter, P. (2001). Multifaceted physiological response allows yeast to adapt to the loss of the signal recognition particle-dependent protein-targeting pathway. *Mol. Biol. Cell* 12, 577–588.
- Naamati, A., Regev-Rudzki, N., Galperin, S., Lill, R., and Pines, O. (2009). Dual targeting of Nfs1 and discovery of its novel processing enzyme, Icp55. *J. Biol. Chem.* 284, 30200–30208.
- Ng, D.T., Brown, J.D., and Walter, P. (1996). Signal sequences specify the targeting route to the endoplasmic reticulum membrane. *J. Cell Biol.* 134, 269–278.
- Okreglak, V., and Walter, P. (2014). The conserved AAA-ATPase Msp1 confers organelle specificity to tail-anchored proteins. *Proc. Natl. Acad. Sci.* 201405755.
- Panzner, S., Dreier, L., Hartmann, E., Kostka, S., and Rapoport, T.A. (1995). Posttranslational protein transport in yeast reconstituted with a purified complex of Sec proteins and Kar2p. *Cell* 81, 561–570.
- Randall, L.L. (1983). Translocation of domains of nascent periplasmic proteins across the cytoplasmic membrane is independent of elongation. *Cell* 33, 231–240.
- Saint-Georges, Y., Garcia, M., Delaveau, T., Jourden, L., Le Crom, S., Lemoine, S., Tanty, V., Devaux, F., and Jacq, C. (2008). Yeast Mitochondrial Biogenesis: A Role for the PUF RNA-Binding Protein Puf3p in mRNA Localization. *PLoS ONE* 3, e2293.
- Sasarman, F., Brunel-Guitton, C., Antonicka, H., Wai, T., Shoubridge, E.A., and LSFC Consortium (2010). LRPPRC and SLIRP interact in a ribonucleoprotein complex that regulates posttranscriptional gene expression in mitochondria. *Mol. Biol. Cell* 21, 1315–1323.
- Schmidt, O., Pfanner, N., and Meisinger, C. (2010). Mitochondrial protein import: from proteomics to functional mechanisms. *Nat. Rev. Mol. Cell Biol.* 11, 655–667.
- Schuldiner, M., Metz, J., Schmid, V., Denic, V., Rakwalska, M., Schmitt, H.D., Schwappach, B., and Weissman, J.S. (2008). The GET complex mediates insertion of tail-anchored proteins into the ER membrane. *Cell* 134, 634–645.

Schülke, N., Sepuri, N.B., and Pain, D. (1997). In vivo zippering of inner and outer mitochondrial membranes by a stable translocation intermediate. *Proc. Natl. Acad. Sci. U. S. A.* 94, 7314–7319.

Spiller, M.P., and Stirling, C.J. (2011). Preferential targeting of a signal recognition particle-dependent precursor to the Ssh1p translocon in yeast. *J. Biol. Chem.* 286, 21953–21960.

St Johnston, D. (2005). Moving messages: the intracellular localization of mRNAs. *Nat. Rev. Mol. Cell Biol.* 6, 363–375.

Stefanovic, S., and Hegde, R.S. (2007). Identification of a Targeting Factor for Posttranslational Membrane Protein Insertion into the ER. *Cell* 128, 1147–1159.

Suissa, M., and Schatz, G. (1982). Import of proteins into mitochondria. Translatable mRNAs for imported mitochondrial proteins are present in free as well as mitochondria-bound cytoplasmic polysomes. *J. Biol. Chem.* 257, 13048–13055.

Walter, P., and Blobel, G. (1980). Purification of a membrane-associated protein complex required for protein translocation across the endoplasmic reticulum. *Proc. Natl. Acad. Sci. U. S. A.* 77, 7112–7116.

Wienhues, U., Becker, K., Schleyer, M., Guiard, B., Tropschug, M., Horwich, A.L., Pfanner, N., and Neupert, W. (1991). Protein folding causes an arrest of preprotein translocation into mitochondria in vivo. *J. Cell Biol.* 115, 1601–1609.

Willer, M., Jermy, A.J., Steel, G.J., Garside, H.J., Carter, S., and Stirling, C.J. (2003). An in vitro assay using overexpressed yeast SRP demonstrates that cotranslational translocation is dependent upon the J-domain of Sec63p. *Biochemistry (Mosc.)* 42, 7171–7177.

Wittke, S., Dünwald, M., Albertsen, M., and Johnsson, N. (2002). Recognition of a subset of signal sequences by Ssh1p, a Sec61p-related protein in the membrane of endoplasmic reticulum of yeast *Saccharomyces cerevisiae*. *Mol. Biol. Cell* 13, 2223–2232.

Woodson, J.D., and Chory, J. (2008). Coordination of gene expression between organellar and nuclear genomes. *Nat. Rev. Genet.* 9, 383–395.

Yogev, O., and Pines, O. (2011). Dual targeting of mitochondrial proteins: Mechanism, regulation and function. *Biochim. Biophys. Acta BBA - Biomembr.* 1808, 1012–1020.

Young, B.P. (2001). Sec63p and Kar2p are required for the translocation of SRP-dependent precursors into the yeast endoplasmic reticulum in vivo. *EMBO J.* 20, 262–271.

CHAPTER TWO

Development of the proximity-specific ribosome profiling approach

INTRODUCTION

Genome-wide approaches for studying the spatial control of protein synthesis have required careful biochemical fractionation of ribosomes from the compartment of interest (Stephens and Nicchitta, 2007), or are limited to bulk labeling of nascent chains using methods such as RiboPuroMycylation (David et al., 2012) or FUNCAT (Dieterich et al., 2010). These caveats have restricted both the location of possible analyses as well as the within-gene resolution of the data obtained. These considerations motivated us to develop a generalizable strategy to globally monitor translation in a manner that preserves *in vivo* spatiotemporal information about the site of synthesis. Our approach involves the integration of two separate, previously established approaches: ribosome profiling and proximity biotinylation (Fig. 3-1A).

Ribosome profiling, the deep sequencing of ribosome-protected mRNA fragments (Ingolia et al., 2009), quantitatively reports on genome-wide translation with sub-codon resolution. Compared to approaches such as RNA-seq that utilize transcript abundance as a proxy for protein levels, ribosome profiling allows for the direct monitoring of the process of translation. The scale and depth afforded by deep sequencing make ribosome profiling particularly powerful as both an “omics” approach and as a high-precision tool for the mechanistic study of the process of protein synthesis (Ingolia, 2014). By monitoring the translational activity of ribosomes derived from a subcellular compartment of interest, we reasoned that we could precisely map specific proteins, even codon-resolved domains of proteins, to their site of synthesis.

To faithfully preserve *in vivo* spatiotemporal information, we utilized a proximity biotinylation approach that has been used previously to characterize protein-protein interactions in live cells (Fernández-Suárez et al., 2008). Specifically, one protein of interest is fused to the *Escherichia coli* biotin ligase BirA and the second to an *in vitro*-selected 15 amino acid BirA acceptor peptide, termed an ‘AviTag.’ Spatial proximity of the two fusion proteins

results in site-specific biotinylation of the AviTag by BirA. For our application, we target BirA to a subcellular compartment to selectively biotinylate Avi-tagged ribosomes in that locale, in live cells with all membranes and spatial relations intact. Post-lysis purification and sequencing of biotinylated ribosome footprints then enables quantification of the translational activity specific to the subcellular site of interest, through analysis of gene- and codon-level enrichments computed by comparison to footprints derived from whole-cell monosomes. Together we call this approach proximity-specific ribosome profiling.

COMPONENTS OF PROXIMITY-SPECIFIC RIBOSOME PROFILING

To establish the proximity-specific ribosome profiling method, we implemented the following five steps: (a) introduction of a non-perturbing ribosome tag consisting of a TEV protease-cleavable AviTag; (b) genetic targeting of BirA to a subcellular location of interest; (c) temporal control of ribosome biotinylation *in vivo*; (d) inhibition of post-lysis biotinylation; and (e) selective isolation of biotinylated ribosomes and specific elution via TEV cleavage (Fig. 3-1A). We developed and validated these steps in the budding yeast *Saccharomyces cerevisiae* as well as in the human HEK293 cell line.

a. Tagging the ribosome

Informed by a recent structure of the yeast 80S ribosome (Ben-Shem et al., 2011), we expressed Avi-tagged versions of several candidate ribosomal proteins with surface-accessible termini. Our studies identified multiple subunits that when tagged and expressed from their endogenous loci, including the natural 3' UTR, were incorporated into ribosomes and covered growth defects seen in deletion mutants. These included C-terminally tagged *RPL16* and *RPS2* [also called uL13 and uS5 (Ban et al., 2014)], which were used for

subsequent experiments (Fig. 3-1B). N-terminally tagged Rpl10a was used for mammalian studies (Heiman et al., 2008) (Fig. 2-1A).

b. BirA localization

In principle, proximity-specific ribosome profiling enables the monitoring of translation at any location to which one can target a BirA fusion protein. To validate our approach in yeast, we constructed three different ER-localized BirA fusion proteins, as well as cytosolic and mitochondrial controls (Fig. 3-1C). To broadly capture the translational activity of all ER-associated ribosomes, we localized BirA to the ER using the C-terminal tail-anchor (TA) from *UBC6* (Kornmann et al., 2009). To more specifically examine translation at the two known translocation entry points to the ER, we fused BirA to *SEC63*, a member of the SEC complex that specifically associates with the Sec61 translocon (Finke et al., 1996), and to *SSH1*, a paralog of the canonical Sec61 translocon. For the mitochondrial studies, we used a BirA fusion to *OM45*, a major constituent of the mitochondrial outer membrane (MOM). In mammalian cells, we utilized a BirA fusion to Sec61 β that uniformly labeled the ER (Shibata et al., 2008). In all cases, the BirA fusion proteins showed the expected localization (Fig. 3-1C; Fig. 2-1B).

c. Inducing biotinylation

Because of the potential cycling of ribosomes between different cellular localizations, particularly following translation termination, it was critical to be able to induce rapid ribosome biotinylation while also suppressing constitutive background BirA activity. Biotin is an essential co-factor for the synthesis of fatty acids, and thus cannot be omitted from growth media. However, by titrating biotin levels in the growth media we were able to suppress BirA activity to undetectable levels without impacting cell growth (Fernández-

Suárez et al., 2008) (Fig. 2-2A).

An alternative approach for controlling biotinylation activity is through the introduction of mutations, either to the BirA enzyme or AviTag substrate, which increase the K_m for biotin in the reaction above intracellular levels dictated by standard growth media. We explored this approach using several previously characterized BirA mutants. Although several mutants did reduce background biotinylation, the negative effects of the mutations on the k_{cat} of labeling were found to be prohibitive (Fig. 2-2C). Similar results were obtained when testing mutations to the AviTag itself (Fernández-Suárez et al., 2008) (unpublished data). These observations led us to favor defined concentrations of biotin in the media as the most robust approach for rapidly inducing biotinylation in our system.

Indeed, brief biotin pulses were sufficient to give a robust biotinylation signal in live cells grown in biotin-defined media (Fig. 2-2B; Fig. 3-1E) Together, this procedure allowed us to achieve rapid (on the timescale of polypeptide synthesis) and efficient biotinylation of both our 40S and 60S Avi-tagged ribosomes using a cytosolic BirA. In marked contrast, ER-localized BirAs failed to label the 40S Avi-tagged ribosomal subunit but retained the ability to robustly label the 60S Avi-tagged subunit (Fig. 3-1E). Based on the length of our BirA tether, which is too short to allow biotinylation of the 40S subunit of a docked, translocating ribosome, this result demonstrates the specific biotinylation of oriented translocating ribosomes over those that passively encounter the ER membrane.

d. Quenching post-lysis biotinylation

Preserving the pattern of biotinylation established *in vivo* was paramount to the success of our method. However, subsequent to cell harvesting and cryogenic lysis, a substantial amount of non-specific biotinylation was observed to occur in lysates (Fig 3-1D; Fig. 2-3). We evaluated three different approaches for quenching the observed post-lysis biotinylation:

competitive inhibition using excess Avi-peptide, chelation of biotin using streptavidin, and removal of ATP and biotin through by lysate de-salting (Fig. 2-3). Of the approaches tested, lysate de-salting proved to be the most effective and straightforward approach for attenuating non-specific labeling.

e. Isolating biotinylated ribosomes

Finally, after isolation of all cellular monosomes from lysates using the standard ribosome profiling protocol, we sought to specifically purify biotinylated ribosomes. Key steps in our optimized pulldown procedure include the use of high-affinity magnetic streptavidin beads over the porous and thus less-specific agarose-based equivalent, high-salt washes to reduce background, and specific elution via a TEV protease site engineered into our AviTag handle (Fig. 2-4).

CONCLUSION

With the experimental components of proximity-specific ribosome profiling in place, we next sought to apply our method to comprehensively characterize subcellular translation at two fundamental cellular compartments: the ER and the mitochondria. This is the focus of Chapters III and IV, where we explore several features of protein targeting to these organelles that have remained difficult to assay with other methodologies. Together, these chapters establish the utility of the proximity-specific ribosome profiling method as a comprehensive yet high-precision tool for gaining insight into the complex principles governing protein targeting and folding *in vivo*, and showcase the broad spectrum of experimental questions whose study it enables.

FIGURES

Figure 2-1

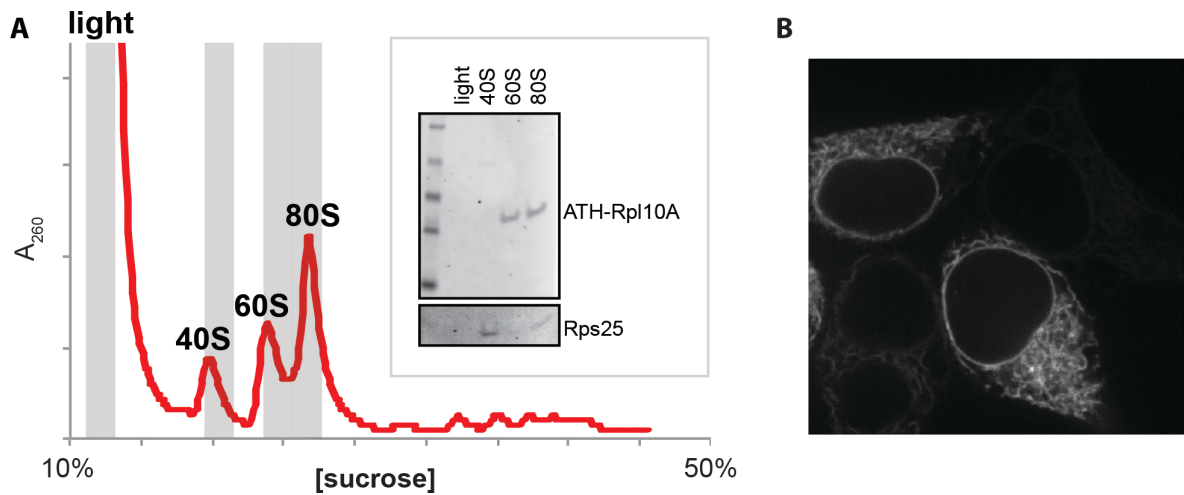


Fig. 2-1. Components for proximity-specific ribosome profiling in 293T cells (A) Polysome profile from 293T cells expressing Avi-TEV-HA-Rpl10a. Fractions analyzed by western blot are indicated with vertical bars. (Inset) western blot analysis of sucrose gradient fractions. ATH-Rpl10A was found only in the 60S and 80S fractions, consistent with incorporation into functional ribosomes. Rps25 is shown below as an endogenous small subunit control. **(B)** Confocal microscopy of BirA-mCherry-Sec61 β .

Figure 2-2

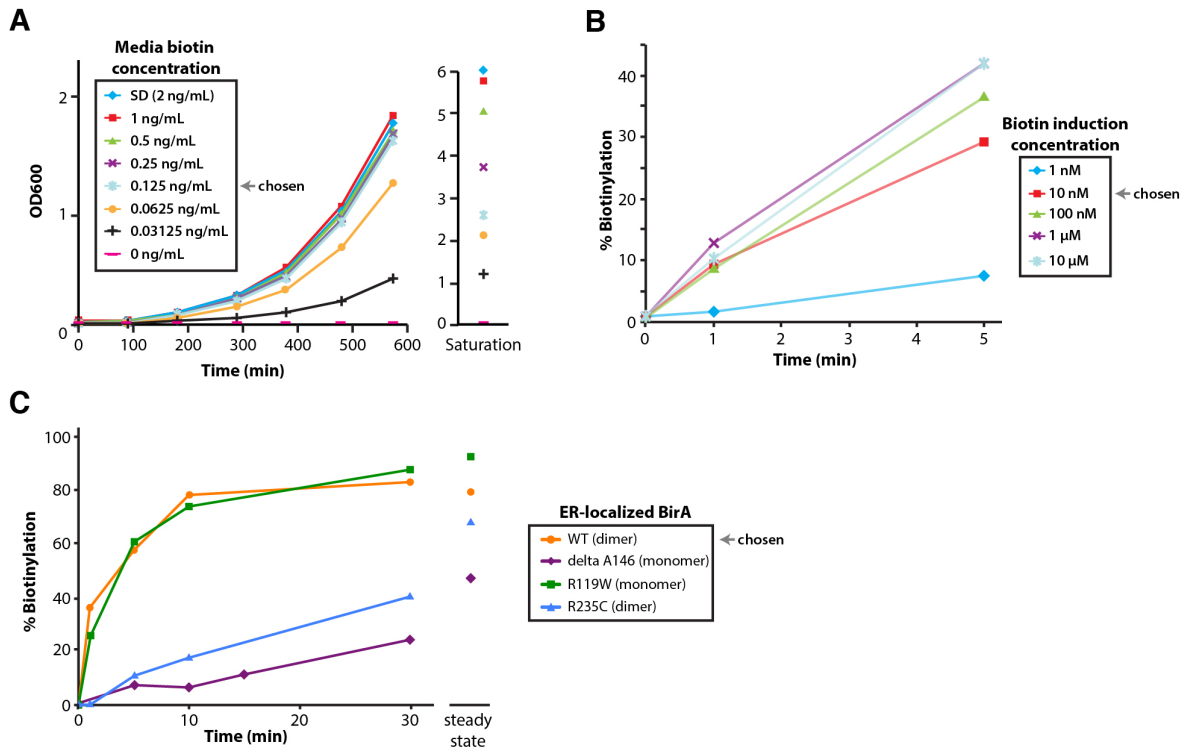
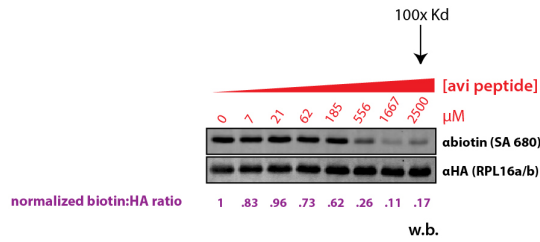
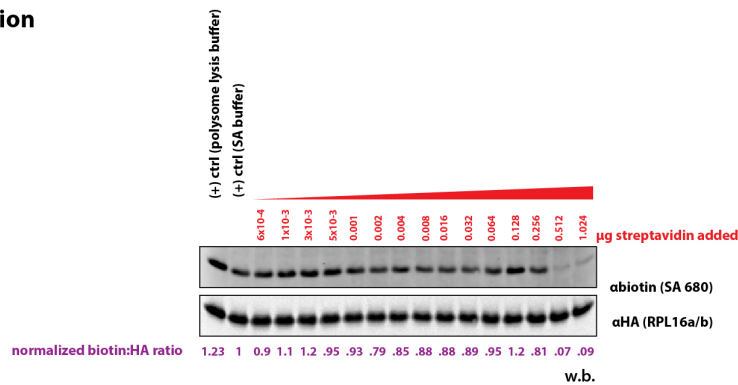


Figure 2-2 Inducing biotinylation in yeast. (A) Growth curves of yeast cells grown in SD –biotin media supplemented with the indicated concentrations of D-biotin. (B) Time course of Rpl16 biotinylation by Sec63-BirA. The growth media was supplemented with D-biotin to the indicated concentration at time zero. (C) The effect of several characterized BirA mutations on the kinetics of Rpl16 labeling. BirAs were expressed under the control of a PGK1 promoter and were localized to the ER using the Ubc6 tail anchor sequence.

A Competitive inhibition



B Biotin chelation



C Lysate de-salting

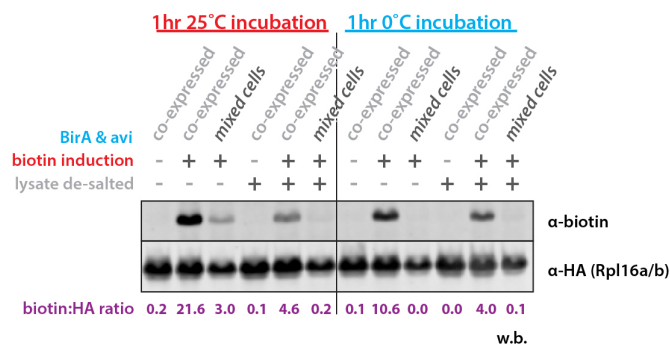


Figure 2-3. Control of post-lysis biotinylation in yeast. (A) Lysates derived from cells expressing pPGK1-BirA were mixed with separate lysates derived from cells expressing Avi-tagged RPL16. Prior to mixing, each extract was mixed with half the indicated quantity of Avi peptide. Mixed lysates were incubated for one hour at 25°C as a mock RNaseI digest and the extent of post-lysis biotinylation evaluated. Biotin:HA ratios were normalized to 0 μM peptide. (B) As in (A), but each extract was mixed with half the indicated quantity of streptavidin prior to mixing. (C) as in (A), except pPGK1-BirA and Avi-tagged RPL16 were either co-expressed in the same cell, or separate cells and mixed prior to lysis. After [mixing,] cryogenic lysis and clarification, the indicated extracts were desalted to remove free ATP and biotin, and incubated as indicated. In all cases, cells were induced with 10 nM biotin for seven minutes prior to harvesting.

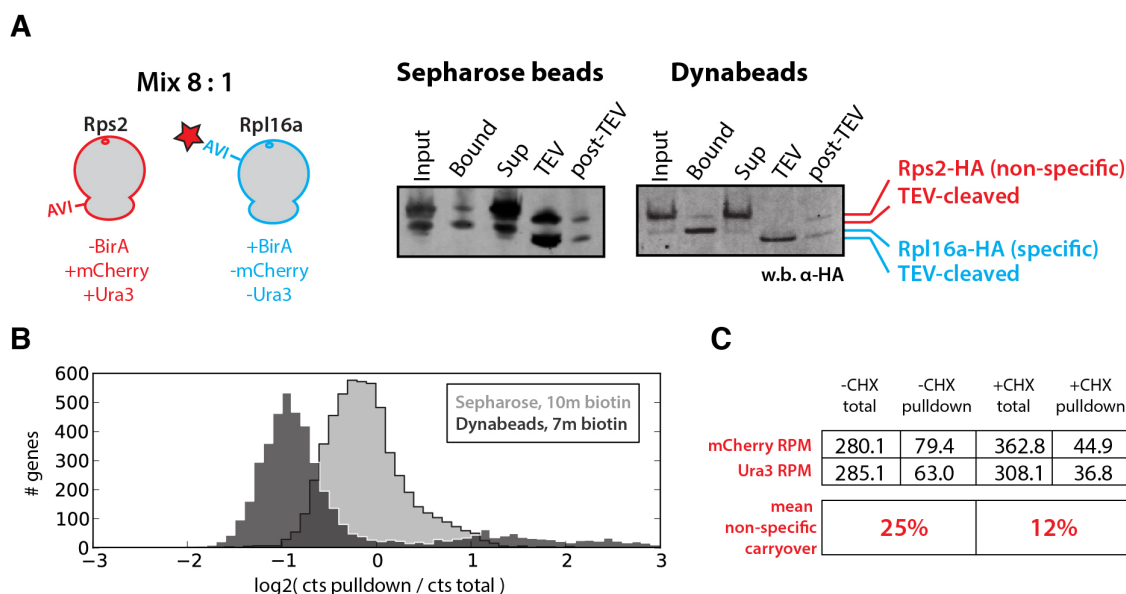


Fig. 2-4. Pulldown specificity. (A) Extracts from cells expressing Avi-tagged Rps2, mCherry, and Ura3 in the absence of BirA or Avi-tagged Rpl16a in the presence of BirA were digested with RNaseI and the resulting monosomes fractionated on a 10-50% sucrose gradient. Avi-tagged Rps2 monosomes were mixed with biotinylated Avi-tagged Rpl16a monosomes in an 8:1 molar ratio. Biotinylated ribosomes were purified on Streptactin sepharose (IBA) or MyOneC1 streptavidin dynabeads (Invitrogen). Contamination was assessed by monitoring the presence of Rps2 throughout the pulldown. (B) Histograms of gene-level enrichments obtained from proximity-specific ribosome profiling experiments (BirA-mVenus-Ubc6) using either Sepharose or Dynabeads for ribosome pulldowns as indicated. Dynabeads yield cleaner separation of the enriched and dis-enriched populations (C) as in (A), monosomes from cells expressing Avi-tagged Rps2, mCherry, and Ura3 in the absence of BirA were spiked into proximity-specific ribosome profiling input monosomes at 1:9 stoichiometry to Avi-tagged Rpl16 in the presence of BirA-mVenus-Ubc6 and profiled using Dynabeads. Non-specific footprint carryover (undetectable by Western blot) ranged from 12-25%, or -2 to -3 log₂ units.

REFERENCES

- Ban, N., Beckmann, R., Cate, J.H.D., Dinman, J.D., Dragon, F., Ellis, S.R., Lafontaine, D.L.J., Lindahl, L., Liljas, A., Lipton, J.M., et al. (2014). A new system for naming ribosomal proteins. *Curr. Opin. Struct. Biol.* *24*, 165–169.
- David, A., Dolan, B.P., Hickman, H.D., Knowlton, J.J., Clavarino, G., Pierre, P., Bennink, J.R., and Yewdell, J.W. (2012). Nuclear translation visualized by ribosome-bound nascent chain puromycylation. *J. Cell Biol.* *197*, 45–57.
- Dieterich, D.C., Hodas, J.J.L., Gouzer, G., Shadrin, I.Y., Ngo, J.T., Triller, A., Tirrell, D.A., and Schuman, E.M. (2010). In situ visualization and dynamics of newly synthesized proteins in rat hippocampal neurons. *Nat. Neurosci.* *13*, 897–905.
- Fernández-Suárez, M., Chen, T.S., and Ting, A.Y. (2008). Protein–Protein Interaction Detection in Vitro and in Cells by Proximity Biotinylation. *J. Am. Chem. Soc.* *130*, 9251–9253.
- Finke, K., Plath, K., Panzner, S., Prehn, S., Rapoport, T.A., Hartmann, E., and Sommer, T. (1996). A second trimeric complex containing homologs of the Sec61p complex functions in protein transport across the ER membrane of *S. cerevisiae*. *EMBO J.* *15*, 1482–1494.
- Heiman, M., Schaefer, A., Gong, S., Peterson, J.D., Day, M., Ramsey, K.E., Suárez-Fariñas, M., Schwarz, C., Stephan, D.A., Surmeier, D.J., et al. (2008). A Translational Profiling Approach for the Molecular Characterization of CNS Cell Types. *Cell* *135*, 738–748.
- Ingolia, N.T. (2014). Ribosome profiling: new views of translation, from single codons to genome scale. *Nat. Rev. Genet.* *15*, 205–213.
- Ingolia, N.T., Ghaemmaghami, S., Newman, J.R.S., and Weissman, J.S. (2009). Genome-Wide Analysis in Vivo of Translation with Nucleotide Resolution Using Ribosome Profiling. *Science* *324*, 218–223.
- Jiang, Y., Cheng, Z., Mandon, E.C., and Gilmore, R. (2008). An interaction between the SRP receptor and the translocon is critical during cotranslational protein translocation. *J. Cell Biol.* *180*, 1149–1161.
- Kornmann, B., Currie, E., Collins, S.R., Schuldiner, M., Nunnari, J., Weissman, J.S., and Walter, P. (2009). An ER-Mitochondria Tethering Complex Revealed by a Synthetic Biology Screen. *Science* *325*, 477–481.
- Ben-Shem, A., Garreau de Loubresse, N., Melnikov, S., Jenner, L., Yusupova, G., and Yusupov, M. (2011). The Structure of the Eukaryotic Ribosome at 3.0 Å Resolution. *Science* *334*, 1524–1529.
- Shibata, Y., Voss, C., Rist, J.M., Hu, J., Rapoport, T.A., Prinz, W.A., and Voeltz, G.K. (2008). The Reticulon and Dp1/Yop1p Proteins Form Immobile Oligomers in the Tubular Endoplasmic Reticulum. *J. Biol. Chem.* *283*, 18892–18904.

Stephens, S.B., and Nicchitta, C.V. (2007). In vitro and tissue culture methods for analysis of translation initiation on the endoplasmic reticulum. *Methods Enzymol.* 431, 47–60.

CHAPTER THREE

Principles of co-translational translocation at the endoplasmic reticulum

INTRODUCTION

The biochemical mechanisms by which secretory proteins translocate into the ER are well-characterized and have been elucidated over decades, primarily through *in vitro* studies that focused on a single pathway and a highly limited number of substrates. Our proximity-specific ribosome profiling approach provides a high-resolution systems-level synthesis of how these mechanisms function together *in vivo* to support efficient co-translational translocation. We integrate our findings to substantially build on (and even rethink) the textbook model for translation at the ER providing a more sophisticated understanding of how efficient targeting of diverse substrates is achieved within the cell.

This view was enabled by both the depth and precision of our measurements. At the gene-level, our analysis demonstrated that co-translational insertion into the ER is principally determined by the location of the signal sequence within the protein, irrespective of the protein's dependence on the signal recognition particle (SRP), which was widely thought to be required for co-translational insertion. The codon-level resolution intrinsic to our data revealed the precise nascent-chain requirements for each secretory or transmembrane protein to engage the translocon *in vivo*. We found that ribosomes interact with the SRP-dependent translocon as soon as the signal sequence emerges from the peptide exit tunnel and, surprisingly, that ribosomes interact with the SEC complex while the signal sequence is still in the exit tunnel. This analysis also uncovered a novel class of proteins that engage the translocon in the looped-conformation required for signal sequence processing, whereas all other proteins undergo a major conformational rearrangement within the translocon to achieve this topology.

RESULTS

Validation of proximity-specific ribosome profiling

We performed proximity-specific ribosome profiling in *S. cerevisiae* using the three different ER-localized BirA constructs as well as the cytosolic and mitochondrially-localized controls (Fig. 3-2A), and in mammalian HEK293 cells using an ER-localized BirA fusion protein. For each experiment, brief treatment with the translation elongation inhibitor cycloheximide (CHX), which preserves the ribosome position along an mRNA, was followed by a biotin pulse. Subsequent to processing and sequencing, we determined an enrichment value for each gene by taking the \log_2 ratio of ribosome footprint densities in the matched streptavidin-pulldown versus input whole-cell ribosome profiling samples. Enrichment metrics obtained from the same BirA were highly reproducible between replicates (Fig. 3-2B; Ssh1 Pearson $r = 0.97$; Sec63 Pearson $r = 0.98$).

Targeting of BirA to the cytosol yielded a narrow range of enrichment values (90% of genes fell within -0.2 to $+0.2$ \log_2 enrichment units) demonstrating that our protocol for isolating biotinylated monosomes introduced minimal bias. We detected a modest but highly significant ($p < 1 \times 10^{-15}$, KS-test) depletion of secreted genes, consistent with the expected lower accessibility of ER-docked ribosomes. By contrast, BirA targeted to mitochondria produced a clear bimodal distribution, enriching for genes annotated to localize to this cellular compartment (Elstner et al., 2009). An in-depth analysis of translation at the MOM is presented in an accompanying manuscript (Williams et al.). Targeting of BirA to the ER membrane inverted the mitochondrial enrichment pattern, cleanly separating secreted proteins from those synthesized in the cytosol or targeted to mitochondria. Ssh1, Sec63, and Ubc6 ER-localized BirA fusion constructs all labeled ribosomes translating similar sets of secretome genes [defined in (Ast et al., 2013)], though we observed striking differences in the point during translation at which RNCs interact with these BirA fusions (explored below). To determine whether other gene categories were significantly overrepresented in the enriched populations of our ER datasets, we performed GO-term analysis on gene categories

in yeast and mammalian cells that were enriched above a threshold derived from a receiver operator characteristic analysis (Fig. 3-2C; Fig. 3-2D; Fig. 3-S1). In both yeast and HEK293 cells, enriched gene sets were exclusively from the secretome. However, a substantial number of mammalian transcripts predicted by Phobius (Käll et al., 2004) to encode secretory proteins were not enriched in our assay. This set of genes was enriched in GO-terms for mitochondria and other locations, arguing that these proteins represent false positives in the computationally-predicted secretory gene set. This discrepancy serves to highlight both the sensitivity and utility of our approach for experimentally defining proteins that are targeted to specific cellular compartments.

We noted that peroxisomal proteins exhibited heterogeneous ER translational enrichment. The peroxisome is a highly conserved organelle responsible for lipid catabolism whose mechanism of biogenesis has been controversial. Specifically, there is evidence for both *de novo* peroxisome generation from ER-derived vesicles, as well as for derivation from pre-existing peroxisomes through growth and fission (Smith and Aitchison, 2013). Our data reveal that 16 of 54 yeast peroxisomal proteins showed clear co-translational enrichment. Consistent with previous targeted studies in yeast (van der Zand et al., 2010), a unifying determinant for this ER-targeting is the presence of one or more TMDs –we found no evidence for the enrichment of peroxisomal matrix proteins. Importantly, this partitioning was also seen in mammalian cells (Fig. 3-S2). Our data thus suggest that the peroxisome obtains TMD-containing proteins from the ER and matrix proteins exclusively from the cytosol across cell types.

Co-translational targeting of SRP-dependent and -independent substrates in vivo

While a subset of proteins are strictly reliant on SRP for ER targeting, a process which is thought to be obligatorily co-translational, import of other proteins occurs efficiently without

SRP when measured both *in vitro* and *in vivo* (Ast et al., 2013; Ng et al., 1996). A recent *in silico* analysis revealed that roughly 40% of yeast secretome substrates use the less-studied SRP-independent pathway (Ast et al., 2013). SRP-independent translocation depends on translocon accessory factors as well as the luminal chaperone Kar2/BiP and *in vitro* can occur efficiently after translation (Panzner et al., 1995; Young, 2001) (Fig. 3-3A).

Remarkably, we found that the vast majority of secretory proteins undergo co-translational targeting *in vivo*, irrespective of their dependence on SRP (Fig. 3-3B). This pattern held for 162 genes experimentally validated as SRP-dependent or -independent (Ast et al., 2013; Ng et al., 1996), as well as for an additional 756 genes whose SRP-dependence was predicted using a hydrophathy-based analysis (Ast et al., 2013).

It was a formal possibility that the apparent co-translational ER enrichment of these translating messages was a result of brief treatment with the translation elongation inhibitor (CHX) prior to biotinylation, as this provides extra time for the RNC complex to engage the translocon. We evaluated this possibility by omitting translation inhibitors and labeling with biotin for one minute, a timescale comparable to a single round of polypeptide synthesis. For the large majority of SRP-dependent and -independent substrates, levels of translational enrichment were not dependent on CHX (Fig. 3-3B). Intriguingly, ribosomes translating a small minority of the SRP-independent proteins lost their enrichment, suggesting that in an unperturbed setting these proteins translocate post-translationally.

We thus conclude that SRP-independence is not synonymous with post-translational translocation; import concurrent with protein synthesis is the principal route into the ER *in vivo*. By effectively coupling translation and translocation for the large majority of proteins entering the secretory pathway, the cell minimizes the possible issues associated with having a cytosolic cohort of un-translocated, aggregation-prone proteins (Brandman et al., 2012).

Understanding how the cell achieves co-translational translocation of SRP- independent messages remains an open question.

Comprehensive analysis of co- versus post-translational translocation in vivo

Having uncoupled SRP-independence from post-translational translocation, we sought to better understand the determinants for partitioning between the co- and post-translational import pathways. To classify genes based on their ER translational enrichment, we systematically identified genes whose enrichments were dependent on CHX using a support vector machine (SVM) classifier trained to distinguish between roughly 140 proteins characterized empirically as being CHX-dependent or -independent (see experimental methods). This SVM analysis enabled us to systematically characterize the import of proteins as being either co-translational (CHX-independent), co-translational translocation that is dependent upon (or enhanced by) treatment with a translation inhibitor (CHX-dependent), or obligatorily post-translational (dis-enriched) (Fig. 3-3C).

Consistent with the above observations (Fig. 3-3B), this analysis revealed that the large majority (681 of 837) of Phobius-predicted secretory genes are translated at the ER independent of CHX. Of the remaining predicted secretory proteins, 63 were dependent on CHX for enrichment, whereas 93 were not enriched under any condition tested. The latter dis-enriched group contained nearly all of the roughly 50 annotated TA proteins whose C-terminal TMDs preclude co-translational recognition and that are known to be targeted to the ER post-translationally through the GET pathway (Schuldiner et al., 2008; Stefanovic and Hegde, 2007).

What then accounts for the remaining proteins whose translocation is not strictly co-translational? The position of an ER-targeting signal within a protein imposes restrictions on when during synthesis targeting may occur, and thus might be an important determinant.

Indeed, robustly enriched CHX-independent secreted genes and dis-enriched TA genes fall on the opposite sides of this spectrum, with extreme N- and C-terminal ER targeting signals, respectively (Fig. 3-3D). Interestingly, the targeting signals of proteins dependent on CHX for co-translational targeting to the ER fall in-between these two extremes, often present far downstream in relation to their overall gene length. In regard to co-translational recognition of their targeting signals, these proteins likely face constraints similar to those of TA genes in which there is little or no time for the ribosome to reach the ER surface after synthesis of the hydrophobic TMD. Given the broad spectrum of signal positions near the C-terminus, the partitioning of these between the SEC and GET machineries is an interesting question.

Our SVM classification also revealed co-translational, CHX-independent ER enrichment of 70 genes for which no hydrophobic domains were detected by Phobius (Fig. 3-3C). However, the majority of these were predicted to contain a hydrophobic domain by alternate hydrophobic prediction algorithms (TMHMM or SignalP) (Krogh et al., 2001; Petersen et al., 2011) (Fig. 3-3E). These genes thus likely represent genuine secretory proteins that were missed by Phobius, highlighting the value of our studies as an experimental complement to computational algorithms for globally identifying secreted and transmembrane proteins.

Timing and specificity of co-translational targeting to the ER

We next asked when during translation RNCs are recruited to the ER, which is expected to depend upon the mechanism of recruitment. For example, SRP binds preferentially to short nascent chains containing cytosolically-accessible SSs (Siegel and Walter, 1988) and halts translation elongation until the RNC reaches the ER. By contrast, SRP-independent transport through the SEC complex relies on a poorly-understood network of cytosolic chaperones, none of which are known to arrest translation.

In yeast, translocation occurs through two paralogous channels, Sec61 and Ssh1. The essential Sec61 translocon associates either with several accessory factors, including both essential (Sec63 and Sec62) and non-essential (Sec66 and Sec72) peripheral components to form the SEC complex, or separately with the SRP receptor (SR). By contrast, the non-essential Ssh1 is a simpler translocon thought to interact peripherally only with SR (Jiang et al., 2008). We reasoned that fusing BirA to specific complexes would allow us to globally monitor the timing and specificity of translocation of substrates through these distinct translocons (Fig. 3-4A).

Ssh1

Ssh1 is expected to receive RNCs exclusively from SRP and should therefore only interact with the ribosomes after the SS emerges from the ribosome peptide exit tunnel. Consistent with this model, for all secreted and TMD proteins regardless of the location of their targeting sequence, BirA-Ssh1 enrichment begins only after the hydrophobic domain is fully accessible (~60 amino acids from the start of targeting sequence) (Kowarik et al., 2002) (Fig. 3-4B; Fig. 3-S3A).

Type II signal anchors (SAs) and cleavable SSs are oriented in a looped conformation with their N-termini facing the cytosol (Fig. 3-4C). Models of how SSs and SAs achieve this topology within the translocon differ in the efficiency of RNC recruitment; i.e. the head-first model stipulates early RNC binding and subsequent signal inversion, whereas looped insertion model requires delayed RNC binding in which the nascent chain is correctly oriented before binding (Fig. 3-4C). Studies of model substrates are consistent with the head-first model (Devaraneni et al., 2011; Goder, 2003). Efficient targeting to Ssh1 immediately following SS translation would necessitate such a mechanism. Our studies revealed that the timing of engagement was bimodally distributed: the majority of the secretome is efficiently

recruited to Ssh1 immediately following exit of the SS from the ribosome, however, a prominent subset of proteins with cleavable SSs engaged only after enough synthesis (~120 amino acids) to allow the nascent chain to acquire a looped topology (Fig. 3-4B,C). These results argue that both head-first and looped-insertion occur *in vivo*, depending on the protein. Remarkably, we found that deletion of *SEC66*, a nonessential component of the SEC complex, exclusively impacts translocation of the newly identified looped-insertion substrates (Fig. 3-4C-E, Fig. 3-S4). These results illustrate that proximity-specific ribosome profiling is useful in deciphering how distinct translocon components enable the efficient handling of diverse targeting sequences and topologies (Devaraneni et al., 2011; Fons, 2003; Görlich et al., 1992).

Sec63

Sec63 mediates translocation in an SRP-independent manner. Without SRP-mediated translational pausing in the cytosol, Sec63 translocation naïvely might have been expected to result in a delayed and more broadly distributed timing of ER targeting. Surprisingly, the opposite was observed; Sec63 begins to interact with RNCs translating secretory proteins well before the emergence of the SS from the ribosome (Fig. 3-4F) and maximal engagement occurs shortly after the SS is fully solvent accessible. These data suggest that Sec63 interacts with ribosomes through two distinct modes: one depends on the presence of an accessible SS and the second reflects an interaction with ribosomes while the targeting sequence is in the exit tunnel. Consistent with this interpretation, acute loss of SRP function using a temperature-sensitive SRP allele (Fig. 3-S5) did not impact the early engagement, but did compromise the late (post-SS emergence) enrichment (Fig. 3-4G).

Dynamics of ER-associated ribosomes

Upon translation termination, ER-associated ribosomes can either immediately dissociate and return to the pool of cytosolic ribosomes or preferentially undergo multiple rounds of translation on ER-associated mRNAs. To investigate these dynamics in the context of living cells, we harvested samples for proximity-specific profiling after increasing lengths of biotinylation time in the absence of CHX. Enrichment for secretory messages is expected to decrease at a rate proportional to the timescale at which biotinylated ribosomes, originating from the ER, exchange into the cytosol and begin translating cytosolic messages (Fig. 3-5A). We observed rapid collapse of our bimodal enrichment distribution into a single population on the order of minutes, although secreted messages remain on the enriched side of the distribution at all time points tested as expected from continual biotinylation of ER ribosomes (Fig. 3-5B). Based on a median gene length of ~425 codons and a translation rate of ~5.5 codons per second (Bonven and Gulløv, 1979), translation of a single secretory protein is expected to take ~77 seconds. We thus conclude from these observations that ribosomes at the yeast ER are highly dynamic, freely exchanging into the cytosol within at most a few rounds of translation.

DISCUSSION

Here we present a proximity-based ribosome profiling strategy that allows one to monitor translation for any location at which it is possible to target a BirA fusion protein. We apply this strategy to dissect modes of co-translational translocation into the ER. Nearly one quarter of the proteome is imported into the ER – accordingly, this process has been the focus of intense research. Much of this previous work, however, has explored the behavior of a small group of model substrates often outside of a cellular context. Proximity-specific ribosome profiling allowed us to simultaneously probe the ER engagement of nascent chains across the full proteome *in vivo*, in the context of competing and redundant trafficking

pathways. This comprehensive characterization revealed several principles of how cells integrate distinct targeting pathways with the translocation machinery to allow for unexpectedly robust co-translational ER import of a diverse set of substrates.

Foremost is the critical role of the timing of translation of the ER targeting sequence relative to translation termination for determining the propensity of a protein to undergo import co-translationally. This stands in contrast to the view that co-translational import is dictated by the factors which mediate targeting (e.g., SRP). It had previously been appreciated that TA proteins must insert post-translationally since the targeting sequence is obscured prior to translation termination. However, these represent a single point on a broader spectrum of signal positions. Proteins with targeting domains near the C-terminus typically are targeted as RNCs only when the kinetics of translation are crippled (Fig. 3-3D). By contrast, the predominantly SRP-independent (Ast et al., 2013) set of substrates containing N-terminal SSs are robustly co-translationally targeted. Congruently, we also find that the full range of co-translational substrates are able to engage both the essential SEC and 'alternate' Ssh1 translocons. Moreover, we find that RNCs are able to interact with the SEC complex prior to SS exposure, despite the fact there are no known mechanisms for coordinating translation and recruitment to SEC.

The above findings suggest a model wherein a pioneering round of translation is responsible for recruiting the RNC to the ER surface, following which the message remains tethered to the ER by ongoing translation by downstream ribosomes (Fig. 3-5C). Consistent with our understanding of SRP function, SRP likely plays a critical role in establishing specificity and ensuring translocation competency through its ability to halt translation. Subsequent rounds of ribosome initiation in the context of this ER-tethered mRNA would obviate the need for SRP to survey every translation event, particularly for messages with extensive downstream regions that can accommodate multiple ribosomes. Such a mechanism

is consistent with the observed 1:50 stoichiometry of SRP to the ribosome (Ogg and Walter, 1995), and would simplify the problem of cellular protein sorting while minimizing the toxicity associated with solvent-exposed hydrophobic domains. Indeed, inhibition of translocation is known to induce a massive cytoplasmic stress response (Brandman et al., 2012) underscoring the danger of having ER-targeted proteins in the cytosol even when they can be post-translationally translocated.

Our studies also revealed a novel class of signal sequences that emphasize an intimate connection between the timing of import and protein topology, mediated by translocon accessory factors. The bimodal timing of targeting to Ssh1 suggests that insertion can occur in either a head-first or looped orientation. Our results implicate Sec66 in mediating the import of those proteins that undergo looped insertion. This functionality may be necessary for certain substrates whose insertion kinetics would preclude re-orientation within the translocon. A clear future application of our method is probing the elusive roles of other translocon accessory factors, such as TRAM and TRAP in mammals (Fons, 2003; Görlich et al., 1992).

A final principle that emerged from our studies is the dynamic nature of ER-associated ribosomes in yeast, which cycle readily between cellular compartments. This is in contrast to evidence from *in vitro* exchange experiments that showed stable association of the 60S ribosome subunit with the mammalian ER (Borgese et al., 1973). It will be interesting to explore whether such a pattern holds in more specialized secretory cells, such as plasma cells, which rely on efficient translation at the ER. Indeed, electron micrographs have revealed the presence of circular polysomes (Christensen et al., 1987) in these cells, consistent with a 'closed-loop model' of translation (Wells et al., 1998) that is presumed to promote efficient translation re-initiation.

The principles uncovered here highlight the ability of proximity-specific ribosome profiling to synergize with prior mechanistic studies of ER trafficking pathways. Diverse biological systems localize mRNAs to generate cellular structure and function, yet compared to the ER, much less is known about how co-translational protein trafficking contributes to asymmetry at these sites. More generally, this approach enables the profiling of subpools of ribosomes that interact, even transiently, with cellular proteins of interest, e.g. those involved in protein folding, quality control, trafficking, and post-translational modification. As a flexible, precise, and global method, proximity-specific ribosome profiling provides a new tool for exploring the interface between translation and cell biology.

Acknowledgements

We thank members of the Weissman and P. Walter labs for discussion; E. Costa, M. Kampmann, P. Kimmig, K. Kostova, and M. Smith for critical reading of the manuscript. Technical assistance was provided by the UCSF CAT, NIC and the CFR. This research was supported by the Center for RNA Systems Biology and the Howard Hughes Medical Institute. C.H.J. is the Rebecca Ridley Kry Fellow of the Damon Runyon Cancer Research Foundation (DRG-2085-11). C.C.W. is an NSF GRF.

FIGURES

Figure 3-1

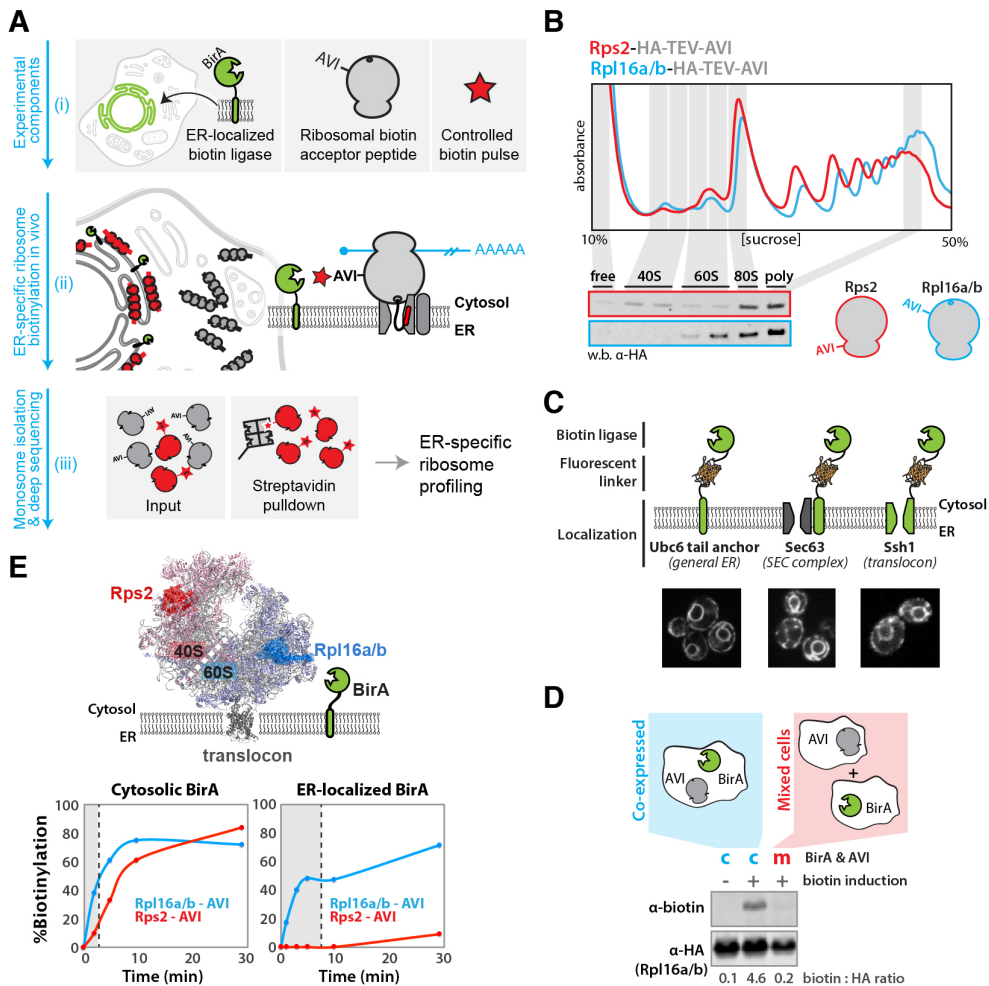


Fig. 3-1. A system for in vivo proximity-dependent ribosome biotinylation to monitor local protein synthesis at the ER. (A) Schematic for proximity-specific ribosome profiling (i) The *E. coli* biotin ligase BirA is localized to a subcellular site of interest in cells expressing an Avi-tagged ribosomal protein and grown in low-biotin conditions (ii) A biotin pulse is applied resulting in specific biotinylation of ribosomes in close physical proximity to the localized BirA (iii) Ribosome profiling of paired input (grey and red) and isolated biotinylated (red) monosomes reveals codon-resolved translational enrichment specific to the BirA locale. (B) Fractionation of yeast lysates derived from strains containing scarless C-terminal Rps2 or Rpl16a/b HA-TEV-AviTags on 10-50% sucrose gradients. Polysome traces demonstrate proper ribosomal assembly and incorporation of tags into polysomes demonstrates their non-perturbative nature. (C) ER localization of BirA fusion proteins used in this study. BirA-mVenus-Ubc6, Sec63-mVenus-BirA and BirA-mVenus-Ssh1 all localize to the perinuclear and cortical ER. (D) Western blot analysis demonstrating that biotinylation of ribosomal AviTags does not occur before the addition of excess biotin nor post-lysis in our assay. (E) Biotinylation kinetics of 40S and 60S AviTags by BirAs localized to the cytosol or ER (Sec63). Favorable kinetics are achieved independent of localization and preferential 60S biotinylation demonstrates the specificity of the ER-localized ligase for oriented ER ribosomes. Shaded regions indicate biotinylation times used in subsequent sequencing experiments.

Figure 3-2

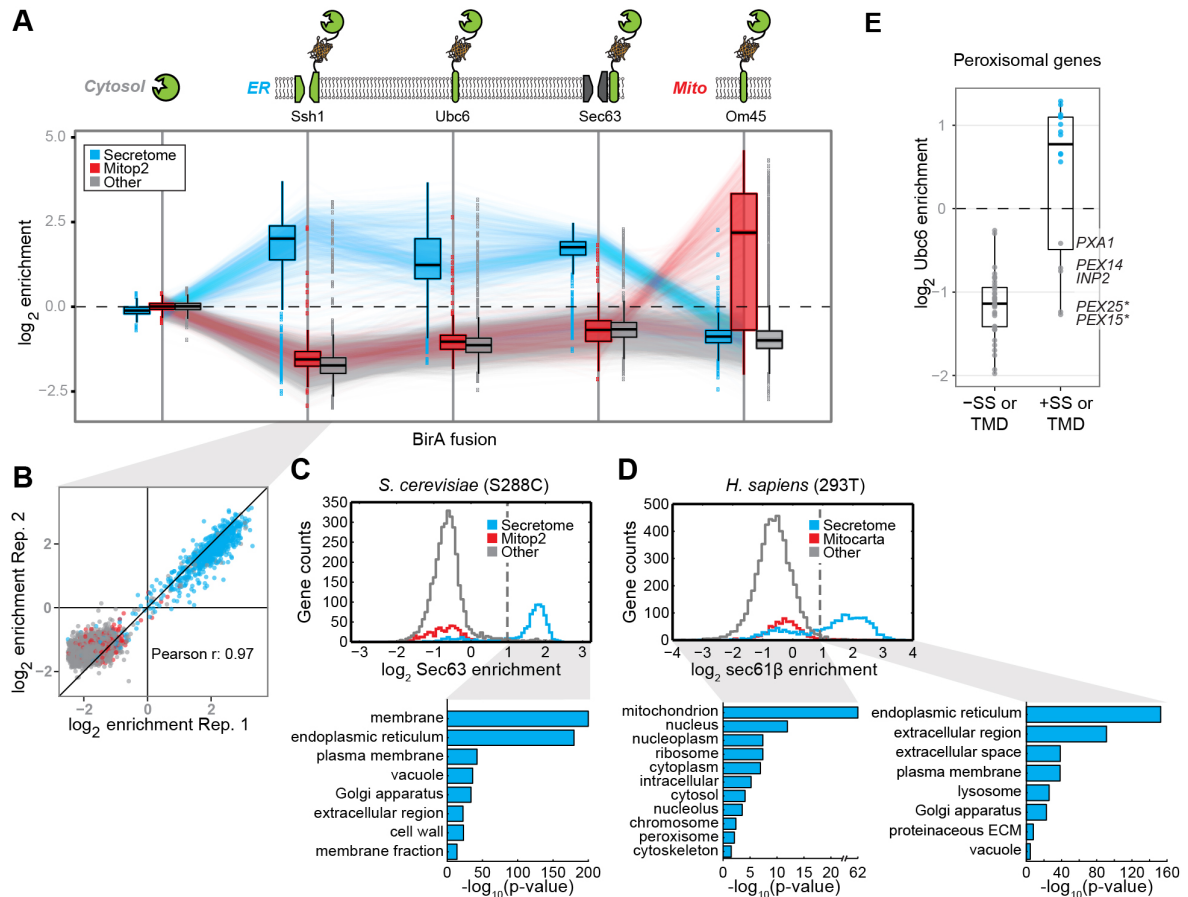


Fig. 3-2. Specificity of proximity-dependent ribosome profiling across multiple systems. (A) Boxplots of the log₂ enrichment distributions for secretome (blue), curated mitochondrial (red), and all other (grey) gene categories obtained from proximity-specific ribosome profiling experiments in yeast using different BirA fusions. Biotinylation was carried out in the presence of CHX for two minutes (cytosolic, mitochondria) or seven minutes (ER). Enrichments were computed for each reliably-expressed gene as the log₂ ratio of biotinylated footprint density (RPM) over the corresponding density from the matched input whole-cell ribosome profiling experiment. Where possible, lines connect the same gene across experiments. **(B)** Enrichments shown for representative proximity-specific ribosome profiling replicates using the BirA-Ssh1 fusion protein. Colors match those in (A). **(C)** Histograms of log₂ enrichments for Sec63-BirA in yeast. Enrichment thresholds were determined by ROC analysis (Fig. 3-S1). Shown below are the corresponding enrichment analyses of GO-slim cellular components for robustly enriched genes versus expressed genes. Colors match those in (A). **(D)** As in (C) for BirA-Sec61β in HEK293T cells. Additionally, GO-term analysis of dis-enriched secretome genes versus expressed secretome genes is shown. **(E)** Gene enrichments obtained with the general BirA-Ubc6 ER marker in yeast, for well-expressed CHX-independent peroxisomal genes. SS and TMD annotations were predicted by SignalP and TMHMM, respectively. * denotes necessarily post-translational tail-anchor TMDs.

Figure 3-3

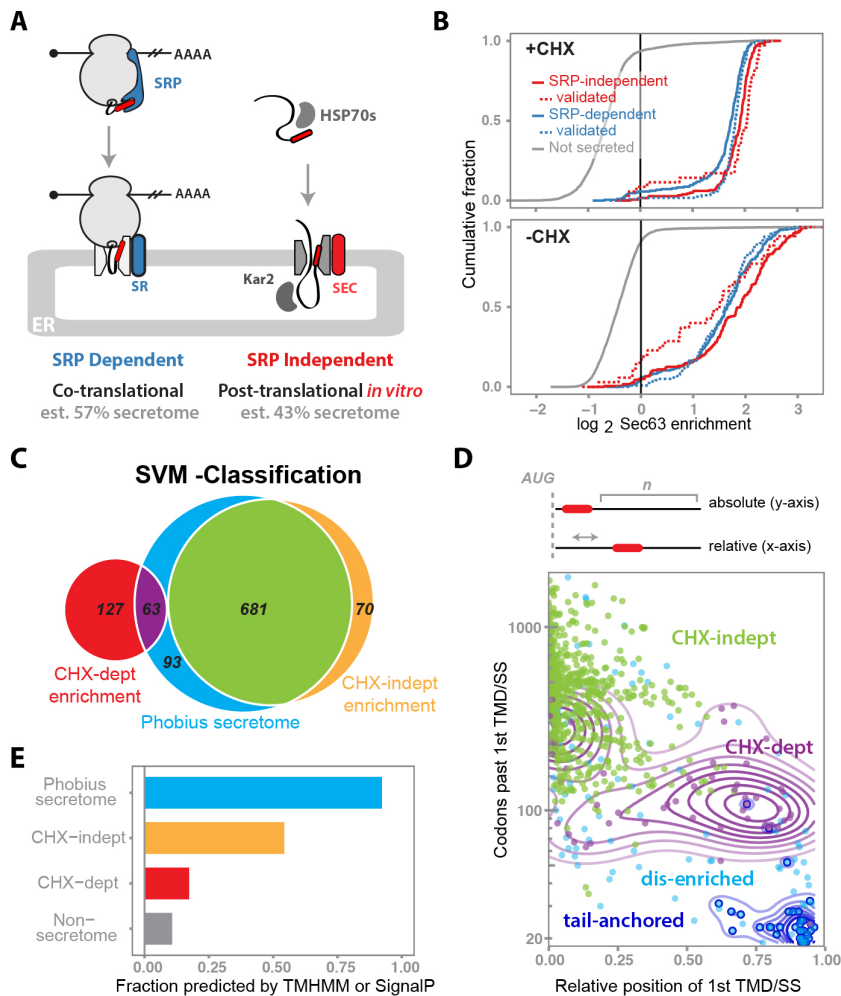


Fig. 3-3. Global characterization of co- vs post-translational translocation in vivo. (A) Overview of current models for SRP-dependent and –independent targeting to and translocation into the ER. Predictions for the proportion of substrates that partition between pathways are taken from Ast et al., 2013. (B) Cumulative distribution of the Sec63-BirA log₂ enrichments for SRP-dependent (blue), –independent (red), and non-secreted (grey) genes with or without CHX. Biochemically validated genes (dashed lines) were consolidated from Ast et al., 2013 and Ng et al., 1996. (C) Venn diagram summarizing the SVM classifications for CHX-dependence in the context of the Phobius-predicted secretome. The Sec63-BirA +CHX enrichment profile was fit as a mixture of two normal distributions and all genes enriched above the 99th percentile of the dis-enriched distribution were classified by the SVM. (D) Number of codons downstream of the first hydrophobic domain of Phobius-secretome genes versus the position of this domain relative to overall gene length, plotted for genes in different SVM-classified enrichment categories. Contour lines are added for specific gene sets for visual clarity and represent Gaussian density fits of the corresponding points in that set. Colors match those in (C) with the tail-anchored genes (dark-blue) overlaid as open circles. (E) Proportion of genes for which a hydrophobic feature was predicted by either TMHMM or SignalP, for different gene categories. Colors and gene sets match those in (C) with the addition of non-secretome genes (grey) as predicted by Phobius.

Figure 3-4

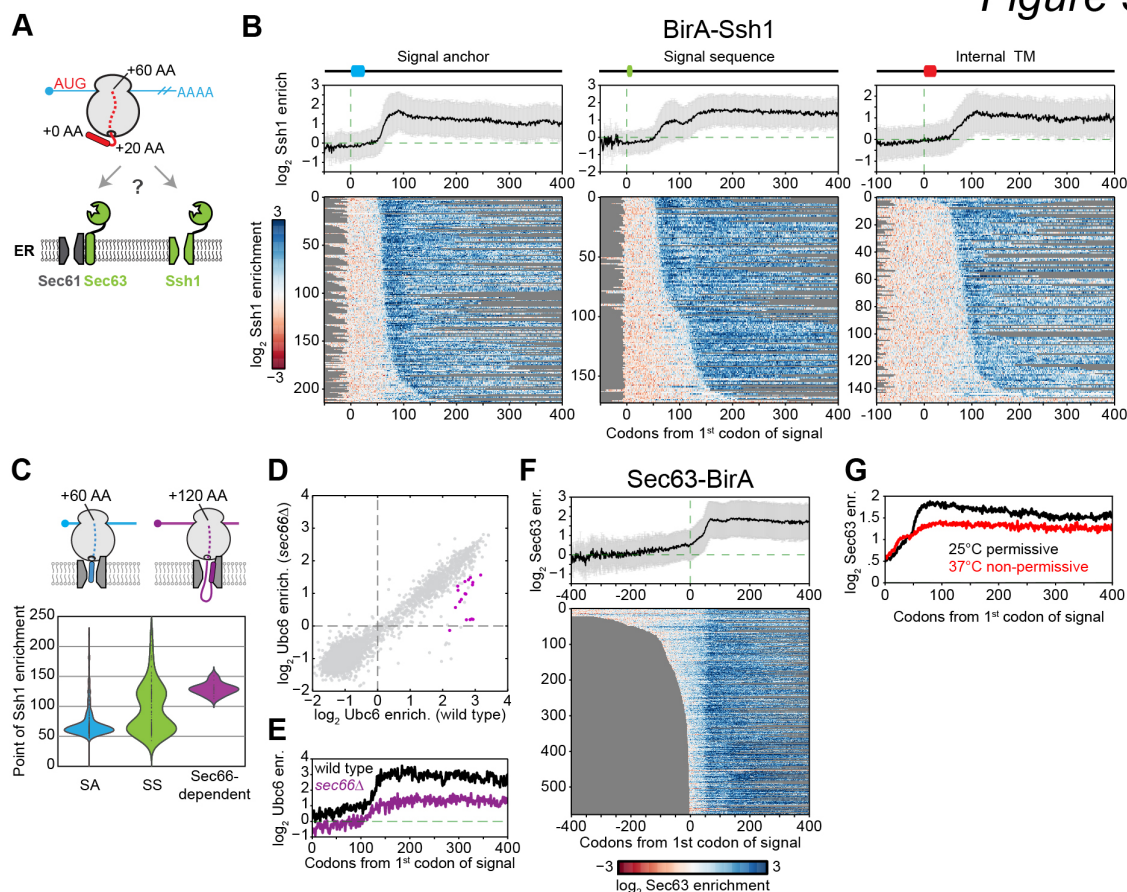


Fig.3-4. Timing and specificity of co-translational targeting to the ER. (A) Schematic of the yeast translocon-specific BirAs used to examine ribosome accessibility at two translational entry points into the ER. (B) Meta-gene plots of log₂ BirA-Ssh1 enrichment per codon as a function of ribosome position relative to the first codon of the first Phobius-predicted hydrophobic element for the indicated signal class. Heat maps below represent single-gene positional enrichments used to derive the corresponding averaged meta-gene plot, sorted by increasing distance to the point at which enrichment occurs. (C) Violin plot showing the distribution of the point of enrichment for BirA-Ssh1 relative to the first hydrophobic element, for different types of hydrophobic features and Sec66-dependent genes as defined in (D). Shown above are two RNC conformations consistent with nascent chain lengths. (D) Gene enrichments obtained with the general BirA-Ubc6 ER marker in yeast in wild type vs *sec66Δ* backgrounds. Sec66-dependent genes are defined in (Fig. 3-S4). (E) Meta-gene plot as in (B) of log₂ BirA-Ubc6 enrichments for Sec66-dependent genes in wild type (black) and *sec66Δ* (purple) backgrounds. (F) Meta-gene plot as in (B) of log₂ Sec63-BirA enrichments. (G) Meta-gene plot of log₂ enrichments as in (F) in a *sec65-1* SRP ts background at the permissive (25°C, black) and non-permissive (37°C, red) temperatures

Figure 3-5

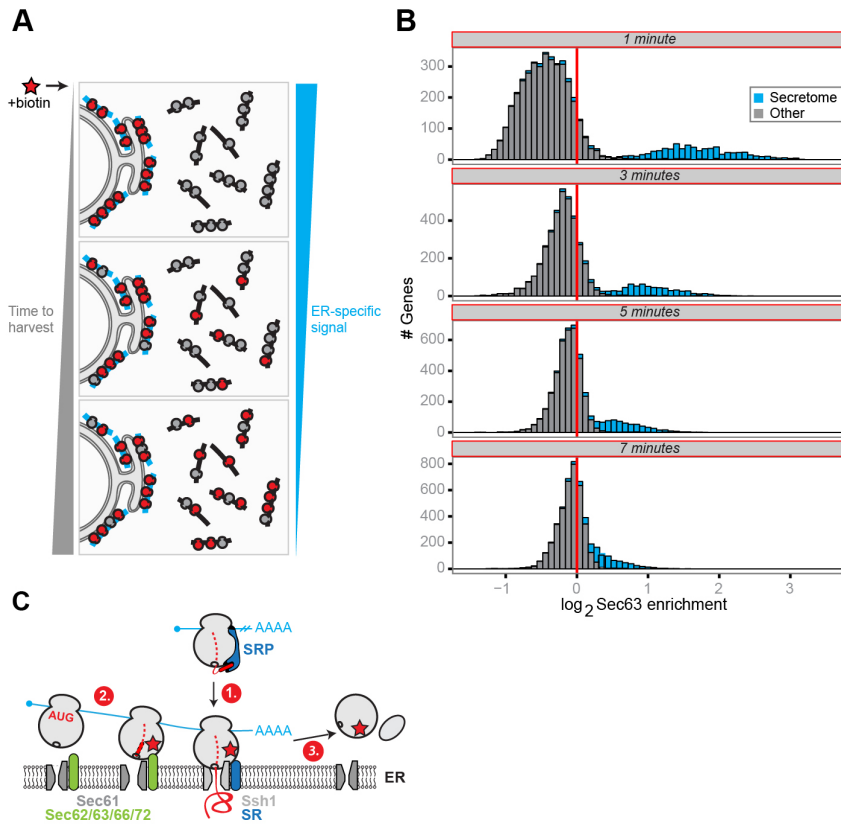


Fig. 3-5. Dynamics of ER-associated ribosomes in vivo. (A) Overview of the pulse labeling experiment to assay the kinetics of ribosome exchange from the ER in vivo. (B) Histograms of log₂ Sec63-BirA enrichment values for well-expressed secretome (blue) and all other (grey) genes over the exchange time course. Times represent the total time of ribosome biotinylation in the absence of CHX. (C) Working model consistent with the positional enrichments observed for the translocon-specific BirAs and ribosome recycling. (1) Initial recruitment to the ER depends on a fully accessible signal sequence. (2) Ribosomes translating ER-tethered mRNAs can interact with SEC early. (3) Upon termination, ribosomes recycle into the cytosolic pool.

MATERIALS AND METHODS

Except where explicitly noted, methods for yeast match those used for the equivalent HEK293T experiment or analysis. A complete list of yeast strains, cell lines, plasmids, and primers can be found in Tables 3-1 to 3-4, respectively.

1 Experimental

1.1 Strain construction

The endogenous copies of *RPL16a*, *RPL16b*, and *RPS2* were C-terminally tagged with an engineered HA-TEV-AviTag sequence to allow for detection by western blot, biotinylation, and specific elution after streptavidin pulldown via TEV protease cleavage. The 3' UTRs of these proteins were preserved by scarless tagging using the pop-in/pop-out method described in (Rothstein, 1991) with the plasmids p1, p2, and p3. Yeast strain y2 in which both *RPL16* paralogs were tagged was obtained by crossing and sporulation of the two individually tagged paralog strains.

Cytosolic BirA was integrated genomically at the *LEU2* locus and was driven by a PGK1 promoter (p4). The endogenous copies of *SEC63* and *OM45* were C-terminally tagged with -mVenus-BirA::HIS5 using p5. *SEC63* is an essential gene and the functionality of the tagged copy was inferred from a lack of growth defect. The *SSH1* gene was cloned to create BirA-mVenus-Ssh1::HIS5 (p6) and BirA-Heh2-Ssh1::HIS5 (p7) cassettes, which were used to replace the genomic copy for expression under the control of the endogenous promoter. Heh2 contains a 120 residue, intrinsically disordered linker (Meinema et al., 2011) that was reasoned to be less prone to proteolytic cleavage than a (GS)_n linker of an equivalent length. This long linker was used in the context of Ssh1 to confirm that the positional enrichments observed with the mVenus linker (Fig. 3-4B) were not dependent on linker length (Fig. 3-S3A). Inhibition of translocation is known to induce a cytosolic stress response, thus an Hsf1

reporter (Brandman et al., 2012) was used to evaluate the functionality of tagged Ssh1 (Fig. 3-S3B). To match the expression level of other ER-localized BirAs, BirA-mVenus-ubc6 was expressed under control of the SEC63 promoter from the *HIS3* locus (p8).

Deletion of *SEC66* was performed by one-step gene replacement as described in (Longtine et al., 1998). The temperature sensitive allele of *SEC65* was introduced with a *sec65-1::KAN* cassette (a kind gift from Peter Walter).

1.2 Cell line construction

HEK293T cell lines were derived by sequential lentiviral infection. Tagged ribosomal proteins were expressed from the EF1a promoter as Puro-T2a-AviTag-TEV-HA-mmuRPL10A (60S) or Puro-T2A-hsRPS2-HA-TEV-AviTag (40S). BirA was expressed from a Dox-inducible promoter on the pINDUCER11 backbone (Meerbrey et al., 2011). Cytosolic BirA was expressed as an N-terminal 3X FLAG tag, whereas ER-localized BirA was expressed as a 3X FLAG-mCherry-sec61 β fusion. pINDUCER11-infected cells were FACS sorted to generate stable polyclonal cell lines.

1.3 Microscopy

The localization of BirA fusion proteins was assessed by live cell microscopy on a Spinning-Disc confocal. Black 96 well glass (#1.5) bottom plates were treated with 0.1 mg/mL Concanavalin A for 10 minutes. The ConA was aspirated and plates allowed to dry. 200 μ L of yeast cells grown to OD < 0.2 in SD media were applied to each well and allowed to adhere for approximately 15 minutes before imaging. Mammalian cells were cultured on glass (#1.5) bottom 35 mm dishes.

Plates were imaged on a 100X Nikon Apo TIRF 1.49 oil objective. Images were acquired from a Photometrics Evolve EMCCD attached to a Nikon Ti-E microscope equipped

with a Yokogawa CSU-22 spinning disc in micromanager. mVenus was excited by a 491 nm Cobalt Calypso DPSS laser line, emission measured at 525 nm. mCherry fluorescence was excited by a 561 nm Coherent Sapphire DPSS laser line, emission measured at 610 nm.

1.4 Media and growth conditions

a. Yeast

To prevent constitutive biotinylation, yeast were grown in biotin-free, synthetic defined media (1.7 g/L YNB-Biotin [Sunrise Science Products], 5 g/L Ammonium sulfate, 20 g/L dextrose, complete amino acids) supplemented with d-biotin (Sigma) to a final concentration of 0.125 ng/mL, at 30°C with vigorous shaking. Under these conditions, steady-state ribosome biotinylation was undetectable and we observed minimal impact on log-phase doubling time, although the OD₆₀₀ at which saturation occurred was lowered to roughly 2-3 (Fig. 2-2). In a typical sequencing experiment, 25-50 mL of an overnight culture was used to inoculate a 300-750 mL culture at an OD₆₀₀ of 0.05-0.1, and biotin induction was performed at mid-log phase with an OD₆₀₀ of 0.6-0.8.

Growth of the yeast strain y12 expressing the *sec65-1* temperature-sensitive allele of *SEC65* (Stirling et al., 1992) was carried out as above at the permissive temperature of 25°C. To test the effect of acute loss of SRP, cultures were shifted to the non-permissive temperature of 37°C for 30 minutes prior to biotin induction, using a water bath to ensure a rapid temperature transition.

b. HEK293T

Fetal bovine serum was depleted of biotin with strep-tactin sepharose (IBA). Four mL of strep-tactin slurry were washed thrice in 20 mL PBS then incubated with 150 mL FBS at room temperature for 2.5 hours. Strep-tactin beads were pelleted at 300 x g for four minutes

and the supernatant filtered. Cells were cultured in DMEM with 10% biotin-depleted FBS, 2 mM L-Glutamine and 100 ng/ μ L doxycycline (Sigma).

1.5 Biotin induction and harvesting

Biotin induction was carried out at mid-log phase either in the presence or absence of cycloheximide (CHX), as required by the experiment. If used, CHX was added to media 2 minutes prior to the addition of biotin, at a final concentration of 100 μ g/mL. To induce biotinylation, d-biotin was added to the media to a final concentration of 10 nM (50 μ M for HEK293T cells) and biotinylation was allowed to proceed for the desired amount of time at the same temperature as growth. We note that this induction concentration maintains an appropriate balance between favorable biotinylation kinetics (Fig. 2-2B) and effective quenching of the post-lysis biotinylation reaction via de-salting (Fig. 2-3C).

For large-scale sequencing experiments, cells were harvested by filtration onto 0.45 μ m pore size nitrocellulose filters (Whatman), scraped from the membrane, and immediately submerged in ℓ N₂. All times specified in the text or methods represent the total time from biotin induction to ℓ N₂.

For small-scale western-blot experiments, 1 mL aliquots were taken from 5-10 mL induced cultures and placed into pre-chilled, 1.5 mL siliconized microcentrifuge tubes. Samples were then spun at 20,000 x g at 4°C for 30 seconds, the supernatant removed, and the pellet-containing tubes immediately placed in ℓ N₂. We note that the use of siliconized tubes was essential for efficient yeast pelleting in synthetic defined media.

1.6 Western blotting and biotinylation quantification

Lysates were prepared from pelleted yeast (see 1.5) by resuspending frozen pellets in 30-50 μ L Laemmli buffer, followed by denaturation at 70°C for 10 minutes, and clarification at room temperature by spinning at 20,000 x g for 10 minutes. Lysates from monosome

aliquots were prepared using the appropriate dilution with 4x NuPage LDS buffer (Novex), supplemented with 20 mM fresh DTT, followed by denaturation for 10 minutes at 70°C, or 99°C if preparing from streptavidin beads (e.g., for pulldown aliquots).

Adherent mammalian cells were rinsed in ice cold PBS and lysed directly in plates in polysome lysis buffer (20 mM Tris pH 7.5, 150 mM NaCl, 5 MgCl₂, 2% Triton X-100, 1 mM DTT) on ice for 5 minutes and clarified by spinning at 20,000 x g for 5 minutes. Lysates were then denatured in NuPAGE LDS supplemented with 20 mM DTT at 70°C for 10 minutes.

Lysates were run on 4-12% Bis-tris gels, transferred to PDVF membranes using the XCell II system (Invitrogen) according to the manufacturer's instructions, blocked with Odyssey blocking buffer, and subsequently probed. The HA epitope tag was detected using a mouse anti-HA antibody at a 1:5,000 dilution (Roche 12CA5) or the high-affinity rat anti-HA antibody at a 1:1,000 dilution (Roche 3F10). Licor IRDye800 anti-mouse (Odyssey), or IR800 anti-rat (Rockland) secondary antibodies were then used at 1:20,000 or 1:5,000 dilutions, respectively. Biotin was detected directly (Fig. 3-1D, Fig. 2-3) using Streptavidin AlexaFluor 680 (Molecular Probes) at a 1:5,000 dilution in TBST and a 10 minute incubation period. This was carried out either directly after blocking or probing with secondary antibody. All blots were visualized using the Licor (Odyssey) system.

Percent biotinylation was quantified by probing for HA in a streptavidin shift assay, in which clarified lysates were mixed with excess unlabeled streptavidin (Rockland) prior to electrophoresis and immunoblotting. Biotinylated AviTags shift to a higher molecular weight than the corresponding, non-biotinylated AviTags, and percent biotinylation was computed from the fraction of total signal (shifted + unshifted) that was shifted.

1.7 Monosome isolation and lysate desalting

a. Yeast

In a typical experiment, 650 μL of polysome lysis buffer (20 mM Tris pH 8.0, 140 mM KCl, 1.5 mM MgCl_2 , 100 $\mu\text{g}/\text{mL}$ CHX, 1% Triton X-100) was dripped into a 50 mL conical tube filled with and immersed in ℓN_2 , containing the harvested yeast strip/pellet from a mid-log phase 300 mL biotin-induced culture (see 1.5). The frozen cell, buffer mixture was cryogenically pulverized for six, three minute x 15 Hz cycles on a Retsch MM301 mixer mill. Sample chambers were pre-chilled in ℓN_2 and re-chilled between each cycle.

Upon thawing, a substantial amount of non-specific post-lysis biotinylation was observed at temperatures above 4°C. Pulverized cells were thus thawed and kept as close to 0-4°C as possible prior to a very brief clarification step performed by spinning for 2 minutes at 4°C and 20,000 x g on a tabletop centrifuge. The supernatant was then immediately loaded onto pre-chilled, 2 mL Zeba de-salt spin column previously equilibrated with polysome gradient buffer (20 mM Tris pH 8.0, 140 mM KCl, 5 mM MgCl_2 , 100 $\mu\text{g}/\text{mL}$ CHX, 0.5 mM DTT) according to the manufacturer's instructions. This step depletes the lysates of biotin and ATP, effectively quenching the post-lysis biotinylation reaction (Fig. 3-1D, 2-2C). Aliquots of this extract were flash-frozen in ℓN_2 , typical yields were 0.5-1 mL of extract with A_{260} of 100-300.

A 200 μL aliquot of the above lysate was treated with 7.5 U RNaseI (Ambion) per 50 A_{260} units of lysate, and incubated for 1 hour at room temperature on an overhead roller. Reactions were then quenched with 10 μL SUPERase-In RNase inhibitor (Ambion) and loaded onto sucrose density gradients (10-50% w/v) prepared with the polysome gradient buffer described above. Gradients were made in Sw-41 ultracentrifuge tubes (Seton Scientific) using a BioComp Gradient Master (BioComp Instruments) according to the manufacturer's instructions. Samples were spun for 3 hours at 4°C and 35,000 rpm in an Sw-41 rotor (Beckmann Coulter). Fractionation was performed on the Gradient Master using a BioRad EM-1 Econo UV monitor to continually monitored A_{260} values. Monosome peaks

were collected, flash-frozen in ℓN_2 , and stored at -80°C . Typical yields were 2-3 mL of monosomes with A_{260} of 2-5.

b. HEK293T

Cells were rinsed with ice-cold PBS containing 100 $\mu\text{g}/\text{mL}$ cycloheximide, then lysed directly in 15 cm plates by addition of 800 μL polysome lysis buffer supplemented with 100 $\mu\text{g}/\text{mL}$ CHX, 1U/mL apyrase (NEB), 24U/mL Turbo DNase (Ambion). Cells were scraped and collected into a 1.5 mL tube and kept on ice for ~ 10 minutes. Lysates were cleared by spinning for 5 minutes at 4°C and 20,000 $\times g$ on a tabletop centrifuge. The supernatant was then immediately loaded onto pre-chilled, 2 mL Zeba de-salt spin column previously equilibrated with polysome gradient buffer (20 mM Tris pH 7.5, 150 mM NaCl, 5 mM MgCl_2 , 100 $\mu\text{g}/\text{mL}$ CHX, 1 mM DTT) according to the manufacturer's instructions.

1.8 Streptavidin pulldown of biotinylated ribosomes

Biotinylated ribosomes were isolated from the total monosome fraction (see 1.7) using MyOne streptavidin C1 magnetic DynaBeads (Invitrogen). We note that the use of magnetic beads resulted in significantly less non-specific binding as compared to the agarose-bead-based equivalent (Fig. 2-4). The volume of beads used per pulldown was scaled based on 187 μL (1.87 mg) beads per 15 pmol of biotinylated ribosomes, as estimated from the manufacturer's instructions. The pmol of biotinylated ribosomes in a given volume was calculated from (i) the fraction of biotinylated ribosomes as estimated from a streptavidin shift assay (see 1.6) and (ii) the total concentration of 80S ribosomes in the fraction, determined by the A_{260} and using an extinction coefficient of $5 \times 10^7 \text{ cm}^{-1}\text{M}^{-1}$ (Algire et al., 2002).

Prior to binding, beads were washed 2x with one volume (equal to the initial bead volume) of Buffer A (100 mM NaOH, 50 mM NaCl), 1x with one volume of Buffer B (100 mM NaCl), and 1x with one volume of low-salt binding Buffer C (20 mM Tris pH 8.0, 140 mM KCl, 5 mM MgCl₂, 100 µg/mL CHX, 0.5 mM DTT, 0.1% Triton X-100). Triton X-100 was added to monosome fractions containing 15 pmol of biotinylated ribosomes, to a final concentration of 0.01%. This solution was added to washed beads and the pulldown was allowed to proceed on an overhead roller for 1 hr at 4°C. The supernatant was removed and the beads were washed 3x with 1 mL high-salt wash Buffer D (20 mM Tris pH 8.0, 500 mM KCl, 5 mM MgCl₂, 100 µg/mL CHX, 0.5 mM DTT, 0.1% Triton X-100), each for 20 minutes at 4°C. After the third wash, beads were re-equilibrated in low-salt Buffer C by resuspension in 1 mL, then transferred to a new tube and resuspended in a smaller volume (200 µL) of Buffer C in preparation for elution by TEV protease cleavage. Cleavage was performed by incubation on a nutator with in-house TEV protease for 1 hr at room temperature. Three volumes of Trizol LS (Ambion) were added to both the TEV eluate and a separate, matched input sample consisting of 10-20 pmol of total monosomes.

1.9 Library generation

10-20 pmol of monosomes in Trizol LS were extracted using 200 µL chloroform per 750 µL Trizol LS. RNA was precipitated for at least 1 hour at -30°C using GlycoBlue (Invitrogen) and an equal volume of isopropanol, pelleted, resuspended in 10 µL nuclease-free water, and resolved on a 10 or 15% TBE-urea gel. Oligoribonucleotide size standards in neighboring lanes were used to excise roughly 24-32 nt ribosome footprints, which were passively eluted overnight at 4°C in 200-400 µL 0.3 M NaCl.

Ribosome footprints were then precipitated with GlycoBlue and 2.5 volumes ethanol, resuspended directly in 8 µL 1.25x T4 polynucleotide kinase (PNK) buffer (New England

Biolabs [NEB]), and dephosphorylated with 2 μ L PNK for 1 hour at 37°C. This solution was used directly for ligation to 0.5 μ g 3' miRNA cloning linker 1 (Integrated DNA Technologies) upon addition of 8 μ L 50% PEG (NEB), 1 μ L 10x truncated T4 RNA ligase 2 K227Q (rnl2) buffer (NEB), and in-house rnl2 enzyme. Ligation proceeded for 3 hours at 25°C at which point RNA was precipitated for at least 1 hour at -30°C, purified on a 10% TBE-urea gel, eluted, and precipitated as above.

rRNA contaminants were removed from ligation products as described (Brar et al., 2012) for yeast or using the Ribo-Zero kit (Epicentre) for mammalian cells. rRNA-depleted ligation products were then reverse-transcribed in a 16.7 μ L reaction using SuperScript III (Invitrogen) for 30 minutes at 48°C. RNA template was hydrolyzed for 20 minutes at 98°C after addition of 1/10 volume 1 M NaOH. Equi-molar HCl was added to quench the reaction and cDNAs were precipitated at -30°C for at least 1 hour and subsequently purified on a 10% TBE-urea gel, eluted overnight, precipitated, and resuspended in 15 μ L nuclease-free water.

cDNAs were circularized using CirLigase (Epicentre) in a 20 μ L reaction for 1.5 hours at 60°C according to the manufacturer's instructions. Circularized products were amplified by 8-16 cycles of PCR using oNTI231 and any of several Illumina indexing primers (IDT) using Phusion polymerase (Finnzymes) in a 17 μ L reaction. PCR amplicons were gel purified on 8% non-denaturing TBE gels, eluted, precipitated, resuspended in 10 μ L EB, and quantified using the Bioanalyzer High Sensitivity DNA assay (Agilent Technologies). 2 nM dilutions were multiplexed as needed and sequenced via a single-end run on an Illumina HiSeq sequencer, using version 3 clustering and sequencing kits with a 6-cycle index read (Illumina).

2 Computational analyses

2.1 Footprint sequence alignment

Sequencing reads were demultiplexed and stripped of 3' cloning adapters using in-house scripts. Reads shorter than 18 nt long were discarded and the remainder were aligned using Bowtie 0.12.9 (Langmead et al., 2009). For all yeast experiments the UCSC sacCer3 genome and annotations (April 2011) were used. Reads were mapped sequentially to Bowtie indices composed of rRNAs, tRNAs, and finally all chromosomes. Only uniquely-mapped, zero-mismatch reads from the final genomic alignment were used for subsequent analyses. These alignments were assigned to a specific P-site nucleotide, scaled according to length based on a 15-nt offset from the 3' end of reads.

2.2 Gene enrichments

Gene-level enrichments were computed by taking the \log_2 ratio of biotinylated footprint density (reads per million) within a gene coding sequence (CDS) over the corresponding density in the matched input ribosome profiling experiment. Yeast genes were excluded from all analyses if they met any of the following criteria: had fewer than 100 CDS-mapping footprints in the input sample of a particular experiment; annotated as 'dubious' in the SGD database; CDS overlaps another same-strand CDS; gene maps to the mitochondrial chromosome. Additionally, with the exception of Fig. 3-3, genes classified as being CHX-dependent for co-translational ER enrichment were excluded from all analyses.

2.3 Gene ontology analysis

The threshold to determine the enriched gene set was determined by ROC analysis. The secretome (yeast) or set of Phobius-predicted proteins (HEK293T) were considered true positives. The true negative set included all genes that were not in the secretome and not predicted by Phobius (Käll et al., 2004), SignalP (Petersen et al., 2011) or TMHMM (Krogh et al., 2001).

GO term enrichment was performed on the set of genes above the ROC-defined threshold compared to a background set of all expressed genes with an RPKM \geq 5. For the HEK293T dataset, GO term enrichment was determined for the Phobius-predicted genes below threshold versus the background set. Enrichments were calculated using the GeneCodis webtool (Tabas-Madrid et al., 2012).

2.4 Gene categorization

a. Peroxisomal annotations

To generate a comprehensive list of yeast peroxisomal genes, SGD was queried for *S. cerevisiae* genes annotated with the GO component name 'peroxisome', or any term below this on the GO hierarchy. Transmembrane domain annotations for these genes were consolidated from Phobius predictions (see 2.8) and the UniProtKB database.

b. SRP-dependence

Genes predicted as SRP-dependent or -independent were taken from (Ast et al., 2013). Genes biochemically validated as falling in either category were consolidated from (Ast et al., 2013) and (Ng et al., 1996).

c. Mitochondrial annotations

Mitochondrial annotations were taken from the mitop2 database (Eltner et al., 2009) for yeast and from the MitoCarta database (Pagliarini et al., 2008) for human.

2.5 CHX-dependence classifier

Due to the kinetics of biotinylation, co-translational ER enrichment in the absence of CHX increases along the length of a gene, confounding a simple fold-change-based approach for classifying CHX-dependence. An SVM classifier was therefore implemented to systematically

and objectively characterize this phenomena, using the R package 'caret' (version 5.17-7). The bimodal distribution of \log_2 gene enrichments from a Sec63-BirA, 7 minute +CHX biotinylation experiment were fit to two normal distributions, and all genes with enrichments greater than the 99th percentile of the dis-enriched population (\log_2 enrichment of 0.25) were classified as CHX-dependent or -independent.

Eight features were used for classification: \log_2 gene enrichments for one minute -CHX, two minute +CHX, and seven minute +CHX libraries; gene enrichment ranks (computed in the context of all genes, not only those classified) for one minute -CHX, two minute +CHX, and seven minute +CHX libraries; gene length; the slope of enrichment across a gene CDS obtained from a linear model fit (generally positive for CHX-independent enrichment). Features were chosen based on their predictive ability as determined with the recursive feature elimination utility in the caret package.

Our combined training and test sets consisted of 62 CHX-dependent and 74 CHX-independent genes which were determined empirically as falling into either category by name-blind inspection of the same features used to train the classifier. The final SVM was trained on 75% of the empirically-classified genes, with three rounds of 10-fold cross-validation using scaled-feature values and a linear kernel function, which outperformed the Gaussian equivalent. Testing on the final 25% of genes yielded zero classification error. Of 941 enriched genes, 190 were classified as CHX-dependent (Fig. 3-3C). A full list of manually-curated and SVM-classified genes can be found in Table S5 of the original publication.

2.6 Positional enrichments

Secretome genes were classified based on the type and location of their first hydrophobic feature predicted by Phobius as follows: cleavable signal-sequence, signal anchor (first TMD

beginning within the first 60 codons), or internal transmembrane (first TMD beginning after the first 60 codons). Genes were required to have an average of ≥ 3 footprints per codon. For each gene, the \log_2 enrichment at each codon was calculated after addition of a 0.1 RPM pseudocount. For metagene analyses, genes were aligned by the first codon of the first hydrophobic feature and the geometric mean of \log_2 enrichments was calculated at each codon.

The point of enrichment for each gene was calculated by sliding a 21 codon window from the N- to C-terminus and maximizing the difference between the C-terminal window and the N-terminal window. The N-terminus was padded with ten codons of 0 \log_2 enrichment to allow for genes that begin enriched. Heatmaps display genes sorted by their point of enrichment.

2.7 Sec66-dependent genes

Genes that were dependent on the non-essential SEC complex component Sec66 to maintain robust co-translational ER enrichment were identified by comparing the BirA-Ubc6 ER enrichments obtained with a two minute biotin pulse in the presence of CHX, in a wild type strain (y8) versus a *sec66* Δ strain (y13). \log_2 enrichments were separately normalized by subtracting the mean enrichment and dividing by the standard deviation of enrichments for the corresponding experiment. Genes were then binned by the minimum number of sequencing counts in either the wild type or *sec66* Δ input sample, and the difference between the normalized enrichments in the two samples was compared within each bin. Negative outliers were defined as those genes whose enrichments decreased in the *sec66* Δ sample by at least three standard deviations compared to the other genes in that bin (Fig. 3-S4). A full list of Sec66-dependent genes can be found in Table S5 of the original publication.

2.8 Hydrophobic domain predictions

Except where explicitly noted, all hydrophobic domain annotations were based on Phobius (Käll et al., 2004) predictions using the default settings. Alternative predictions of transmembrane domains and signal peptides (Fig. 3-3D) were based on the output of TMHMM 2.0 (Krogh et al., 2001) and SignalP 4.0 (Petersen et al., 2011) with the default settings.

SUPPLEMENTAL FIGURES

Figure 3-S1

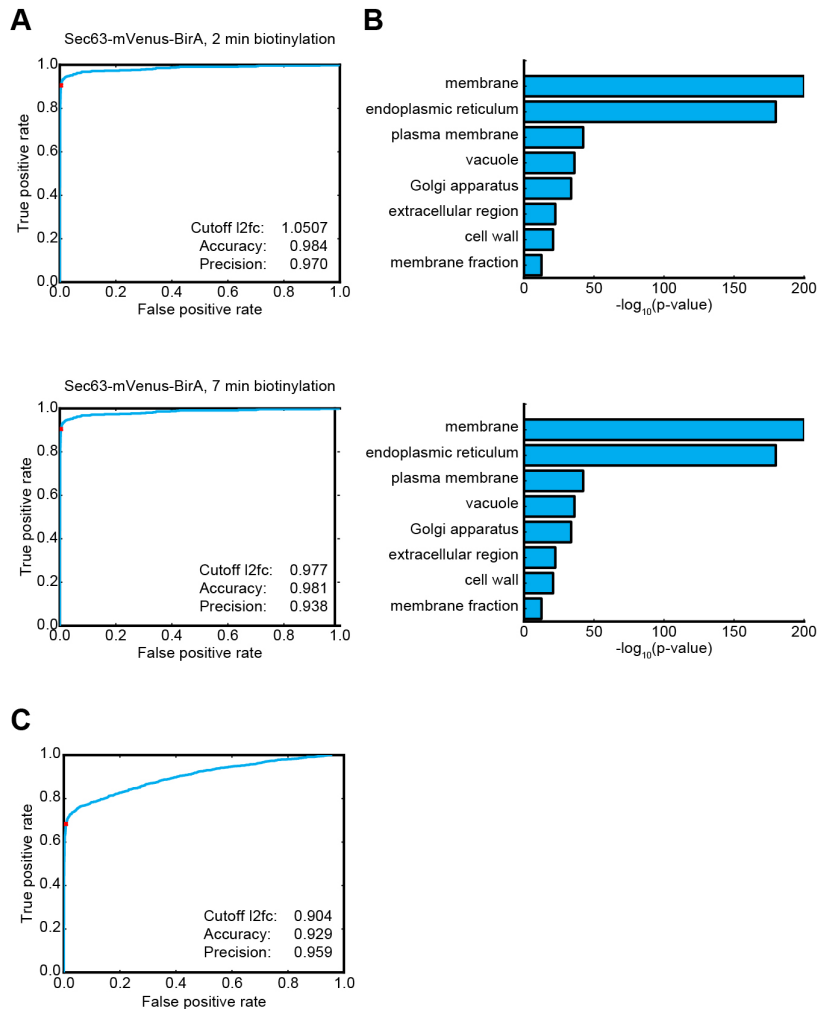


Fig. 3-S1. ROC for enrichment thresholds. (A) ROC curves were used to evaluate the ability of enrichment thresholds to classify genes translated at the ER. The secretome was used as the true positive set and true negatives were defined as the set of genes that were not in the secretome and were furthermore not predicted by SignalP or TMHMM. The threshold that maximized accuracy was used to classify genes as enriched for downstream analyses. Above are data from cells expressing Sec63-BirA biotinylated for two minutes; the below data are from cells treated with biotin for seven minutes. (B) Genes enriched at or above threshold were compared to the background set of all expressed genes in a cellular-component GO-term analysis, ordered as in (A). (C) Analysis as in (A) for enrichment in 293T cells expressing BirA-Sec61 β treated with biotin for ten minutes. The true positives were defined as all genes predicted by Phobius, and true negatives defined as all genes that were not predicted by Phobius, SignalP or TMHMM.

Figure 3-S2

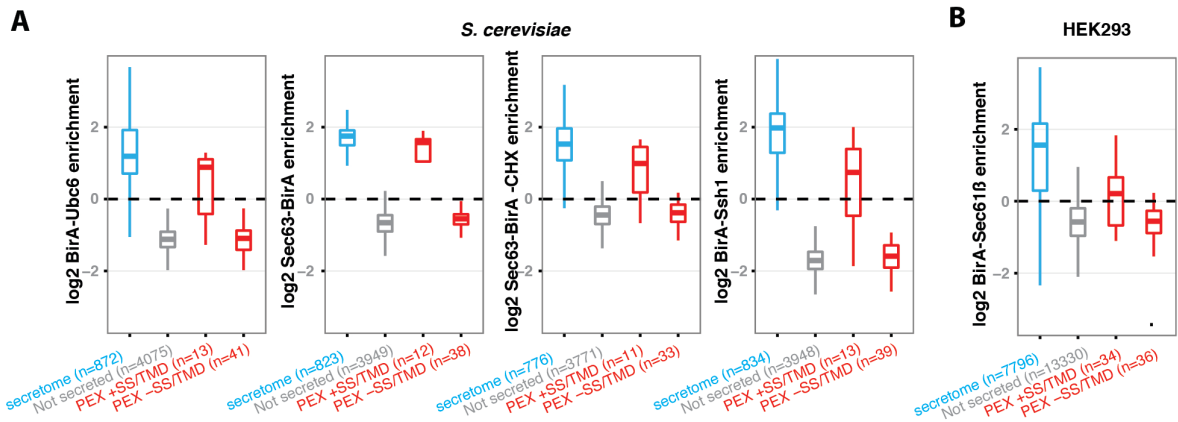


Fig. 3-S2. TMD-containing peroxisomal proteins are synthesized at the ER The distributions of ER enrichments for genes in the indicated groups from ER-localized BirA fusions (see y-axis) in yeast (**A**) and mammalian cells (**B**). Cells were treated with CHX unless otherwise indicated.

Figure 3-S3

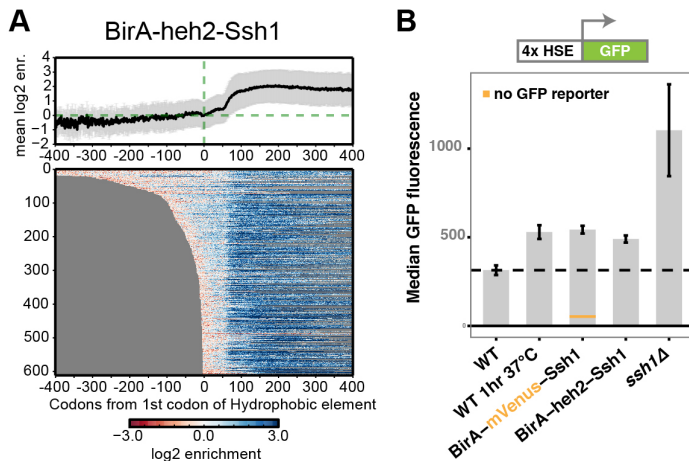


Fig. 3-S3. Functionality of Ssh1 BirA fusion proteins (**A**) Meta-gene plot of log₂ BirA-heh2linker-Ssh1 enrichments as a function of ribosome position in the context of the long, intrinsically disordered linker from Heh2p. The observed pattern closely matches that obtained with a mVenus linker (Fig. 3-4B), suggesting that it reflects properties intrinsic to the Ssh1 translocon as opposed to steric restriction. (**B**) An Hsf1 reporter was used to assay cytosolic stress in order to assess the functionality of Ssh1 fusion proteins. Error bars represent standard error of two replicates. Background fluorescence from the spectral overlap of mVenus is marked in orange and contributed minimally to the measured GFP reporter levels.

Figure 3-S4

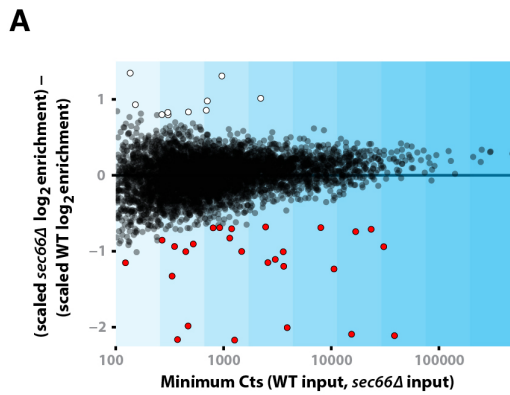


Fig. 3-S4. Determination of Sec66-dependence by fold change (A) Plotted are the differences in scaled (to zero-mean and unit variance) log₂ ER enrichment for *sec66Δ* libraries minus wildtype for each gene as a function minimum expression level of the input counts between wildtype and *sec66Δ* libraries. Genes were divided into nine logarithmically spaced bins (shaded blue vertical bars) and genes beyond three standard deviations from the mean in each bin were considered significantly different. Genes that increase in their ER-association are plotted as white circles, genes that decreased are plotted as red circles.

Figure 3-S5

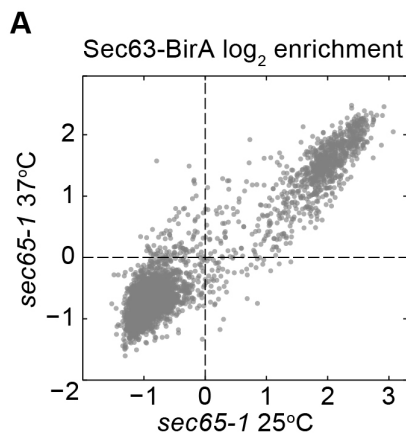


Fig. 3-S5. Gene-level translational enrichments upon acute loss of SRP (A) Scatter plot of ER enrichments in *sec65-1* cells under permissive or restrictive conditions.

Supplemental tables

Table 3-1. Yeast strains

name	description	parent	MAT	genotype
y1	Rps2-HA-TEV-AVI	BY4741	a	<i>rps2::RPS2-HA-TEV-AVI leu2Δ0 ura3Δ0 met15Δ0 his3Δ1</i>
y1.1	Rps2-HA-TEV-AVI	BY4742	α	<i>rps2::RPS2-HA-TEV-AVI ura3Δ0 lys2Δ0 his3Δ1 leu2Δ0</i>
y2	Rpl16a-HA-TEV-AVI; Rl16b-HA-TEV-AVI	sporulate d	a	<i>rpl16a::RPL16a-HA-TEV-AVI rpl16b::RPL16b-HA-TEV-AVI ura3Δ0 met15Δ0 his3Δ1 leu2Δ0</i>
y3	pPGK-BirA (cytosolic)	BY4741	a	<i>leu2::pPGK1-BirA::LEU2 ura3Δ0 met15Δ0 his3Δ1</i>
y4	pPGK-BirA (cytosolic); Rps2-HA-TEV-AVI	y1	a	<i>rps2::RPS2-HA-TEV-AVI leu2::pPGK1- BirA::LEU2 ura3Δ0 met15Δ0 his3Δ1</i>
y5	pPGK-BirA (cytosolic); Rpl16a/b-HA-TEV-AVI	y2	a	<i>rpl16a::RPL16a-HA-TEV-AVI rpl16b::RPL16b- HA-TEV-AVI leu2::pPGK1-BirA::LEU2 ura3Δ0 met15Δ0 his3Δ1</i>
y6	Sec63-mVenus-BirA; Rps2-HA-TEV-AVI	y1.1	α	<i>sec63::SEC63-mVenus-BirA::HIS5 rps2::RPS2- HA-TEV-AVI ura3Δ0 lys2Δ0 his3Δ1 leu2Δ0</i>
y7	Sec63-mVenus-BirA; Rpl16a/b-HA-TEV-AVI	y2	a	<i>sec63::SEC63-mVenus-BirA::HIS5 rpl16a::RPL16a-HA-TEV-AVI rpl16b::RPL16b- HA-TEV-AVI ura3Δ0 met15Δ0 his3Δ1</i>
y8	BirA-mVenus-Ubc6; Rpl16a/b-HA-TEV-AVI	y2	a	<i>his3::pSEC63-BirA-mVenus-UBC6 TA::HIS5 rpl16a::RPL16a-HA-TEV-AVI rpl16b::RPL16b- HA-TEV-AVI ura3Δ0 met15Δ0 leu2Δ0</i>
y9	BirA-mVenus-Ssh1; Rpl16a/b-HA-TEV-AVI	y2	a	<i>ssh1::BirA-mVenus-SSH1::HIS5 rpl16a::RPL16a-HA-TEV-AVI rpl16b::RPL16b- HA-TEV-AVI ura3Δ0 met15Δ0 his3Δ1</i>
y10	BirA-heh2 linker-Ssh1; Rpl16a/b-HA-TEV-AVI	y2	a	<i>ssh1::BirA-heh2-SSH1::HIS5 rpl16a::RPL16a- HA-TEV-AVI rpl16b::RPL16b-HA-TEV-AVI ura3Δ0 met15Δ0 his3Δ1</i>
y11	Om45-mVenus-BirA; Rpl16a/b-HA-TEV-AVI	y2	a	<i>om45::OM45-mVenus-BirA::HIS5 rpl16a::RPL16a-HA-TEV-AVI rpl16b::RPL16b- HA-TEV-AVI leu2::pPGK1-BirA::LEU2 ura3Δ0 met15Δ0 his3Δ1</i>
y12	sec65-1 SRP-ts; Sec63-mVenus-BirA; Rpl16a/b-HA-TEV-AVI	y7	a	<i>sec65::sec65-1::KAN sec63::SEC63-mVenus- BirA::HIS5 rpl16a::RPL16a-HA-TEV-AVI rpl16b::RPL16b-HA-TEV-AVI ura3Δ0 met15Δ0 his3Δ1</i>
y13	sec66Δ; BirA-mVenus- Ubc6; Rpl16a/b-HA- TEV-AVI	y8	a	<i>sec66Δ::KAN his3::pSEC63-BirA-mVenus-UBC6 TA-HIS5 rpl16a::RPL16a-HA-TEV-AVI rpl16b::RPL16b-HA-TEV-AVI ura3Δ0 met15Δ0 leu2Δ0</i>
y14	Hsf1 reporter	BY4741	a	<i>ura3::4x-HSFE-GFP::URA3 leu2Δ0 met15Δ0 his3Δ1</i>
y15	BirA-mVenus-Ssh1; Hsf1 reporter	y14	a	<i>ssh1::BirA-mVenus-SSH1::HIS5 ura3::4x-HSFE- GFP::URA3 leu2Δ0 met15Δ0 his3Δ1</i>
y16	BirA-heh2 linker-Ssh1; Hsf1 reporter	y14	a	<i>ssh1::BirA-heh2-SSH1::HIS5 ura3::4x-HSFE- GFP::URA3 leu2Δ0 met15Δ0 his3Δ1</i>
y17	ssh1Δ; Hsf1 reporter	y14	a	<i>ssh1Δ::KAN ura3::4x-HSFE-GFP::URA3 leu2Δ0 met15Δ0 his3Δ1</i>

Table 3-2. Cell lines

name	description	genotype
293T_1	ATH-mouseRPL10A	pMK1047-puro-ATHmmuRPL10A lenti infection
293T_2	ATH-mouseRPL10A + ER birA	pMK1047-puro-ATHmmuRPL10A lenti infection; pINDUCER11 3xFLAG-BirA-mCherry-Sec61 β infection

Table 3-3. Plasmids

name	description
p1	RPS2 HA-TEV-AVI pop-in/pop-out plasmid
p2	RPL16a HA-TEV-AVI pop-in/pop-out plasmid
p3	RPL16b HA-TEV-AVI pop-in/pop-out plasmid
p4	pRS305 pPGK1-BirA::LEU2
p5	pFA6a -mVenus-BirA::HIS5
p6	pFA6a BirA-mVenus-SSH1::HIS5
p7	pFA6a BirA-heh2 linker-SSH1::HIS5
p8	pFA6a pSEC63-BirA-mVenus-UBC6-TA::HIS5
p9	pCJ28 pEF1a-puro-T2A-ATHmmuRPL10A(cDNA)
p10	pCJ33 N-terminal 3X FLAG fused to wt BirA - mCherry - sec61 β

Table 3-4. Primers

name	purpose	sequence (5' to 3')
oGAB11	28mer size standard	agucacuuagcgauguacacugacugug/3phos/
oCJ200	Primer for reverse transcription of sequencing libraries	5'/5phos/GATCGTCCGGACTGTAGAACTCTGAACCTGTGCG/isp18/CA AGCAGAAGACGGCATAACGAGATATTGATGGTGCCTACAG
oNTI231	Amplification of sequencing libraries, paired with indexing primer	caagcagaagacggcatcacga
Linker 1	3' Cloning adaptor for RNA footprints	appctgtagccatcaat/3ddc
oCW52	C-terminally tag SEC63 with -mVenus-BirA::HIS5, forward primer	cgatacggatacagaagctgaagatgatgaatcaccagaaATCGGTGACGGTGCTG GTTT
oCW53	C-terminally tag SEC63 with -mVenus-BirA::HIS5, reverse primer	tctaagagctaaatgaaaaactactaatcacttataTGGATGGCGGCGTTAGTA TC
oCW304	Replace genomic SSH1 with BirA-linker-Ssh1::HIS5, forward primer	ttttagcacatttgccccccactctccattgttttagtATGAAGGATAACACCGTGC C
oCW305	Replace genomic SSH1 with BirA-linker-Ssh1::HIS5, reverse primer	tcttctttttgttttcttcttttttcttttctttgtTGGATGGCGGCGTTAGTATC
oCW294	Target SEC63p-BirA-mVenus-Ubc6::HIS5 to HIS3 locus, forward primer	atactaaaaaatgagcaggcaagataaacgaagcaagGGTGCTCTCATGCTTCT TTTC
oCW252	Target SEC63p-BirA-mVenus-Ubc6::HIS5 to HIS3 locus, reverse primer	atgtatatatcgtatgctgcagcttaataatcggtgTGGATGGCGGCGTTAGTA TC
oJD75	Target pPGK1-BirA::LEU2 to LEU2 locus, forward primer	cattattacgccctcttgacctctaatcatgaatgttctcGAGCAGATTGTACTGAGA G
oJD76	Target pPGK1-BirA::LEU2 to LEU2 locus, reverse primer	tgagccattagatcaatttgcttacctgtattcctttaCAGGAAACAGCTATGACCAT G
oCW272	Delete Sec66 with KAN MX cassette, forward primer	tacaggaaagaggtacgcacaactacttgagttgccaatGACATGGAGGCCAGAA TAC
oCW273	Delete Sec66 with KAN MX cassette, reverse primer	cactgaacgagcgaatacatatctttgcacacagtaggcaTGGATGGCGGCGTTAGT ATC

REFERENCES

- Algire, M.A., Maag, D., Savio, P., Acker, M.G., Tarun, S.Z., Jr, Sachs, A.B., Asano, K., Nielsen, K.H., Olsen, D.S., Phan, L., et al. (2002). Development and characterization of a reconstituted yeast translation initiation system. *RNA N. Y. N* 8, 382–397.
- Ast, T., Cohen, G., and Schuldiner, M. (2013). A Network of Cytosolic Factors Targets SRP-Independent Proteins to the Endoplasmic Reticulum. *Cell* 152, 1134–1145.
- Bonven, B., and Gulløv, K. (1979). Peptide chain elongation rate and ribosomal activity in *Saccharomyces cerevisiae* as a function of the growth rate. *Mol. Gen. Genet. MGG* 170, 225–230.
- Borgese, D., Blobel, G., and Sabatini, D.D. (1973). In vitro exchange of ribosomal subunits between free and membrane-bound ribosomes. *J. Mol. Biol.* 74, 415–438.
- Brandman, O., Stewart-Ornstein, J., Wong, D., Larson, A., Williams, C.C., Li, G.-W., Zhou, S., King, D., Shen, P.S., Weibezahn, J., et al. (2012). A Ribosome-Bound Quality Control Complex Triggers Degradation of Nascent Peptides and Signals Translation Stress. *Cell* 151, 1042–1054.
- Brar, G.A., Yassour, M., Friedman, N., Regev, A., Ingolia, N.T., and Weissman, J.S. (2012). High-resolution view of the yeast meiotic program revealed by ribosome profiling. *Science* 335, 552–557.
- Christensen, A.K., Kahn, L.E., and Bourne, C.M. (1987). Circular polysomes predominate on the rough endoplasmic reticulum of somatotropes and mammatropes in the rat anterior pituitary. *Am. J. Anat.* 178, 1–10.
- Devaraneni, P.K., Conti, B., Matsumura, Y., Yang, Z., Johnson, A.E., and Skach, W.R. (2011). Stepwise Insertion and Inversion of a Type II Signal Anchor Sequence in the Ribosome-Sec61 Translocon Complex. *Cell* 146, 134–147.
- Elstner, M., Andreoli, C., Klopstock, T., Meitinger, T., and Prokisch, H. (2009). The mitochondrial proteome database: MitoP2. *Methods Enzymol.* 457, 3–20.
- Fons, R.D. (2003). Substrate-specific function of the translocon-associated protein complex during translocation across the ER membrane. *J. Cell Biol.* 160, 529–539.
- Goder, V. (2003). Molecular mechanism of signal sequence orientation in the endoplasmic reticulum. *EMBO J.* 22, 3645–3653.
- Görlich, D., Hartmann, E., Prehn, S., and Rapoport, T.A. (1992). A protein of the endoplasmic reticulum involved early in polypeptide translocation. *Nature* 357, 47–52.
- Jiang, Y., Cheng, Z., Mandon, E.C., and Gilmore, R. (2008). An interaction between the SRP receptor and the translocon is critical during cotranslational protein translocation. *J. Cell Biol.* 180, 1149–1161.

Käll, L., Krogh, A., and Sonnhammer, E.L. (2004). A Combined Transmembrane Topology and Signal Peptide Prediction Method. *J. Mol. Biol.* 338, 1027–1036.

Kowarik, M., Küng, S., Martoglio, B., and Helenius, A. (2002). Protein folding during cotranslational translocation in the endoplasmic reticulum. *Mol. Cell* 10, 769–778.

Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E.L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* 305, 567–580.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10, R25.

Longtine, M.S., McKenzie, A., 3rd, Demarini, D.J., Shah, N.G., Wach, A., Brachat, A., Philippsen, P., and Pringle, J.R. (1998). Additional modules for versatile and economical PCR-based gene deletion and modification in *Saccharomyces cerevisiae*. *Yeast* Chichester Engl. 14, 953–961.

Meerbrey, K.L., Hu, G., Kessler, J.D., Roarty, K., Li, M.Z., Fang, J.E., Herschkowitz, J.I., Burrows, A.E., Ciccia, A., Sun, T., et al. (2011). The pINDUCER lentiviral toolkit for inducible RNA interference in vitro and in vivo. *Proc. Natl. Acad. Sci.* 201019736.

Meinema, A.C., Laba, J.K., Hapsari, R.A., Otten, R., Mulder, F.A.A., Kralt, A., van den Bogaart, G., Lusk, C.P., Poolman, B., and Veenhoff, L.M. (2011). Long Unfolded Linkers Facilitate Membrane Protein Import Through the Nuclear Pore Complex. *Science* 333, 90–93.

Ng, D.T., Brown, J.D., and Walter, P. (1996). Signal sequences specify the targeting route to the endoplasmic reticulum membrane. *J. Cell Biol.* 134, 269–278.

Ogg, S.C., and Walter, P. (1995). SRP samples nascent chains for the presence of signal sequences by interacting with ribosomes at a discrete step during translation elongation. *Cell* 81, 1075–1084.

Pagliarini, D.J., Calvo, S.E., Chang, B., Sheth, S.A., Vafai, S.B., Ong, S.-E., Walford, G.A., Sugiana, C., Boneh, A., Chen, W.K., et al. (2008). A mitochondrial protein compendium elucidates complex I disease biology. *Cell* 134, 112–123.

Panzner, S., Dreier, L., Hartmann, E., Kostka, S., and Rapoport, T.A. (1995). Posttranslational protein transport in yeast reconstituted with a purified complex of Sec proteins and Kar2p. *Cell* 81, 561–570.

Petersen, T.N., Brunak, S., von Heijne, G., and Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* 8, 785–786.

Rothstein, R. (1991). Targeting, disruption, replacement, and allele rescue: integrative DNA transformation in yeast. *Methods Enzymol.* 194, 281–301.

Schuldiner, M., Metz, J., Schmid, V., Denic, V., Rakwalska, M., Schmitt, H.D., Schwappach, B., and Weissman, J.S. (2008). The GET complex mediates insertion of tail-anchored proteins into the ER membrane. *Cell* 134, 634–645.

- Siegel, V., and Walter, P. (1988). The affinity of signal recognition particle for presecretory proteins is dependent on nascent chain length. *EMBO J.* 7, 1769–1775.
- Smith, J.J., and Aitchison, J.D. (2013). Peroxisomes take shape. *Nat. Rev. Mol. Cell Biol.* 14, 803–817.
- Stefanovic, S., and Hegde, R.S. (2007). Identification of a Targeting Factor for Posttranslational Membrane Protein Insertion into the ER. *Cell* 128, 1147–1159.
- Stirling, C.J., Rothblatt, J., Hosobuchi, M., Deshaies, R., and Schekman, R. (1992). Protein translocation mutants defective in the insertion of integral membrane proteins into the endoplasmic reticulum. *Mol. Biol. Cell* 3, 129–142.
- Tabas-Madrid, D., Nogales-Cadenas, R., and Pascual-Montano, A. (2012). GeneCodis3: a non-redundant and modular enrichment analysis tool for functional genomics. *Nucleic Acids Res.* 40, W478–483.
- Wells, S.E., Hillner, P.E., Vale, R.D., and Sachs, A.B. (1998). Circularization of mRNA by eukaryotic translation initiation factors. *Mol. Cell* 2, 135–140.
- Williams, C.C., Jan, C.H., and Weissman, J.S. Trafficking and plasticity of the mitochondrial proteome revealed by proximity-specific ribosome profiling. *Science* (submitted).
- Young, B.P. (2001). Sec63p and Kar2p are required for the translocation of SRP-dependent precursors into the yeast endoplasmic reticulum in vivo. *EMBO J.* 20, 262–271.
- Van der Zand, A., Braakman, I., and Tabak, H.F. (2010). Peroxisomal Membrane Proteins Insert into the Endoplasmic Reticulum. *Mol. Biol. Cell* 21, 2057–2065.

CHAPTER FOUR

Trafficking and plasticity of the mitochondrial proteome

INTRODUCTION

The vast majority of mitochondrial proteins are encoded within the nuclear genome and must be targeted to their final mitochondrial destination. Here, we utilize proximity-specific ribosome profiling to explore several areas of protein targeting to the mitochondria that have remained difficult to assay with other methodologies.

First, our approach enabled us to define an important subset of proteins that are targeted co-translationally to mitochondria *in vivo*, resolving a decades-long debate in the field regarding the mechanism of protein insertion. Second, the sensitivity of our approach enabled us to confidently identify dozens of previously unappreciated proteins that target to the mitochondria, highlighting the utility of our method as a proteomics tool. Third, synthesis of our mitochondrial and endoplasmic reticulum (ER) data sets provides the first spatially resolved proteomic maps of proteins which target to the two organelles in living cells, without the confounding impurities associated with biochemical preparations. Comparison of these data emphasized the exquisite specificity of co-translational protein trafficking *in vivo* and revealed that soluble mitochondrial proteins more readily evolve extra-mitochondrial localizations over time compared to proteins of the secretory pathway. Finally, our ribosome profiling data revealed an intriguing mechanism whereby alternative translation initiation sites enables localization of the fumarate reductase Osm1 to both the mitochondria and ER. The discovery of an ER form of Osm1 suggests an answer to a long-standing question of what drives oxidative folding in the ER.

RESULTS

Mitochondria are ancient membrane-bound organelles that are derived from an endosymbiotic α -proteobacterium. The modern mitochondrial proteome is encoded by genes that are both eukaryotic and prokaryotic in origin; in *Saccharomyces cerevisiae*, ~99% of

mitochondrial proteins are encoded in the nuclear genome (Elstner et al., 2009) where they comprise ~15% of protein-coding genes (Cherry et al., 2012). These proteins are translated in the cytosol and must be sorted into the matrix, inner membrane (IM), intermembrane space (IMS), and outer membrane (OM) by a complex network of channels and chaperones. How mitochondrial proteins evolved into nuclear-encoded genes is an open question that is intimately connected to the mechanism of targeting to the organelle.

While the predominant view is that protein import into mitochondria occurs post-translationally on unfolded polypeptides, a role for co-translational translocation has been suggested for some substrates (Fox, 2012). Moreover, ribosomes have been observed on the OM by electron microscopy (Kellems et al., 1975) and numerous mRNAs associate closely with mitochondria (Marc et al., 2002; Saint-Georges et al., 2008; Suissa and Schatz, 1982), suggesting a broader role for co-translational protein insertion. However, those studies generally employed translation elongation inhibitors, such as cycloheximide (CHX), to preserve ribosome nascent-chain complexes (RNCs); such reagents confound the co-translational interpretation as they increase the association of RNCs with mitochondria by providing those containing exposed targeting sequences an artificially-prolonged time to engage the translocon.

To directly evaluate which proteins are translated on the surface of mitochondria in an unperturbed context, we applied the proximity-specific ribosome profiling technique established in the accompanying manuscript (Jan et al., 2014). This approach involves *in vivo* biotinylation of Avi-tagged ribosomes that are in contact with a spatially localized biotin ligase (BirA), followed by readout of their translational activity by deep sequencing of ribosome-protected fragments. To selectively mark ribosomes on the mitochondrial surface, we fused BirA to the C-terminus of OM45, a major component of the OM (Fig. 4-1A). Whereas a cytosolic BirA was capable of efficiently labeling both Avi-tagged large (Rpl16 /

uL13) and small (Rps2 / uS5) subunits (Jan et al., 2014), Om45-BirA was only able to biotinylate the large subunit indicating that labeled ribosomes are constrained at the mitochondria with their peptide exit tunnels facing the OM (Fig. 4-1B). This pattern was also observed with ER-localized BirA (Jan et al., 2014), where ribosomes are constrained due to translocon-binding during co-translational insertion, arguing that ribosome marked by Om45-BirA are engaged in mitochondrial protein translocation.

Proximity-specific profiling experiments were performed using a brief (two-minute) pulse of biotin in the absence of any translation inhibitors, to preserve the *in vivo* rates of translation and targeting. Gene ontology analysis confirmed that the enriched genes were overwhelmingly mitochondrial (Fig. 4-S1) and a distinct subset of expressed mitop2 mitochondrially localized reference genes (151/574) were co-translationally targeted to the mitochondria (Fig. 4-1C). This subset was strongly enriched for IM proteins; most IM proteins were co-translationally targeted whereas only a minor fraction of OM, IMS, and matrix proteins were targeted as RNCs (Fig. 4-1D). By virtue of their transmembrane domains (TMDs), the IM proteins must simultaneously avoid aggregation in the cytosol and errant integration into other membranes. Co-translational insertion may minimize the potential for toxicity associated with the accumulation of membrane proteins in the cytosol and, intriguingly, mirrors the SRP-dependent co-translational insertion of TMD-containing proteins in the bacterial ancestor of mitochondria (Saraogi and Shan, 2011).

Omission of translation elongation inhibitors in the above experiments allowed us to define the proteins that are unambiguously translated at the mitochondrial surface. However, this set of genes represented a subset of the mRNAs that purify with mitochondria in the presence of CHX (Marc et al., 2002). We therefore reasoned that by providing a prolonged time for RNCs to engage mitochondria, CHX pretreatment would yield a more comprehensive view of the mitochondrial proteome. Indeed, when we included CHX, we

maintained mitochondrial specificity (Fig. 4-S1) but observed a large increase in the number of enriched proteins (Fig. 4-2A). For example, 108/183 mitop2-annotated matrix proteins were enriched (Fig. 4-2B).

We focused on the subset of genes that had a high probability of containing N-terminal MTSs, as predicted by MitoProt (Claros and Vincens, 1996). The bimodal distribution of enrichments for this group raised the question of why some RNCs were translocation incompetent. Certain genes related to tRNA function, *VAS1* (Chatton et al., 1988) and *TRM1* (Ellis et al., 1989), were primarily transcribed from alternative promoters that remove their MTSs (Fig. 4-S2). More striking was the difference in protein size between the enriched and depleted matrix proteins. Examination of gene enrichments binned by ORF length revealed a cutoff of ~180 codons, below which very few RNCs and above which the large majority of RNCs were observed at the mitochondria (Fig. 4-2C). These shorter nascent chains may be unavailable for targeting due to shielding by ribosome-associated factors, such as Hsp70s (Willmund et al., 2013). Alternatively, this could reflect a minimum length required for the efficient tethering of initiation-competent mRNAs to mitochondria mediated by 3' RNCs, which is required for efficient co-translational translocation at the ER (Jan et al., 2014). Such a model is consistent with the observed N-terminal (i.e., prior to residue 180; Fig. 4-S3) enrichment of proteins longer than 180 codons.

Regardless of the precise mechanism that prevents shorter RNCs from reaching the mitochondria, enrichment provides a sensitive and reliable measurement of mitochondrial protein localization as the enriched genes were nearly exclusively mitochondrial (Fig. 4-S1). We identified 36 novel proteins translated at the mitochondria with no prior mitochondrial annotation in mitop2, GO, or GFP databases (Cherry et al., 2012; Elstner et al., 2009; Huh et al., 2003). On average, these genes were no less-likely to contain predicted MTSs than the mitop2 reference set (Fig. 4-S1C), arguing that they are *bona fide* residents of this organelle

(Claros and Vincens, 1996). The sensitivity of proximity-dependent ribosome profiling thus complements fluorescence and mass spectrometry-based proteomic approaches for cataloguing subcellular protein localization.

Among the newly identified mitochondrial proteins was Hap1, a heme-responsive transcription factor (Guarente et al., 1984), which acts in the nucleus. We found that a Hap1-GFP fusion expressed from the endogenous locus is found both in the nucleus, as expected, and in the mitochondria as predicted by our translational enrichments (Fig. 4-2D). Heme biosynthesis occurs in the mitochondria and is regulated by Hap1. Because MTS-mediated import requires an energized IM, Hap1 localization may allow for direct sensing of mitochondrial health. Hap1 sequestration in mitochondrial may allow the cell to tune nuclear transcriptional activity through Hap1 localization (Nargund et al., 2012). Additionally, Hap1 may have uncharacterized functions within mitochondria.

We next investigated the evolution of mitochondrial protein localization among paralogs. *Saccharomyces cerevisiae* underwent a whole-genome duplication (WGD) ca. 100 million years ago, making it well suited for exploring changes in protein localization between paralogs (termed ohnologs) (Wolfe, 2000). With the exception of the IM proteins, we find remarkable fluidity in the targeting of ohnologs to mitochondria (Fig. 4-3A). By contrast, targeting of ohnologs to the ER only rarely showed discrepancies between paralog pairs (Fig. 4-3B), thus emphasizing the relative plasticity of the mitochondrial proteome. Ohnologs for which only one gene is mitochondrial are evolving at nearly the same rate as other pairs (Fig. 4-3C), arguing against the hypothesis of single-paralog gene death. IM and secreted TMD proteins reside in a lipid membrane and soluble secreted proteins in an oxidizing environment thereby restricting the ability of these proteins to function outside their normal compartments. By contrast, the matrix and cytosol have similar environments, which would allow proteins to function in either location. This plasticity between the matrix and cytosolic

proteins may have facilitated or have been driven by the metabolic changes associated with a switch toward preferring fermentation over respiration that evolved following the WGD in budding yeast (the Crabtree effect) (Thomson et al., 2005).

We further exploited our mitochondrial and ER protein localization data sets to explore the extent to which individual proteins show dual localization between the two intracellular sites. The tight physical association of these organelles makes the biochemical preparation of pure mitochondria and microsomes challenging, confounding searches for dual-localization in such preparations. For example the ERMES complex, which tethers the ER and mitochondria (Kornmann et al., 2009), was originally thought to be exclusively mitochondrial. Proximity-specific ribosome profiling of the mitochondria and ER correctly resolved components of the ERMES complex to their respective compartments, and provided highly specific catalogs of proteins at each organelle (Fig. 4-4A, 4-4B, 4-S4).

Remarkably, we found that very few proteins exhibited dual localization. We find most proteins with conflicting secretory/mitochondrial annotations (Cherry et al., 2012) to be co-translationally inserted exclusively to the ER, although a small subset are translated in the cytosol and trafficked to one or both compartments, post-translationally (Fig. 4-4B). Among the few counter examples was the phosphatase *PTC7*, which undergoes alternative intron retention leading to targeting of the spliced form to the mitochondria and the unspliced form to the nuclear envelope (Juneau et al., 2009).

The fumurate reductase *OSM1*, also stood out as being robustly targeted to both the mitochondria and ER. While several studies suggest that *Osm1* localizes exclusively to the mitochondria (Enomoto et al., 2002; Huh et al., 2003; Muratsubaki and Enomoto, 1998), high-throughput studies hint at a possible ER function. Specifically, a genetic-interaction data set indicated that *osm1Δ* is synthetic lethal with a temperature-sensitive allele of *ERO1* (Costanzo et al., 2010), the protein responsible for driving disulfide bond formation in the ER.

Additionally, Osm1 was found in the N-glycoproteome, consistent with its entry into the secretory pathway (Zielinska et al., 2012). Quantitative western blot analysis of Osm1 with or without endoglycosidase H treatment revealed that Osm1 exists in both a glycosylated and lower molecular weight unglycosylated form consistent with a dual localization to both the ER and mitochondria (Fig. 4-4C). We confirmed this directly by fluorescence microscopy of GFP-tagged Osm1 expressed from its endogenous locus (Fig 4E).

Given the exquisite specificity of the vast majority of ER SSs and MTSs, how is that Osm1 can be efficiently targeted to both organelles? Ribosome profiling experiments using lactimidomycin (LTM), which leads to the accumulation of ribosomes at translation start sites (Stern-Ginossar et al., 2012), revealed that *OSM1* possesses two functional in-frame start codons (Fig. 4-4D). Intriguingly, the longer and shorter forms are predicted to have ER and MTS signals, respectively. Mutation of the upstream methionine to alanine (Met1Ala) to produce only the short form resulted in strict mitochondrial localization (Fig. 4-4E). By contrast, the Met32Ala mutant, which can only express the long form, localized exclusively to the ER (Fig. 4-4E). The localization of the Met32Ala mutant further established the primacy of ER-trafficking when a protein contains both an ER and a downstream mitochondrial targeting sequence.

Interestingly, we found that Osm1 homologs, including those of non-WGD yeast species, are predicted to contain signal sequences (Byrne and Wolfe, 2005) (Fig. 4-S5), emphasizing the conserved ER localization of this class of fumarate reductases. Given the synthetic lethal interaction between a temperature sensitive allele of *ERO1* and loss of *OSM1*, an appealing function for this ER-localized form is to drive oxidative protein folding. Consistent with this, studies of oxidative folding under anaerobic conditions have suggested that fumarate is the terminal electron acceptor (Liu et al., 2013) and Osm1 generates free oxidized FAD via the reduction of fumarate. Moreover, Ero1 is a flavoprotein and reoxidation

of Ero1 following disulfide bond formation can be driven by oxidized FAD (Gross et al., 2006; Tu and Weissman, 2002). The identification of a conserved ER form of fumarate reductase would thus provide a mechanism for generating free oxidized FAD that is directly accessible to Ero1.

PERSPECTIVE

Proximity-specific ribosome profiling allowed us to elucidate several fundamental aspects of mitochondrial biogenesis. First, we discovered a major role for co-translational targeting, but one that is largely limited to IM proteins. The similarity of co-translational insertion at the ER and of IM proteins reveals a convergent paradigm for trafficking of aggregation-prone alpha-helical TMD-containing proteins, though the mechanism responsible for IM proteins remains an open question. Second, we show that ohnologs encoding soluble proteins exhibit a high degree of fluidity in their mitochondrial localization, facilitating the exchange of function between the mitochondrial matrix and the cytosol. Third, we demonstrate the exquisite specificity of protein trafficking to the ER versus mitochondria *in vivo*. Our data suggest that co-translational trafficking occurs with higher fidelity than post-translational targeting, such as tail anchor insertion (Okreglak and Walter, 2014). How targeting factors, such as SRP (Akopian et al., 2013) and NAC (Alamo et al., 2011), act on dynamic and heterogeneous RNCs to create this specificity *in vivo* remains an important open question that could broadly inform our understanding of how the cell achieves and exploits local mRNA translation.

Acknowledgements

We thank members of the Weissman and P. Walter labs for discussion; E. Costa and V. Okreglak for critical reading of the manuscript. The UCSF CAT and NIC provided technical support. We thank Cristina Policarpi for assistance in strain construction and pilot ribosome profiling experiments, Gloria Brar for sharing unpublished ribosome profiling data, and Peter

Walter for use of the Su9-TagBFP plasmid. This research was supported by the Center for RNA Systems Biology and the Howard Hughes Medical Institute. C.C.W. is an NSF GRF. C.H.J. is the Rebecca Ridley Kry Fellow of the Damon Runyon Cancer Research Foundation (DRG-2085-11).

FIGURES

Figure 4-1

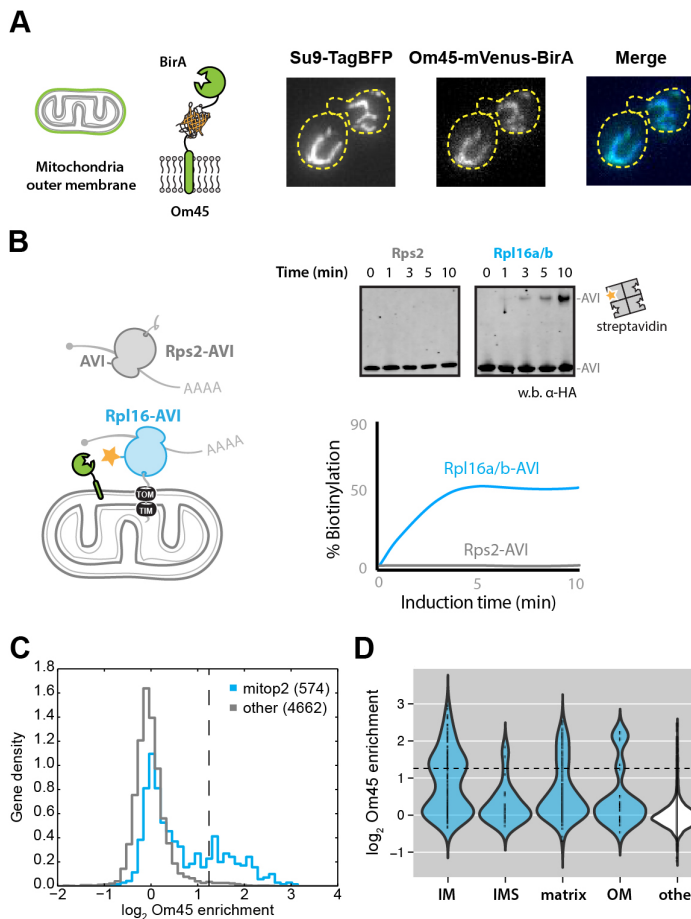


Fig. 4-1. Co-translational mitochondrial protein targeting as revealed by proximity-specific ribosome profiling. (A) Localization of biotin ligase to the outer membrane. Schematic of the Om45 fusion protein and confocal fluorescence images of Om45-mVenus-BirA and Su9-TagBFP are shown. **(B)** Assessment of ribosome orientation at the mitochondrial outer membrane. The kinetics of Avi-tagged Rps2 (grey) and Rpl16a/b (blue) biotinylation by Om45-mVenus-BirA were determined by western blot analysis of streptavidin-shift gels. **(C)** Histogram of \log_2 ratios for biotinylated footprints / input footprints in the absence of CHX. Genes annotated in the mitop2 reference set are shown in blue and enrichment threshold shown as a dashed vertical line. **(D)** Violin plot showing \log_2 gene enrichments for proteins grouped by their annotated location within mitochondria in the absence of CHX arrest. Dashed enrichment threshold matches that in (C).

Figure 4-2

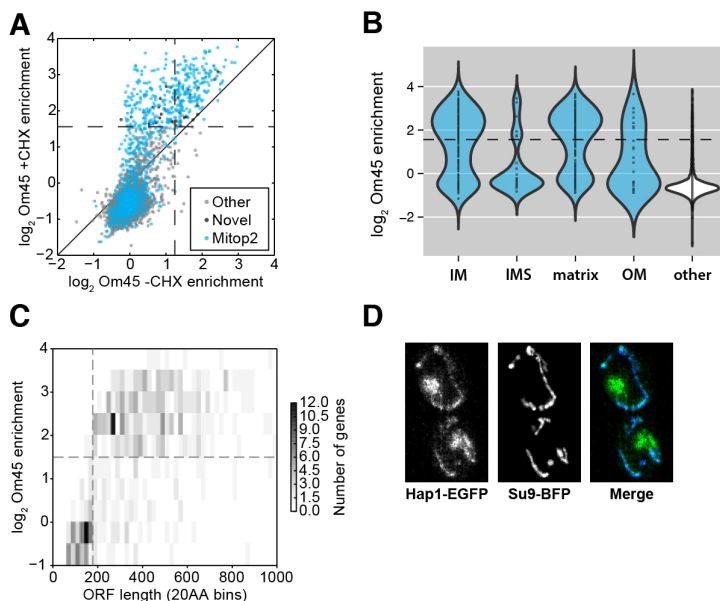


Fig. 4-2. Comprehensive characterization of RNC targeting to the mitochondria. (A) Scatter plot of log₂ gene enrichments for footprints from ribosomes biotinylated by Om45-mVenus-BirA in the presence or absence of CHX. Novel genes represent those which were significantly enriched but had no known mitochondrial annotations. (B) Violin plot showing log₂ gene enrichments for proteins grouped by their location within mitochondria in the presence of CHX arrest. (C) 2D histogram of genes binned by ORF length compared to their log₂ gene enrichment in the presence of CHX. (D) Epifluorescence microscopy of Hap1-EGFP and Su9-TagBFP.

Figure 4-3

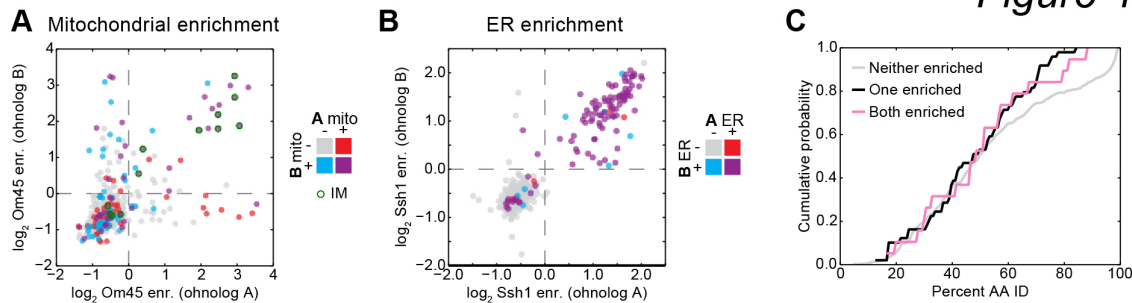


Fig. 4-3. Conservation of paralogous protein localization. (A) Scatter plot of mitochondrial log₂ gene enrichments for ohnologs. X- and Y-axis assignment is arbitrary. (B) As in A, for ER enrichments. (C) Cumulative distributions of percent identity between ohnologs where neither, one, or both paralogs were enriched at the mitochondria.

Figure 4-4

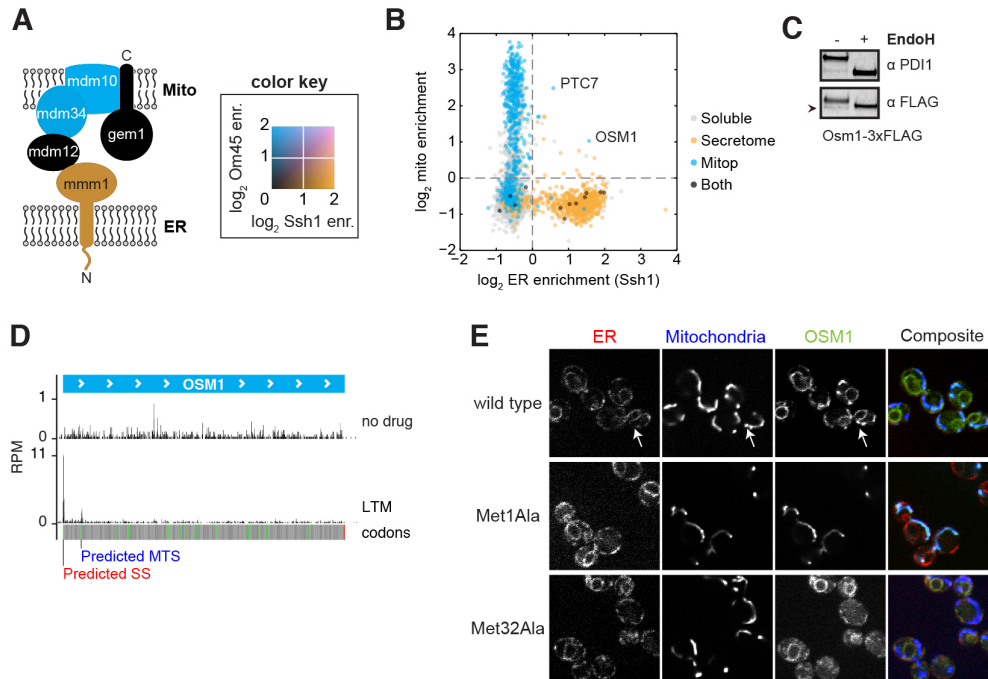


Fig. 4-4. Dual localization of proteins to the mitochondria and ER. (A) The color-mapped \log_2 enrichments for ribosome labeling by Om45 (blue) and Ssh1 (orange) are overlaid for members of the ERMES complex. (B) Scatter plot of ER enrichments compared to mitochondrial enrichments in the presence of CHX. Points are colored by their annotated localization. (C) Western blot analysis of N-linked glycans. Extracts from cells expressing Osm1-3xFLAG were blotted for PDI1 or FLAG with or without EndoH treatment. Arrowhead indicates an Osm1 band migrating below the size of the deglycosylated protein. (D) Ribosome occupancy of Osm1 with or without LTM treatment. In-frame AUG codons are shown in green; predicted localization of proteins made from the indicated start codon are shown. (E) Confocal fluorescence microscopy of Osm1-EGFP localization compared to ER (Sec63-mCherry) or mitochondria (Su9-BFP). Wild type, Met1Ala, and Met32Ala variants of Osm1 are shown.

MATERIALS AND METHODS

1 Experimental

A complete list of all yeast strains, plasmids, and primers used in this study are listed in Tables 4-1 to 4-3, respectively.

1.1 Strain construction

Yeast strains expressing tagged ribosomal proteins and Om45-BirA are described in (Jan et al., 2014). Hap1 was chromosomally tagged at its C-terminus with yeast optimized

GFP::HIS5 (Lee et al., 2013) in W303-1A, which lacks the Ty-insertion present in S288C. This strain was crossed to W303-1B expressing Su9-TagBFP (Okreglak and Walter, 2014). For glycosylation analysis, Osm1 was chromosomally tagged at its C-terminus with 3xFLAG::KAN using p1. The genomic copy of OSM1 in W303 was replaced with WT, M1A, or M32A variants of an Osm1-yoEGFP::LEU2 cassette, cloned from (Lee et al., 2013). Chromosomal integrants were verified with Sanger sequencing.

1.2 Media, proximity-specific ribosome profiling

Growth and biotin induction were performed as described in (Jan et al., 2014). Biotinylation was allowed to proceed for two minutes before harvesting, in the presence or absence of CHX as dictated by the experiment.

1.3 Microscopy

Black 96 well glass (#1.5) bottom plates were treated with 0.1 mg/mL Concanavalin A for 10 minutes. The ConA was aspirated and plates allowed to dry. 200 μ L of yeast cells grown to OD < 0.2 in SD media were applied to each well and allowed to adhere for approximately 15 minutes before imaging. The localization of Hap1 was assessed by live cell microscopy on a widefield epifluorescence microscope illuminated by a SPECTRA X light engine (lumencor). Samples were imaged through a 100X Nikon Plan Apo VC 1.4 oil objective onto a Zyla sCMOS camera (Andor). GFP was measured at 475(ex)/521(em) and BFP measured at 390(ex)/440(em). Om45-BirA and Osm1 localization was determined by live cell Spinning-Disc confocal microscopy. Plates were imaged with a 100X Nikon Apo TIRF 1.49 oil objective onto Photometrics Evolve EMCCD attached to a Nikon Ti-E microscope equipped with a Yokogawa CSU-22 spinning disc in micromanager. GFP and mVenus were excited by a 491 nm Cobalt Calypso DPSS laser line, emission measured at 525 nm. mCherry fluorescence was excited by a 561 nm Coherent Sapphire DPSS laser line, emission measured at 610 nm.

1.4 Glycosylation analysis

Yeast cells expressing Osm1-3XFLAG (y6) were grown to an OD₆₀₀ of 0.6. 1 mL was pelleted at 20,000xg for one minute, then resuspended in 80 µL 1X LDS buffer (Invitrogen) supplemented with 100 mM DTT and 1x complete protease inhibitors (Roche) pre-warmed to 95°C. The lysate was incubated at 95°C for five minutes, vortexed and pelleted at 20,000xg for eight minutes. Five microliters of clarified lysate were supplemented with 4 µL of 0.5M Sodium Citrate (pH 5.5), 10 µL water and 1 µL EndoH (Roche) or water (negative control). The reaction was incubated at 37°C for one hour, then stopped by addition of 6 µL 4X LDS buffer. Proteins were separated on 4-12% NuPAGE bis-tris gels, transferred to PVDF and blotted with mouse anti-FLAG M2 (Sigma) or rabbit anti-PDI. Binding was detected with goat anti-Mouse IgG IR800 and goat anti-Rabbit IgG Alexa680 and imaged on an Odyssey IR scanner (Licor).

2 Computational analyses

Alignment of ribosome protected fragments, gene enrichments, positional enrichments, and gene ontology analysis were performed as described (Jan et al., 2014).

2.1 Gene annotations

a. Mitop2

The manually curated mitochondrial reference set was downloaded from the mitop2 database (<http://www.mitop2.de/>) (Elstner et al., 2009). Mitochondrial sub-localizations were extracted from the mitop2 reference set when available.

b. GO

GO-slim terms were downloaded from the Saccharomyces Genome Database (<http://yeastgenome.org>) and mined for any gene that included the search term “mito” as a cellular component.

c. GFP

Tabular data from the yeast GFP collection was downloaded from (<http://yeastgfp.yeastgenome.org/>) (Huh et al., 2003).

d. Secretome

The secretome was defined in (Ast et al., 2013; Jan et al., 2014)

e. Non-mitochondrial genes

The set of non-mitochondrial genes was defined as those absent from the mitop2 reference set, lacking a mitochondrial GO slim cellular component annotation and lacking a mitochondrial localization in the yeast GFP collection.

f. Whole genome duplication paralogs (Ohnologs)

Paralogs annotated as Ohnologs were downloaded from the Yeast Gene Order Browser (<http://ygob.ucd.ie/>) (Byrne and Wolfe, 2005).

g. MTS prediction

The probabilities of mitochondrial protein localization were determined using the Mitoprot webserver (<http://ihg.gsf.de/ihg/mitoprot.html>) (Claros and Vincens, 1996).

2.2 Data visualization and significance testing

All plots were generated by Python scripts using the matplotlib and seaborn libraries, or R scripts using the ggplot2 library. A pseudocount was added to FPKM measurements prior to calculating \log_2 enrichment values.

SUPPLEMENTARY FIGURES

Figure 4-S1

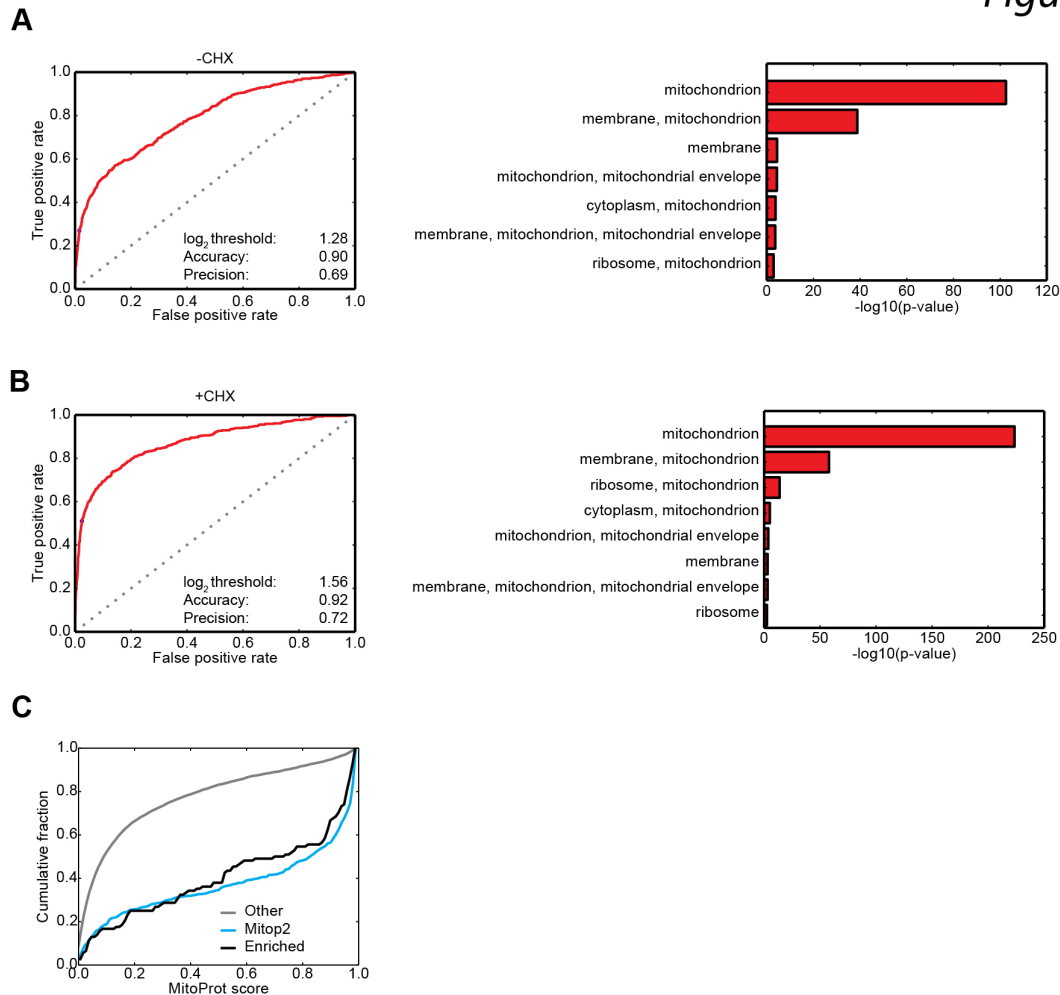


Fig. 4-S1. Gene enrichment analysis. (A) Receiver operator characteristic (ROC) curves were used to evaluate enrichment threshold classification of mitop2 reference set genes (used as the true positive set). The threshold that maximized accuracy was used to classify genes as enriched for downstream analyses. Enrichment of GO-slim cellular component terms in the genes above threshold is shown to the right. (B) As in (A), for experiments performed with CHX pre-treatment. (C) The cumulative distributions of MitoProt predictions scores are shown for mitop2 reference genes, enriched genes with no prior mitochondrial annotations and all other genes.

Figure 4-S2

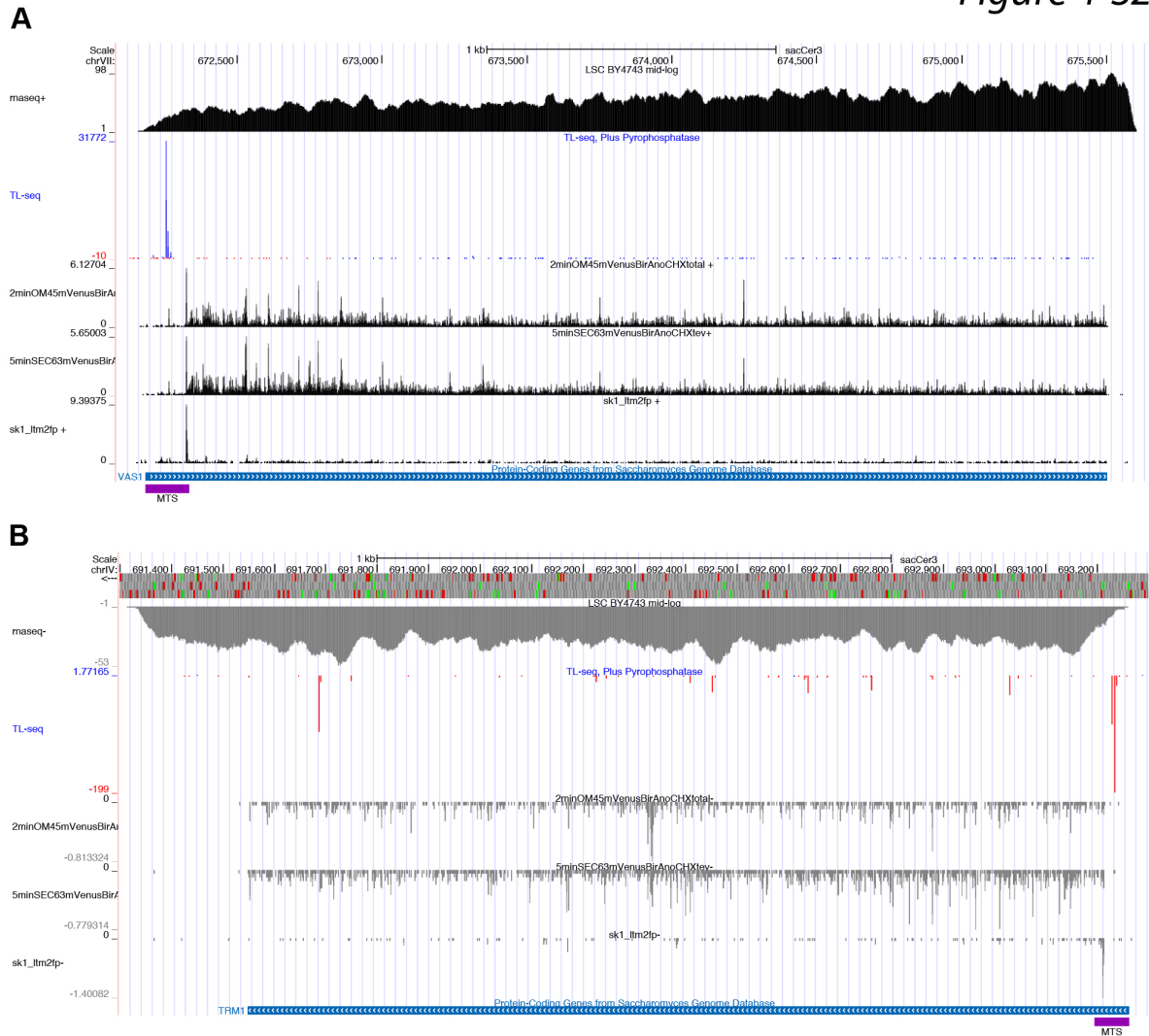


Fig. 4-S2. Alternative initiation removes the MTSs of Vas1 and Trm1. (A) Genome browser track showing RNA-Seq, Transcript Leader-Seq and ribosome profiling data at the Vas1 locus. Peaks in the Transcript-Leader-Seq indicate transcription start sites and peaks in the LTM track indicate translation start sites.

Figure 4-S3

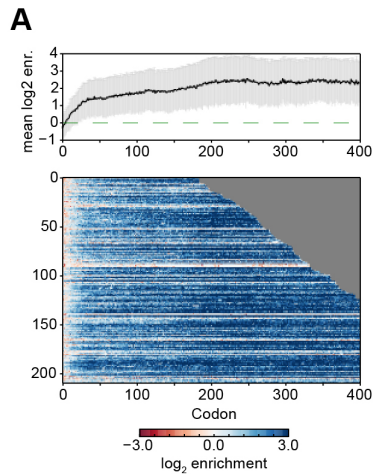


Fig. 4-S3. Position-specific enrichment analysis of mitop2 genes. (A) Shown is a colormap of enrichment at each codon (columns) for each mitop2 gene (rows) sorted by length. Only genes longer than 179 codons were plotted. The geometric mean enrichment and standard deviation are shown above.

Figure 4-S4

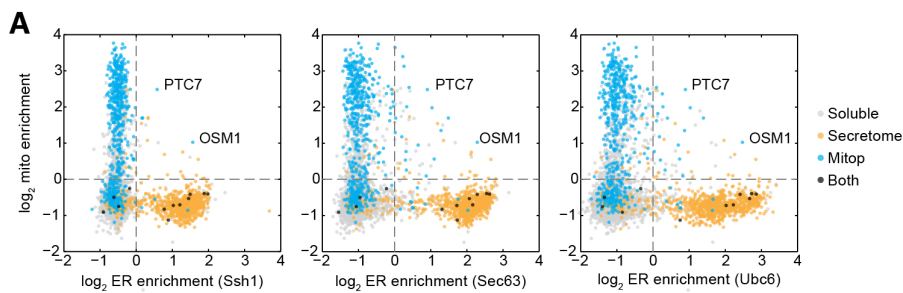


Fig. 4-S4. Consistency of dual localization across ER BirA fusions. (A) Scatter plots of ER enrichments for the indicated BirA fusion compared to mitochondrial enrichments. Points are colored by their annotated localization.

A

Homology (s=single gene, o=osm1, f=frd1)	Phobius prediction (SP=signal peptide, SA=signal anchor, NC=non-cytosolic)
s	APR367W --MLRGRSKTVLTLILLVTVRLLFNRLRSRGPESNRRIDMSGK-VQ--PVVVVGLGLAGLSTGAQLVKN--GVPVILMDKASAIIGNSVKASSINGAGTQVQELLGVY--DSADSFYRD SP
f	TBLA0B02110 -----MINLKKKTLISLLVTIFVAYSILSNDNLKSFINKNSNLSKTEPVVIGTGLAGLSAAYEILSTNPEVKVILLEKQAFVGNMKAASSINGALFSSIQKHFHIE--DSIEAFLED SP
o	TPHA0O00490 -----MPRLRKLILFYLVLVSVFFLVRRSFVNDSDGKEQ-----F--PVVIGSGLAGLSAVNQLLND--DVPVLLDKASSFGNSKASSINGAGTSTQVQKHHVIE--DSELPFLQD SP
s	Ecmj_3562 -----MKCLSRLLRLLLFIFGLILLKLYHNSRQSEDVTEDMGSA-VN--PVVIGSGLAGLTVTSVQL-SKY-KIPIILLDKASSIGNSIKASSINGAGTSTQVQKHHVIE--DSELPFLQD SP
s	KLLA0F16753g -----MRQRYLKILLLFVVVLYVLSLFSKMQATMS-----K--PVVIGSGLAGLTVTSVQL-KF-NIPIILLDKASSIGNSIKASSINGAGTSTQVQKHHVIE--DSELPFLQD SP
o	Kpo1_370.10 -----MPKIKRLLLLVITITVLYYKGFSRVKNHQ-----S--PVVIGSGLAGLTVTSVQLVSSH-NIPVILLEKSNLIGNSIKASSINGAGTSTQVQKHHVIE--DSELPFLQD SP
o	CAGL0I01320g -----MRGSRKLVVYITIALVLLYRKINSNSMY--KAVPK--S--PVVIGTGLAGLTVTSVQLAHTY-KIPVILLDKASSIGNSIKASSINGAGTSTQVQKHHVIE--DSELPFLQD SA
o	KAFR0I02810 -----MRRPVRKLLYVIFLITVVVYTVKQSKTFLRRTASK--A--PVVIGSGLAGLTVTSVQLVSSH-NIPVILLEKSNLIGNSIKASSINGAGTSTQVQKHHVIE--DSELPFLQD SP
o	Suva_12.135 -----MIRSVRRVFIYVSIFVLIILKRLQSLSTD--QM--K--Q--PVVIGSGLAGLTVTSVQLVSSH-NIPVILLEKSNLIGNSIKASSINGAGTSTQVQKHHVIE--DSELPFLQD SP
o	Skud_10.269 -----MARSVRRVFIYVSIFVLIILKRLQSLSTD--QMSLK--Q--PVVIGSGLAGLTVTSVQLVSSH-NIPVILLEKSNLIGNSIKASSINGAGTSTQVQKHHVIE--DSELPFLQD SP
o	Smik_10.329 MVLKLTVMIRSVRRVFIYVSIFVLIILKRLQSLSTD--QMSGK--H--PVVIGSGLAGLTVTSVQLVSSH-NIPVILLEKSNLIGNSIKASSINGAGTSTQVQKHHVIE--DSELPFLQD SA
o	YJR051W (Osm1) -----MIRSVRRVFIYVSIFVLIILKRLQSLSTD--QMSMK--Q--PVVIGSGLAGLTVTSVQLVSSH-NIPVILLEKSNLIGNSIKASSINGAGTSTQVQKHHVIE--DSELPFLQD SP
s	ZYR00F0522g -----MFSIRKRVTVVWVLLSFLVYKHLMLMAT-----R--PVVIGTGLAGLSAGNQLV-KH-KIPVILLDKASSIGNSIKASSINGAGTSTQVQKHHVIE--DSELPFLQD SP
o	KNAG0B00330 -----MRRPIRRLFLNLLIVVIMLKFPLRASKPAKRHISR--PVVIGTGLAGLTVTSVQLVSSH-NIPVILLEKSNLIGNSIKASSINGAGTSTQVQKHHVIE--DSELPFLQD SP
o	NDAI0B01200 -----MLSKTIKRVFVYLAIALILTRNTQNMSSKLLIT-----S--PVVIGSGLAGLTVTSVQLVSSH-NIPVILLEKSNLIGNSIKASSINGAGTSTQVQKHHVIE--DSELPFLQD SP
o	NCAS0A13040 -----MITKPIKRFIYSLITLILLYRHFNTNMSSKLLNLK--K--PVVIGSGLAGLTVTSVQLVSSH-NIPVILLEKSNLIGNSIKASSINGAGTSTQVQKHHVIE--DSELPFLQD SA
s	KLTH0C03146g -----MLTTRLLVGLFVYVLLLSNNSNSNQLPDMSSKPFST--PVVIGSGLAGLTVTSVQLVSSH-NIPVILLEKSNLIGNSIKASSINGAGTSTQVQKHHVIE--DSELPFLQD SP
s	Kwal_27.10088 -----MNSKRVLTFALLCFCLVYLRSSKTSPTS--DMTSKPFST--PVVIGSGLAGLTVTSVQLVSSH-NIPVILLEKSNLIGNSIKASSINGAGTSTQVQKHHVIE--DSELPFLQD SP
s	SAKL0D12056g -----MLQRHKLIVLLLFVLIILYRQST-----QMTTKFTA--PVVIGSGLAGLTVTSVQLVSSH-NIPVILLEKSNLIGNSIKASSINGAGTSTQVQKHHVIE--DSELPFLQD SP
f	Kpo1_1059.11 -----MLKVRRLIVYLIIGFLLYAIRQWSQFPT--QMTSK--K--PVVIGTGLAGLSAGNQLVSSH-NIPVILLEKSNLIGNSIKASSINGAGTSTQVQKHHVIE--DSELPFLQD SP
f	TDEL0C02490 -----MLKVRRLIVYLIIGFLLYAIRQWSQFPT--QMTSK--L--PVVIGSGLAGLTVTSVQLVSSH-NIPVILLEKSNLIGNSIKASSINGAGTSTQVQKHHVIE--DSELPFLQD SP
f	TPHA0F03420 -----MSSR--K--PVVIGTGLAGLSAGNQLVSSH-NIPVILLEKSNLIGNSIKASSINGAGTSTQVQKHHVIE--DSELPFLQD SP
f	TBLA0D03810 -----MLATM--S--PVVIGTGLAGLTVTSVQLVSSH-NIPVILLEKSNLIGNSIKASSINGAGTSTQVQKHHVIE--DSELPFLQD SP
f	CAGL0L01177g -----M--ATVIGTGLAGLTVTSVQLVSSH-NIPVILLEKSNLIGNSIKASSINGAGTSTQVQKHHVIE--DSELPFLQD NC
f	Suva_5.23 -----MSS--S--PVVIGTGLAGLTVTSVQLVSSH-NIPVILLEKSNLIGNSIKASSINGAGTSTQVQKHHVIE--DSELPFLQD SS
f	Skud_5.43 -----MSSS--S--PVVIGTGLAGLTVTSVQLVSSH-NIPVILLEKSNLIGNSIKASSINGAGTSTQVQKHHVIE--DSELPFLQD SS
f	Smik_5.46 -----MSS--S--PVVIGTGLAGLTVTSVQLVSSH-NIPVILLEKSNLIGNSIKASSINGAGTSTQVQKHHVIE--DSELPFLQD SS
f	YEL047C (Frd1) -----MSL--S--PVVIGTGLAGLTVTSVQLVSSH-NIPVILLEKSNLIGNSIKASSINGAGTSTQVQKHHVIE--DSELPFLQD SS
f	KNAG0G00400 -----MSK--D--SVVIGSGLAGLTVTSVQLVSSH-NIPVILLEKSNLIGNSIKASSINGAGTSTQVQKHHVIE--DSELPFLQD SS
f	KAFR0E00440 -----M--S--AVVIGSGLAGLTVTSVQLVSSH-NIPVILLEKSNLIGNSIKASSINGAGTSTQVQKHHVIE--DSELPFLQD SS
f	NDAI0G00250 -----MASTITTDLKNLSIS--E--PVVIGSGLAGLTVTSVQLVSSH-NIPVILLEKSNLIGNSIKASSINGAGTSTQVQKHHVIE--DSELPFLQD NC
f	NCAS0G04200 -----MSNESLATNVEKLNISSD--S--PVVIGSGLAGLTVTSVQLVSSH-NIPVILLEKSNLIGNSIKASSINGAGTSTQVQKHHVIE--DSELPFLQD NC

Fig. 4-S5. Multiple sequence alignment of fumarate reductase homologs. (A) The N-terminal region of the multiple sequence alignment of Frd1/Osm1 homologs from the Yeast Gene Order Browser is shown. Homology to Osm1/Frd1 is indicated in the leftmost column. Signal predictions by Phobius are shown on the right.

Supplemental tables

Table 4-1. Yeast strains

name	description	background	MAT	genotype
y1	Rpl16a/b AVI + Om45-mVenus-BirA	BY4741	a	<i>om45::OM45-mVenus-BirA::HIS5 rpl16a::RPL16a-HA-TEV-AVI rpl16b::RPL16b-HA-TEV-AVI ura3Δ0 met15Δ0 his3Δ1 leu2Δ0</i>
y2	Rpl16a/b AVI + Om45-mVenus-BirA + Su9-tagBFP	BY4741	a	<i>om45::OM45-mVenus-BirA::HIS5 rpl16a::RPL16a-HA-TEV-AVI rpl16b::RPL16b-HA-TEV-AVI p1(pRS316-pTDH3-su9-TagBFP) ura3Δ0 met15Δ0 his3Δ1 leu2Δ0</i>
y3	Rps2 AVI + Om45-mVenus-BirA	BY4741	a	<i>om45::OM45-mVenus-BirA::HIS5 rps2::RPS2-HA-TEV-AVI leu2::PGK1p-BirA-LEU2 ura3Δ0 met15Δ0 his3Δ1</i>
y4	Hap1-yoEGFP	W303-1B	α	<i>hap1::HAP1-yoEGFP::HIS5 leu2-3,112 trp1-1 can1-100 ura3</i>
y5	Hap1-yoEGFP; Su9-BFP	W303	α/a	<i>HAP1/hap1::HAP1-yoEGFP::HIS5 leu2-3,112/leu2-3,112 trp1-1/trp1-1 can1-100/can1-100 ura3/ura3 p1(pRS316-pTDH3-su9-TagBFP)</i>
y6	osm1-3xFLAG	BY4741	a	<i>osm1::OSM1-3xFLAG::KanMX leu2Δ0 ura3Δ0 met15Δ0 his3Δ1</i>
y7	WT OSM1-GFP	W303-1B	α	<i>osm1::OSM1-yoEGFP::LEU2 sec63::sec63-mCherry-BirA::HIS5 p1(pRS316-pTDH3-su9-TagBFP) leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15</i>
y8	M1A OSM1-GFP	W303-1B	α	<i>osm1::OSM1(M1A)-yoEGFP::LEU2 sec63::sec63-mCherry-BirA::HIS5 p1(pRS316-pTDH3-su9-TagBFP) leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15</i>
y9	M32A OSM1-GFP	W303-1B	α	<i>osm1::OSM1(M32A)-yoEGFP::LEU2 sec63::sec63-mCherry-BirA::HIS5 p1(pRS316-pTDH3-su9-TagBFP) leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15</i>

Table 4-2. Plasmids

name	description
p1	pRS316-pTDH3-Su9-TagBFP
p2	pCR2.1 linker-tev-linker-3X-FLAG::KanMX
p3	pFA6a Osm1-yoEGFP::LEU2
p4	pFA6a Osm1(M1A)-yoEGFP::LEU2
p5	pFA6a Osm1(M32A)-yoEGFP::LEU2
p6	pFA6a -mCherry-BirA::HIS5

Table 4-3. Primers

name	purpose	sequence (5' to 3')
oGAB11	28mer size standard	agucacuuagcgauguacacugacugug/3phos/
oCJ200	Primer for reverse transcription of sequencing libraries	5'/5phos/GATCGTCCGACTGTAGAACTCTGAACCTGTCG/i sp18/CAAGCAGAAGACGGCATAACGATATTGATGGTGC CTACAG
oNTI231	Amplification of sequencing libraries, paired with indexing primer	caagcagaagacggcatacga
indexing-primer	Amplification of sequencing libraries, NNNNNN = barcode position	aatgatacggcgaccaccgagatctacacgatcgaagagcacacgtctgaactcc agtcacNNNNNNcgacagggttcagagttc
Linker 1	3' Cloning adaptor for RNA footprints	appctgtaggcaccatcaat/3ddc
oCW102	C-terminally tag OM45 with -mVenus-BirA::HIS5, fwd	tgataagggtgatgtaattctggagctcgaaaaaggacATCGGTGACGG TGCTGGTTT
oCW103	C-terminally tag OM45 with -mVenus-BirA::HIS5, reverse primer	gagaacatgtgaatatgtatatgttatcggggaaccaTGGATGGCGGC GTTAGTATC
oCJ491	OSM1 C-TERM 3XFLAG FWD	CTT TGG AAA GAC AGC TGC GGA TAA CAT AGC AAA

		ATT GTA CGG AGG TTC AGG TTC AGG
oCJ492	OSM1 C-TERM 3XFLAG REV	CGT TGC AAA AAT ATC TTT TAT TTC CCT ATA AAT TCT CAA TCG ATG AAT TCG AGC TCG
oCW500	OSM1 C-term EGFP (WT, M32A) fwd	taccttaataattctatatcatcccagcttaggaaaATGATTAGATCTGT GAGAAGGG
oCW501	OSM1 C-term EGFP (M1A, ATG-->GCT) fwd	taccttaataattctatatcatcccagcttaggaaaGCTATTAGATCTGT GAGAAGGG
oCW503	OSM1 C-term EGFP (WT, M1A, M32A) rev	agtacgttgcaaaaatcttttatttccctataaattcATGCGAATGCACAA GCTAATA
oCJ486	HAP1 C-term GFP tag fwd	ACTAGTAGATTTTTATAGAGCAGATTTTCCAATATGGGAG ggtgacggtgctggttta
oCJ487	HAP1 C-term GFP tag rev	TAAAAATATGTCCCTTCTATTATACATAAAACAAACCGCA tcgatgaattcgagctcg

REFERENCES

- Akopian, D., Shen, K., Zhang, X., and Shan, S. (2013). Signal Recognition Particle: An Essential Protein-Targeting Machine. *Annu. Rev. Biochem.* 82, 693–721.
- Alamo, M. del, Hogan, D.J., Pechmann, S., Albanese, V., Brown, P.O., and Frydman, J. (2011). Defining the Specificity of Cotranslationally Acting Chaperones by Systematic Analysis of mRNAs Associated with Ribosome-Nascent Chain Complexes. *PLoS Biol* 9, e1001100.
- Ast, T., Cohen, G., and Schuldiner, M. (2013). A Network of Cytosolic Factors Targets SRP-Independent Proteins to the Endoplasmic Reticulum. *Cell* 152, 1134–1145.
- Byrne, K.P., and Wolfe, K.H. (2005). The Yeast Gene Order Browser: Combining curated homology and syntenic context reveals gene fate in polyploid species. *Genome Res.* 15, 1456–1461.
- Chatton, B., Walter, P., Ebel, J.P., Lacroute, F., and Fasiolo, F. (1988). The yeast VAS1 gene encodes both mitochondrial and cytoplasmic valyl-tRNA synthetases. *J. Biol. Chem.* 263, 52–57.
- Cherry, J.M., Hong, E.L., Amundsen, C., Balakrishnan, R., Binkley, G., Chan, E.T., Christie, K.R., Costanzo, M.C., Dwight, S.S., Engel, S.R., et al. (2012). *Saccharomyces* Genome Database: the genomics resource of budding yeast. *Nucleic Acids Res.* 40, D700–705.
- Claros, M.G., and Vincens, P. (1996). Computational method to predict mitochondrially imported proteins and their targeting sequences. *Eur. J. Biochem. FEBS* 241, 779–786.
- Costanzo, M., Baryshnikova, A., Bellay, J., Kim, Y., Spear, E.D., Sevier, C.S., Ding, H., Koh, J.L.Y., Toufighi, K., Mostafavi, S., et al. (2010). The genetic landscape of a cell. *Science* 327, 425–431.
- Ellis, S.R., Hopper, A.K., and Martin, N.C. (1989). Amino-terminal extension generated from an upstream AUG codon increases the efficiency of mitochondrial import of yeast N2,N2-dimethylguanosine-specific tRNA methyltransferases. *Mol. Cell. Biol.* 9, 1611–1620.
- Elstner, M., Andreoli, C., Klopstock, T., Meitinger, T., and Prokisch, H. (2009). The mitochondrial proteome database: MitoP2. *Methods Enzymol.* 457, 3–20.
- Enomoto, K., Arikawa, Y., and Muratsubaki, H. (2002). Physiological role of soluble fumarate reductase in redox balancing during anaerobiosis in *Saccharomyces cerevisiae*. *FEMS Microbiol. Lett.* 215, 103–108.
- Fox, T.D. (2012). Mitochondrial Protein Synthesis, Import, and Assembly. *Genetics* 192, 1203–1234.
- Gross, E., Sevier, C.S., Heldman, N., Vitu, E., Bentzur, M., Kaiser, C.A., Thorpe, C., and Fass, D. (2006). Generating disulfides enzymatically: reaction products and electron acceptors of the endoplasmic reticulum thiol oxidase Ero1p. *Proc. Natl. Acad. Sci. U. S. A.* 103, 299–304.

- Guarente, L., Lalonde, B., Gifford, P., and Alani, E. (1984). Distinctly regulated tandem upstream activation sites mediate catabolite repression of the *CYC1* gene of *S. cerevisiae*. *Cell* 36, 503–511.
- Huh, W.-K., Falvo, J.V., Gerke, L.C., Carroll, A.S., Howson, R.W., Weissman, J.S., and O’Shea, E.K. (2003). Global analysis of protein localization in budding yeast. *Nature* 425, 686–691.
- Jan, C.H., Williams, C.C., and Weissman, J.S. (2014). Principles of co-translational translocation at the endoplasmic reticulum revealed by proximity-specific ribosome profiling. *Science* (submitted).
- Juneau, K., Nislow, C., and Davis, R.W. (2009). Alternative splicing of *PTC7* in *Saccharomyces cerevisiae* determines protein localization. *Genetics* 183, 185–194.
- Kellems, R.E., Allison, V.F., and Butow, R.A. (1975). Cytoplasmic type 80S ribosomes associated with yeast mitochondria. IV. Attachment of ribosomes to the outer membrane of isolated mitochondria. *J. Cell Biol.* 65, 1–14.
- Kornmann, B., Currie, E., Collins, S.R., Schuldiner, M., Nunnari, J., Weissman, J.S., and Walter, P. (2009). An ER-Mitochondria Tethering Complex Revealed by a Synthetic Biology Screen. *Science* 325, 477–481.
- Lee, S., Lim, W.A., and Thorn, K.S. (2013). Improved Blue, Green, and Red Fluorescent Protein Tagging Vectors for *S. cerevisiae*. *PLoS ONE* 8, e67902.
- Liu, Z., Österlund, T., Hou, J., Petranovic, D., and Nielsen, J. (2013). Anaerobic α -Amylase Production and Secretion with Fumarate as the Final Electron Acceptor in *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* 79, 2962–2967.
- Marc, P., Margeot, A., Devaux, F., Blugeon, C., Corral-Debrinski, M., and Jacq, C. (2002). Genome-wide analysis of mRNAs targeted to yeast mitochondria. *EMBO Rep.* 3, 159–164.
- Muratsubaki, H., and Enomoto, K. (1998). One of the Fumarate Reductase Isoenzymes from *Saccharomyces cerevisiae* Encoded by the *OSM1* Gene. *Arch. Biochem. Biophys.* 352, 175–181.
- Nargund, A.M., Pellegrino, M.W., Fiorese, C.J., Baker, B.M., and Haynes, C.M. (2012). Mitochondrial Import Efficiency of *ATFS-1* Regulates Mitochondrial UPR Activation. *Science* 337, 587–590.
- Okreglak, V., and Walter, P. (2014). The conserved AAA-ATPase *Msp1* confers organelle specificity to tail-anchored proteins. *Proc. Natl. Acad. Sci.* 201405755.
- Saint-Georges, Y., Garcia, M., Delaveau, T., Jourden, L., Le Crom, S., Lemoine, S., Tanty, V., Devaux, F., and Jacq, C. (2008). Yeast Mitochondrial Biogenesis: A Role for the PUF RNA-Binding Protein *Puf3p* in mRNA Localization. *PLoS ONE* 3, e2293.
- Saraogi, I., and Shan, S. (2011). Molecular mechanism of co-translational protein targeting by the signal recognition particle. *Traffic Cph. Den.* 12, 535–542.

Stern-Ginossar, N., Weisburd, B., Michalski, A., Le, V.T.K., Hein, M.Y., Huang, S.-X., Ma, M., Shen, B., Qian, S.-B., Hengel, H., et al. (2012). Decoding human cytomegalovirus. *Science* 338, 1088–1093.

Suissa, M., and Schatz, G. (1982). Import of proteins into mitochondria. Translatable mRNAs for imported mitochondrial proteins are present in free as well as mitochondria-bound cytoplasmic polysomes. *J. Biol. Chem.* 257, 13048–13055.

Thomson, J.M., Gaucher, E.A., Burgan, M.F., De Kee, D.W., Li, T., Aris, J.P., and Benner, S.A. (2005). Resurrecting ancestral alcohol dehydrogenases from yeast. *Nat. Genet.* 37, 630–635.

Tu, B.P., and Weissman, J.S. (2002). The FAD- and O₂-Dependent Reaction Cycle of Ero1-Mediated Oxidative Protein Folding in the Endoplasmic Reticulum. *Mol. Cell* 10, 983–994.

Willmund, F., del Alamo, M., Pechmann, S., Chen, T., Albanèse, V., Dammer, E.B., Peng, J., and Frydman, J. (2013). The Cotranslational Function of Ribosome-Associated Hsp70 in Eukaryotic Protein Homeostasis. *Cell* 152, 196–209.

Wolfe, K. (2000). Robustness--it's not where you think it is. *Nat. Genet.* 25, 3–4.

Zielinska, D.F., Gnad, F., Schropp, K., Wiśniewski, J.R., and Mann, M. (2012). Mapping N-glycosylation sites across seven evolutionarily distant species reveals a divergent substrate proteome despite a common core machinery. *Mol. Cell* 46, 542–548.

CHAPTER FIVE

Discussion and future perspectives

SUMMARY

Massively parallelized sequencing of nucleic acids has emerged over the last decade as a key tool for studying cellular physiology at a molecular level and on an unprecedented scale (Bentley et al., 2008). In addition to sequencing whole genomes, this high-precision technology is increasingly being exploited to characterize specific subpopulations of DNA or RNA for more targeted investigation of genomic activity and gene expression. These approaches have enabled genome-wide profiling of mRNA expression (Adiconis et al., 2013), characterization of protein-DNA (Johnson et al., 2007) and protein-RNA (Ule et al., 2005) interactions, high-precision mapping of 5' (Arribere and Gilbert, 2013) and 3' (Jan et al., 2011) transcript boundaries, identification of non-coding RNAs (Karginov et al., 2007), analysis of epigenetic modifications (Weber et al., 2005), investigation of RNA structure *in vivo* and *in vitro* (Rouskin et al., 2014), capture of RNA polymerase (Churchman and Weissman, 2012) and ribosome (Ingolia et al., 2009) activity at nucleotide resolution, and quantitative readouts for genome-wide sequencing-based screens (Bassik et al., 2009; Carette et al., 2011).

Although these methods have indisputably informed all areas of biology, they almost universally discard cellular spatial relations. This caveat has limited the utility of high-precision sequencing in the context of dynamic spatiotemporal processes, or more generally those involving heterogeneous populations of RNA or DNA. Proximity-specific ribosome profiling constitutes the first generic approach that captures these dimensions of gene expression without the confounding impurities and limitations of biochemical preparations. At a unique interface of genomics and cell biology, it will enable the exploration of a variety of intriguing subcellular and co-translational processes through isolation of specific subpopulations of ribosomes.

TRANSLATION AT THE ER

The application of proximity-specific ribosome profiling to ER translation revealed numerous systems-level principles that complement years of more targeted studies of single pathways in insolation. While these insights enable a more holistic understanding of co-translational translocation at the ER, they also motivate additional investigations.

The surprising prevalence of co-translational translocation at the ER and the purported contribution of mRNA tethering to its efficiency (Jan et al., 2014) make ER-localized RBPs appealing candidates for further study. Preliminary proximity-specific ribosome profiling experiments carried out in yeast lacking either *BFR1* or *SCP160* (Frey et al., 2001; Lang and Fridovich-Keil, 2000) suggest a minimal contribution of these RBPs to ER translation overall (Fig. 5-1). However, specific mRNAs have been reported to co-migrate with cortical ER for preferential localization to the yeast bud during mitosis (Schmid et al., 2006). Thus, it is possible that the substrates of these factors localize to and traffic with the ER, but are not translocated and are therefore not detected in our assay. However, our ability to robustly detect translation of a non-secreted reporter mRNA, artificially tethered to the ER using the PP7 coat protein and aptamer system (unpublished data), argues against this possibility. This observation also raises the question of why specific sets of non-secretory mRNAs deriving from several cellular compartments are detected at the ER in a CHX-dependent manner (Fig. 3-3C, 5-2). Although cytosolic mRNAs have been observed to biochemically fractionate with the ER (Kraut-Cohen and Gerst, 2010), the functional significance and mechanism of this localization remain unclear.

In yeast, proximity-specific ribosome profiling enabled insights into the functional roles of distinct translocon complexes, which have been difficult to elucidate by other methods. An obvious future application of translocon-specific ribosome profiling is toward the functional characterization of translocon accessory factors in mammals. Although it is

generally agreed that factors, such as TRAM and TRAP, increase translocation efficiency in a substrate-specific manner (Fons, 2003; Görlich et al., 1992), which features they recognize and how they promote translocation mechanistically remain unknown. Comprehensively defining which substrates depend on them for robust import, determining when during translocation they engage nascent proteins, and evaluating the extent to which these factors contribute to ER translation more generally will undoubtedly inform our understanding of translocational regulation (Hegde and Kang, 2008) in higher eukaryotes.

TRANSLATION AT THE MITOCHONDRIAL OUTER MEMBRANE

Proximity-specific ribosome profiling provided direct evidence that mitochondrial import occurs co-translationally for proteins of the IM in yeast, providing a rationale for previous observations of non-universal association of mitochondrial mRNAs with the OM (Fox, 2012). Determining why this class of proteins is preferentially imported co-translationally, and whether this preference is maintained in other cell types, will be important in fully evaluating the significance of the observation.

Comparison of protein targeting to the mitochondria and ER emphasized the utility of proximity-specific ribosome profiling in cases where the inability to biochemically resolve subcellular compartments has confounded previous studies. Despite the seemingly degenerate hydrophobic targeting signals that direct partitioning of substrates between these organelles, our spatial proteomic maps emphasize the fidelity of co-translational trafficking achieved *in vivo*. Intriguingly, the observed interplay of substrates between ribosome-associated factors, such as SRP and NAC, suggests an important role for co-translational processes in establishing this specificity (Fig. 5-3) (Alamo et al., 2011). Characterizing how these targeting factors, in combination with mRNA localization mediated by ER- and mitochondrially-localized RBPs, act to achieve this specificity remains an

important cell biological question.

Regulated dual localization of proteins between cellular compartments seems to represent an interesting exception to this observed targeting specificity. Understanding the rules the cell uses to direct simultaneous localization to distinct compartments, particularly in the context of proteins containing single ambiguous targeting signals, will inform our understanding of intracellular sorting and communication. Importantly, proximity-specific ribosome profiling sensitively detected dual localization events (Fig. 4-4A, 4-4B, 4-S4) and future applications of the approach to additional subcellular sites will also enable more systematic analyses of this phenomenon.

ISOLATION OF RIBOSOME SUBPOPULATIONS

Genetically controllable, proximity-specific biotinylation of ribosomes broadly enables the study of defined subpopulations of ribosomes. In this work, we have applied this labeling approach to study the translational activity of ribosomes localized to two cellular compartments. Exploration of translation at other subcellular locations, such as neuronal processes, will be greatly facilitated by analogous strategies.

In addition to ribosomes that reside in cellular sub-compartments, proximity-specific biotinylation enables the isolation of ribosomes that interact with specific soluble factors, such as chaperones. By extension, monitoring the translational activity of these ribosomes enables precise study of co-translational processes (Oh et al., 2011). Proximity-specific biotinylation represents a more generic strategy for the isolation of factor-associated ribosomes compared to other approaches that require factor-specific cross-linking optimization (Oh et al., 2011). Preliminary ribosome profiling experiments with BirA-tagged Srp72 in yeast demonstrate the feasibility of labeling with ribosome-associated soluble factors. The specificity of labeling by this BirA is demonstrated by the enrichment of secreted

genes (Fig 5-4A). Additionally, the peak positional enrichment observed for secreted genes with Srp72-BirA occurred at ~55 codons (Fig. 5-3B), the expected position of engagement immediately preceding the peak enrichment at ~60 codons with BirA-Ssh1, the dedicated SRP translocon (Fig. 3-4B). Intriguingly, comparable enrichment was observed for the set of secretory proteins containing internal TMDs suggesting that SRP does not have a strict requirement for N-terminal signals (Fig. 5-4B). Because tail-anchored and cytosolic proteins do not exhibit this enrichment (Fig. 5-4B) it does not reflect universal sampling of nascent chains by SRP as has been suggested previously (Ogg and Walter, 1995). It is possible that the specificity of ribosome labeling by other BirA-tagged soluble factors will be sensitive to factor expression levels (Fernández-Suárez et al., 2008). However, it may be possible to reduce non-specific labeling by tuning BirA activity or the AviTag through mutations, as discussed in Chapter II.

Proximity-specific biotinylation also enables robust isolation of ribosomes containing a specific (Avi-tagged) ribosomal protein of interest through constitutive labeling by a cytosolic BirA. This labeling strategy would be particularly useful in the context of 'specialized ribosomes' to directly evaluate the effects of heterogeneous ribosome composition on translational activity (Xue and Barna, 2012). Through extension to whole organisms, this labeling approach could be used to isolate ribosomes in a tissue-, cell type-, or tumor-specific manner by expressing BirA in subpopulations of cells. The cellular heterogeneity of tissues, in particular the brain, has confounded efforts to study biological processes in genetically defined cell populations. Previously, translating ribosome affinity purification (TRAP) has been used to purify whole transcripts from specific cell types expressing an epitope-tagged ribosome (Heiman et al., 2008). The nucleotide resolution and the high-affinity biotin-streptavidin interaction inherent to our system may yield more specific and quantitative measurements of translation.

Finally, in addition to monitoring translational activity by ribosome profiling, one could apply mass spectrometry to identify novel ribosome-associated factors, evaluate ribosome composition, and characterize post-translational modifications specific to ribosome subpopulations. These data would be highly complementary to the ribosome profiling readout, and would more broadly inform our understanding of translational control within cells and organisms.

FIGURES

Figure 5-1

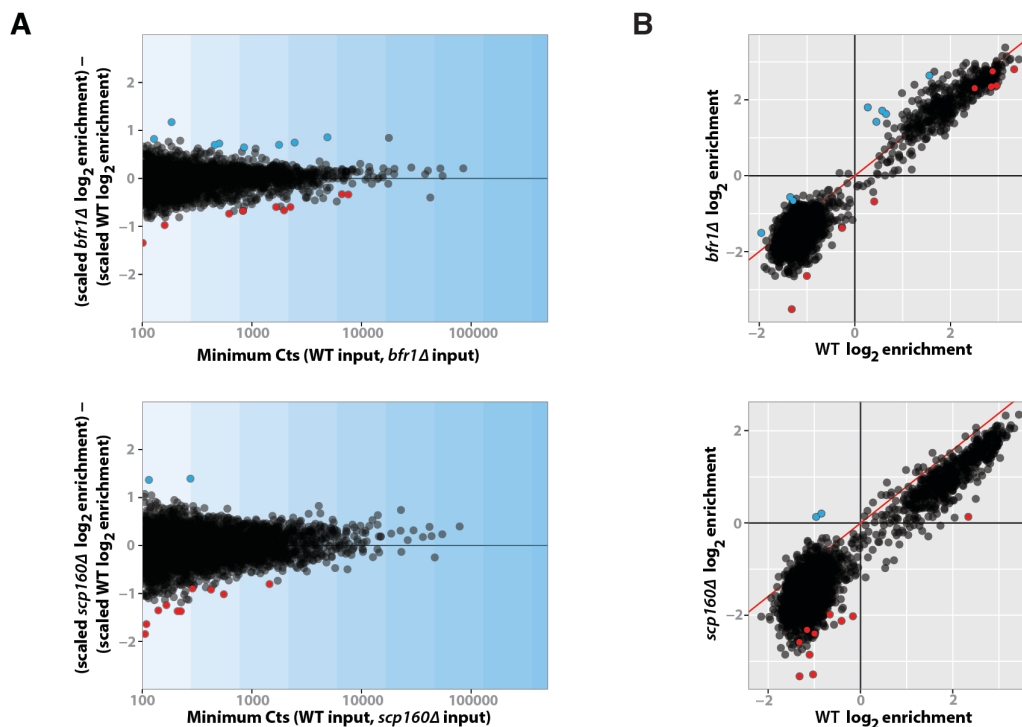


Fig. 5-1. Minimal effects of candidate ER RNA binding protein disruption on gene-level ER enrichment. (A) Plotted are the differences in scaled (to zero-mean and unit variance) log₂ ER enrichments for *bfr1Δ* (top) or *scp160Δ* (bottom) libraries minus wildtype for each gene as a function of minimum sequencing counts between the wildtype or deletion libraries. Genes were divided into nine logarithmically spaced bins (shaded blue vertical bars) and genes beyond three standard deviations from the mean in each bin were considered significant outliers. Positive and negative outliers are plotted as blue and red circles, respectively. (B) log₂ ER enrichments for *bfr1Δ* (top) or *scp160Δ* (bottom) libraries versus WT. Outlier genes as defined in (A) are highlighted. All proximity-specific ribosome profiling experiments were carried out with BirA-mVenus-Ubc6, using two minute biotin pulses in the presence of CHX.

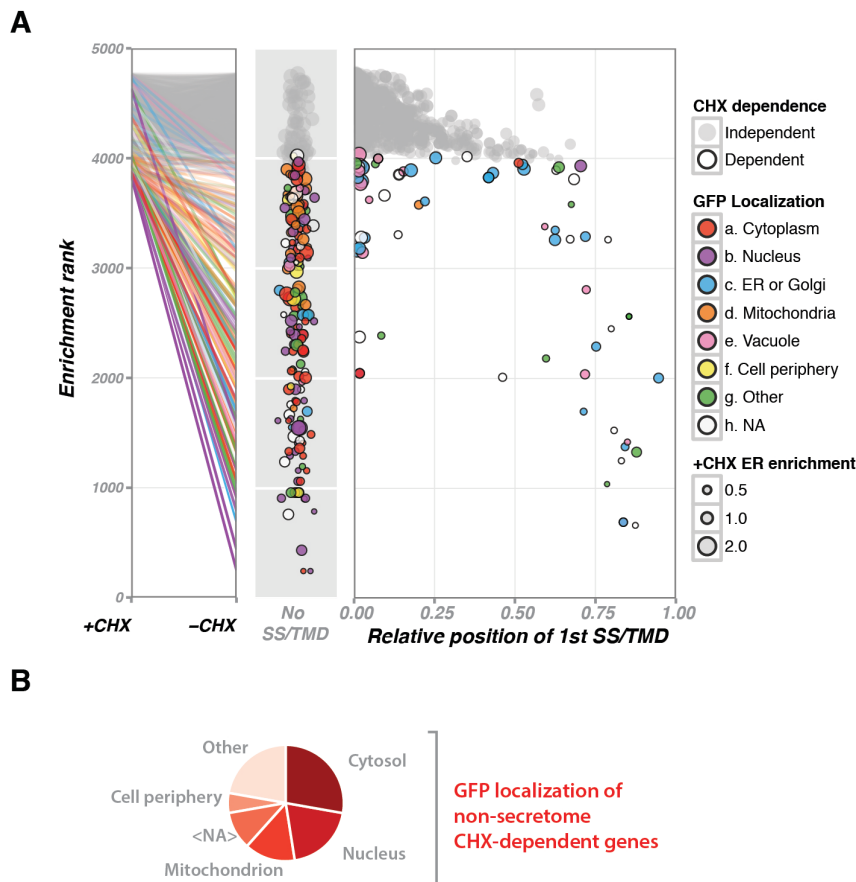


Fig. 5-2. Properties of CHX-dependent genes. (A) Enrichment, enrichment rank, GFP localization, and predicted hydrophobic domain properties of CHX-dependent (colored) and -independent (grey) genes. Each line in the parallel coordinates rank panel displays the gene-level log₂ enrichment ranking of a single gene in the presence or absence of CHX. Each line has a corresponding matched-color point in the neighboring scatter plots whose y-axes also match the -CHX enrichment rankings. Genes are categorically split between scatter plot panels based on whether they are predicted to contain a hydrophobic domain (right) or not (middle). The x-axis of the right scatter plot panel represents the position of the start of the first hydrophobic element relative to the overall gene length. Colors indicate the annotated GFP localization of genes and the size of points indicates the magnitude of the log₂ gene-level enrichment in the presence of CHX. **(B)** Pie chart summarizing the available GFP localization data for CHX-dependent genes not predicted to contain signal sequences or TMDs. GFP localization data taken from the *Saccharomyces* Genome Database (SGD) (Cherry et al., 2012).

Figure 5-3

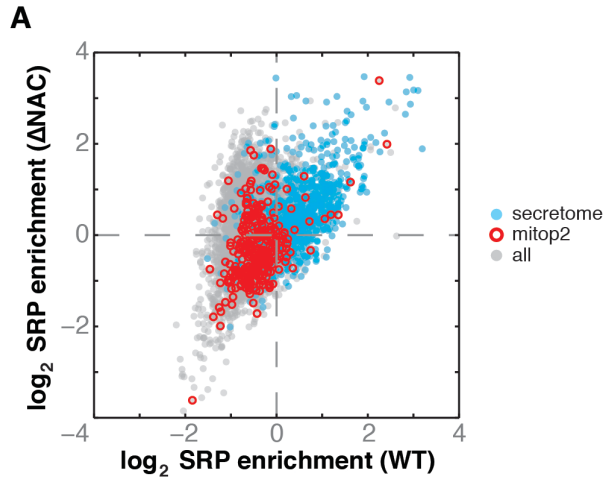


Fig. 5-3. Ribosome-associated factors help maintain ER and mitochondrial targeting specificity.

(A) Translational profile of SRP-bound ribosome-nascent chain complexes in WT or NAC-deleted cells for secretory (blue), mitop2 mitochondrial (red), and other (grey) genes. log₂ enrichments taken from (Alamo et al., 2011).

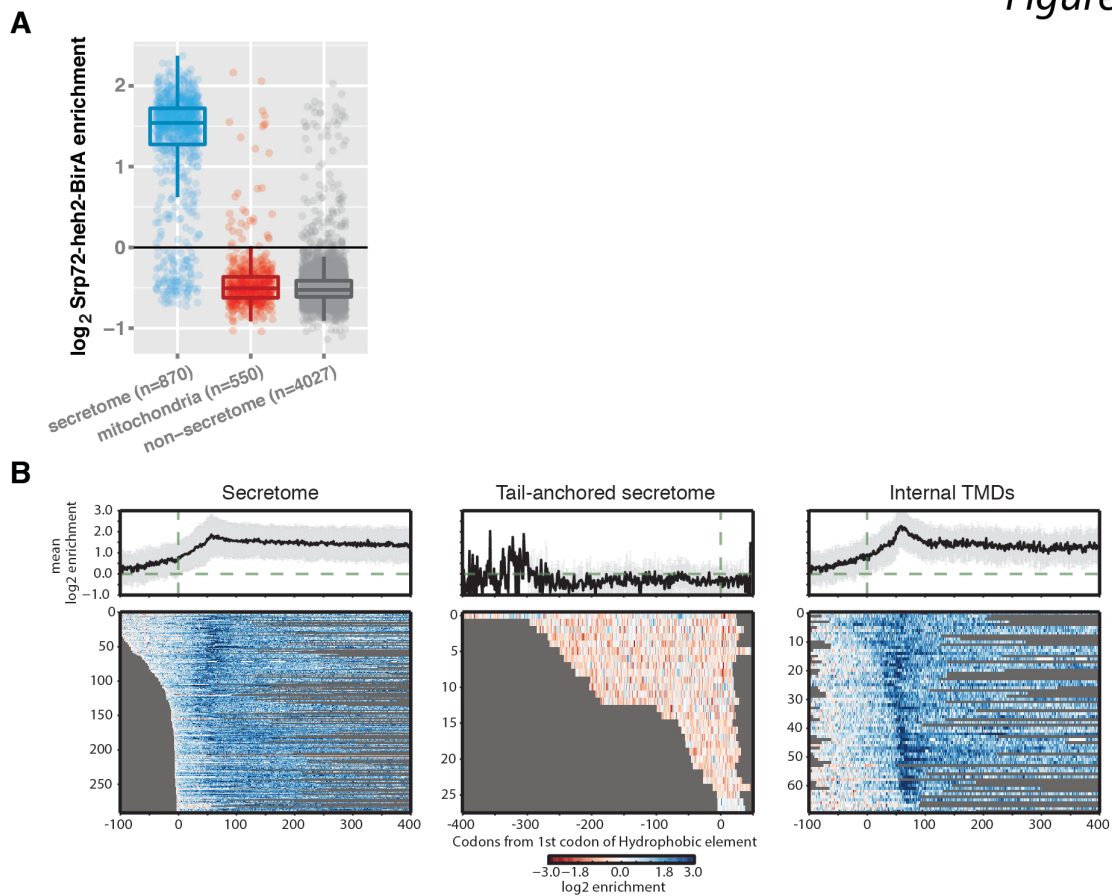


Fig. 5-4. Proximity-specific ribosome profiling with a soluble factor: Srp72-heh2-BirA. (A) Boxplot of \log_2 enrichment for secretome (blue), mitop2 mitochondrial (red), and non-secretome (grey) gene categories from proximity-specific ribosome profiling experiments in yeast using Srp72-heh2-BirA. ORFs annotated as dubious, did not have 100 genome-mapping reads in the input sample, or which overlap other ORFs on the same strand, were filtered. Biotinylation was performed in the presence of CHX for two minutes. (B) Meta-gene plots of \log_2 Srp72-heh2-BirA enrichment per codon as a function of ribosome position relative to the first codon of the first Phobius-predicted hydrophobic element for the indicated signal class. Heat maps below represent single-gene positional enrichments used to derive the corresponding averaged meta-gene plot.

REFERENCES

- Adiconis, X., Borges-Rivera, D., Satija, R., DeLuca, D.S., Busby, M.A., Berlin, A.M., Sivachenko, A., Thompson, D.A., Wysoker, A., Fennell, T., et al. (2013). Comparative analysis of RNA sequencing methods for degraded or low-input samples. *Nat. Methods* 10, 623–629.
- Alamo, M. del, Hogan, D.J., Pechmann, S., Albanese, V., Brown, P.O., and Frydman, J. (2011). Defining the Specificity of Cotranslationally Acting Chaperones by Systematic Analysis of mRNAs Associated with Ribosome-Nascent Chain Complexes. *PLoS Biol* 9, e1001100.
- Arribere, J.A., and Gilbert, W.V. (2013). Roles for transcript leaders in translation and mRNA decay revealed by transcript leader sequencing. *Genome Res.* 23, 977–987.
- Bassik, M.C., Lebbink, R.J., Churchman, L.S., Ingolia, N.T., Patena, W., LeProust, E.M., Schuldiner, M., Weissman, J.S., and McManus, M.T. (2009). Rapid creation and quantitative monitoring of high coverage shRNA libraries. *Nat. Methods* 6, 443–445.
- Bentley, D.R., Balasubramanian, S., Swerdlow, H.P., Smith, G.P., Milton, J., Brown, C.G., Hall, K.P., Evers, D.J., Barnes, C.L., Bignell, H.R., et al. (2008). Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* 456, 53–59.
- Carette, J.E., Guimaraes, C.P., Wuethrich, I., Blomen, V.A., Varadarajan, M., Sun, C., Bell, G., Yuan, B., Muellner, M.K., Nijman, S.M., et al. (2011). Global gene disruption in human cells to assign genes to phenotypes by deep sequencing. *Nat. Biotechnol.* 29, 542–546.
- Churchman, L.S., and Weissman, J.S. (2012). Native elongating transcript sequencing (NET-seq). *Curr. Protoc. Mol. Biol.* Ed. Frederick M Ausubel *Chapter 4*, Unit 4.14.1–17.
- Fernández-Suárez, M., Chen, T.S., and Ting, A.Y. (2008). Protein–Protein Interaction Detection in Vitro and in Cells by Proximity Biotinylation. *J. Am. Chem. Soc.* 130, 9251–9253.
- Fons, R.D. (2003). Substrate-specific function of the translocon-associated protein complex during translocation across the ER membrane. *J. Cell Biol.* 160, 529–539.
- Fox, T.D. (2012). Mitochondrial Protein Synthesis, Import, and Assembly. *Genetics* 192, 1203–1234.
- Frey, S., Pool, M., and Seedorf, M. (2001). Scp160p, an RNA-binding, polysome-associated protein, localizes to the endoplasmic reticulum of *Saccharomyces cerevisiae* in a microtubule-dependent manner. *J. Biol. Chem.* 276, 15905–15912.
- Görlich, D., Hartmann, E., Prehn, S., and Rapoport, T.A. (1992). A protein of the endoplasmic reticulum involved early in polypeptide translocation. *Nature* 357, 47–52.
- Hegde, R.S., and Kang, S.-W. (2008). The concept of translocational regulation. *J. Cell Biol.* 182, 225–232.

Heiman, M., Schaefer, A., Gong, S., Peterson, J.D., Day, M., Ramsey, K.E., Suárez-Fariñas, M., Schwarz, C., Stephan, D.A., Surmeier, D.J., et al. (2008). A Translational Profiling Approach for the Molecular Characterization of CNS Cell Types. *Cell* 135, 738–748.

Ingolia, N.T., Ghaemmaghami, S., Newman, J.R.S., and Weissman, J.S. (2009). Genome-Wide Analysis in Vivo of Translation with Nucleotide Resolution Using Ribosome Profiling. *Science* 324, 218–223.

Jan, C.H., Friedman, R.C., Ruby, J.G., and Bartel, D.P. (2011). Formation, regulation and evolution of *Caenorhabditis elegans* 3'UTRs. *Nature* 469, 97–101.

Jan, C.H., Williams, C.C., and Weissman, J.S. (2014). Principles of co-translational translocation at the endoplasmic reticulum revealed by proximity-specific ribosome profiling. *Science* (submitted).

Johnson, D.S., Mortazavi, A., Myers, R.M., and Wold, B. (2007). Genome-wide mapping of in vivo protein-DNA interactions. *Science* 316, 1497–1502.

Karginov, F.V., Conaco, C., Xuan, Z., Schmidt, B.H., Parker, J.S., Mandel, G., and Hannon, G.J. (2007). A biochemical approach to identifying microRNA targets. *Proc. Natl. Acad. Sci. U. S. A.* 104, 19291–19296.

Kraut-Cohen, J., and Gerst, J.E. (2010). Addressing mRNAs to the ER: cis sequences act up! *Trends Biochem. Sci.* 35, 459–469.

Lang, B.D., and Fridovich-Keil, J.L. (2000). Scp160p, a multiple KH-domain protein, is a component of mRNP complexes in yeast. *Nucleic Acids Res.* 28, 1576–1584.

Ogg, S.C., and Walter, P. (1995). SRP samples nascent chains for the presence of signal sequences by interacting with ribosomes at a discrete step during translation elongation. *Cell* 81, 1075–1084.

Oh, E., Becker, A.H., Sandikci, A., Huber, D., Chaba, R., Gloge, F., Nichols, R.J., Typas, A., Gross, C.A., Kramer, G., et al. (2011). Selective ribosome profiling reveals the cotranslational chaperone action of trigger factor in vivo. *Cell* 147, 1295–1308.

Rouskin, S., Zubradt, M., Washietl, S., Kellis, M., and Weissman, J.S. (2014). Genome-wide probing of RNA structure reveals active unfolding of mRNA structures in vivo. *Nature* 505, 701–705.

Schmid, M., Jaedicke, A., Du, T.-G., and Jansen, R.-P. (2006). Coordination of endoplasmic reticulum and mRNA localization to the yeast bud. *Curr. Biol. CB* 16, 1538–1543.

Ule, J., Jensen, K., Mele, A., and Darnell, R.B. (2005). CLIP: a method for identifying protein-RNA interaction sites in living cells. *Methods San Diego Calif* 37, 376–386.

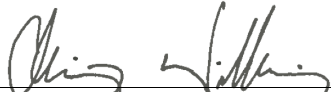
Weber, M., Davies, J.J., Wittig, D., Oakeley, E.J., Haase, M., Lam, W.L., and Schübeler, D. (2005). Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nat. Genet.* 37, 853–862.

Xue, S., and Barna, M. (2012). Specialized ribosomes: a new frontier in gene regulation and organismal biology. *Nat. Rev. Mol. Cell Biol.* 13, 355–369.

PUBLISHING AGREEMENT

It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.



Author Signature

08/20/2014

Date