Distant Chaperones and *N*-Glycan Signals:
New Mechanisms of Secretory Pathway Proteostasis

by

Madeline Y. Wong

B.S., Yale University (2013)

Submitted to the Department of Chemistry
in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Chemistry

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2018

Signature redacted

Signature of Author _____
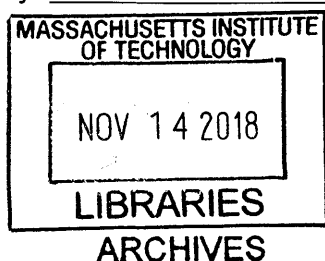Department of Chemistry
August 27, 2018

Signature redacted

Certified by _____
Matthew D. Shoulders
Whitehead Career Development Associate Professor of Chemistry
Thesis Supervisor

Signature redacted

Accepted by _____
Robert W. Field
Haslam and Dewey Professor of Chemistry
Chair, Departmental Committee on Graduate Students

This doctoral thesis has been examined by a
committee of the Department of Chemistry as follows:

Signature redacted

_____ $\mathcal{U}$

<div align="right">

Elizabeth M. Nolan
Associate Professor of Chemistry
Thesis Committee Chair

</div>

Signature redacted

_____ ι

<div align="right">

Matthew D. Shoulders
Whitehead Career Development Associate Professor of Chemistry
Thesis Supervisor

</div>

Signature redacted

_____

<div align="right">

Jeremiah A. Johnson
Associate Professor of Chemistry

</div>

Distant Chaperones and *N*-Glycan Signals:
New Mechanisms of Secretory Pathway Proteostasis

by
Madeline Y. Wong

Submitted to the Department of Chemistry
on August 27, 2018 in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy in
Chemistry

Abstract

Approximately one-third of all cellular proteins traverse the secretory pathway. After translation and folding in the endoplasmic reticulum (ER), proteins are transported through the Golgi to their final locations inside or outside of cells. At each step, proteins are helped by chaperones, which both shepherd proteins towards their native structures and serve as gatekeepers for export. Many proteins in the secretory pathway are also modified by installation of polysaccharides on specific asparagine residues. These *N*-glycans are installed in the ER as uniform precursors, but are trimmed and built up by Golgi glycan maturation enzymes into a striking array of epitopes. *N*-glycans act as a second mechanism to stabilize protein structure and prevent the release of misfolded proteins. Outside the cell, *N*-glycans on cell surfaces and secreted, soluble proteins allow cells to interact with each other, with their environment, and with distal tissues.

During development, cells encounter physiological ER stress incurred by high levels of sustained protein production. Unresolved protein misfolding, on the other hand, results in pathological ER stress and tissue dysfunction. Prior work has used small model substrates to show that cells utilize secretory pathway chaperones and tune *N*-glycosylation to respond to ER stress. This thesis examines how cells use similar strategies to accommodate challenging cargoes such as collagen-I. In the human body, collagen-I constitutes the primary protein component of bone, skin, and other organs; collagen-I misfolding results in pathological ER stress and connective tissue diseases. We therefore set out 1) to identify cellular components required for collagen-I secretion that could be targeted to address disease and 2) to assess the effects of ER stress on both cellular *N*-glycan structures and individual glycoproteins.

Here, we employ a high-throughput assay for collagen-I secretion and find that the cytosolic isoform of Hsp90 is required for collagen-I export. We also show that intracellular stress signaling alters the structures of cell surface and secreted *N*-glycans. Finally, we demonstrate that the collagen-I *N*-glycan buffers collagen-I folding against destabilizing mutations and ER stress. Our results identify potential therapeutic leads for collagen misfolding diseases and point to new mechanisms for maintaining secretory pathway proteostasis.

Thesis Supervisor: Matthew D. Shoulders
Title: Whitehead CD Associate Professor

Acknowledgements

Were the whole realm of nature mine,
That were an offering far too small;
Love so amazing, so divine,
Demands my soul, my life, my all.

- Isaac Watts

CONTENTS

CHAPTER I
A Proteostasis Approach to Treating the Collagenopathies

Summary

The collagen-misfolding diseases, or collagenopathies, encompass mutations in more than 31 genes and show clinical phenotypes in most connective tissues. Single mutations that affect the folding of collagen chains can prevent secretion of sufficient protein levels for tissue function; in other cases, mutant chains accumulate in the endoplasmic reticulum (ER), resulting in sustained ER stress that leads to cellular dysfunction or death. The most severe disease phenotypes often stem from a third class of mutants, whose deposition in the extracellular matrix (ECM) disrupts packing of wild-type collagen fibrils or other ECM components. In addition to these disease mechanisms, loss-of-function variants of chaperones required for collagen folding, or mutations in other ECM proteins that share the same secretion pathway, also contribute to chondro- and skeletal dysplasias. Cells employ a transcriptional program called the Unfolded Protein Response (UPR) that is responsible for restoring homeostasis in the ER upon cellular or organismal insult, and targeting each of the three arms of the UPR has shown promise for diverse protein misfolding diseases. Increasing evidence suggests that UPR-focused therapeutic strategies may be able to mitigate the consequences of collagen misfolding and greatly improve patient outlook. Beyond the UPR, targeting individual components of the cellular proteostasis network also holds therapeutic potential. In this chapter, we discuss application of these strategies to the unique proteostatic challenges of collagen misfolding diseases, and review potential directions for building upon recent advances in this area.

15

Dysregulated Proteostasis and the Collagenopathies

Collagen is a strange, surprising protein. It completely lacks enzymatic activity, and yet regulates processes such as receptor signaling, cell migration, and tissue repair.[1] The structure of mature type I collagen (collagen-I) is a simple right-handed triple helix formed from strands of repeating Gly-X-Y amino acid motifs, but relies on specific sequence properties and post-translational modifications to pre-organize and stabilize the triple helix.[2, 3] Among the twenty-eight other types of collagen, the triple helix is subjected to an impressive array of permutations, from the interrupted helices of the FACIT (Fibril Associated Collagens with Interrupted Triple helices) family to the modular, organizing appendages on the termini of network-forming collagens (reviewed in [1] and [4]).[4] Outside the cell, the collagens assemble with themselves, with each other, and with other ECM components to form and maintain diverse tissues.

Unsurprisingly, delays or difficulties in any of these processes—translation, modification, triple helix assembly, secretion, or fibril formation—cause disease. The affected tissues and phenotypes are as varied as the collagens themselves, but many involve factors such as insufficient or excess protein deposition, cellular stress, inflammation, and apoptosis.[5] Because the molecular basis for disease is not always clear, with some collagen genotypes giving rise to multiple distinct phenotypes,[6] one approach to treating the collagenopathies would be to target the network of cellular proteins responsible for collagen biosynthesis and turnover. Focusing on regulators of collagen protein homeostasis (proteostasis) is not unprecedented: over twenty-five years ago, small molecule inhibitors of the collagen prolyl-4-hydroxylase (P4H) enzyme were presented as potential therapeutic compounds for cases of excess collagen deposition, and compounds that selectively inhibit or activate P4H in the absence of general toxicity are now available.[7-9] Moreover, collagen does not contain clear binding sites or chemical handles; thus, targeting the chaperones and modifying enzymes that regulate its production could expand druggable space while simultaneously addressing multiple aspects of proteostatic imbalance (see Section 3.3 of [10]).[10]

The Unfolded Protein Response

Collagen is synthesized and assembled in the ER, which safeguards organellar proteostasis by means of the Unfolded Protein Response (UPR). Similar to the cytosolic heat shock response, the UPR tunes protein folding and degradation machinery to restore ER proteostasis upon accumulation of misfolded proteins. Activation of the three transmembrane sensors IRE1, ATF6, and PERK produces transcription factors XBP1s, ATF6(1-373), and ATF4, respectively, that together increase levels of chaperones, enhance protein degradation, and stall new protein synthesis to allow the ER to recover. If ER stress continues, however, PERK-induced ATF4 activity will increase levels of a protein termed CHOP in a signaling cascade that leads to apoptosis.

The advent of chemical tools for selective, stress-independent activation of individual arms of the UPR[11] has clarified the overlapping functions of each branch, as well as revealed physiological roles for individual UPR transcription factors.[12] XBP1s activity, for instance, has been shown to regulate processes such as memory formation, innate immunity, and cell non-autonomous signaling.[13-15] ATF6 has been less extensively studied, but has been linked to regulation of ER size and mesodermal differentiation.[16, 17] Importantly, selective activation of individual arms of the UPR also mitigates protein misfolding in both gain-of-function and loss-of-function disease models. In particular, ATF6 activation has been shown to reduce secretion and aggregation of amyloid-forming mutants of transthyretin,[18, 19] as well as increase clearance of (toxic) protein aggregates, either through ER-associated degradation by the proteasome or autophagy.[20] XBP1s activation is also able to reduce secretion of amyloidogenic proteins.[19] Intriguingly, proteomics-based approaches to identify components of the collagen-I proteostasis network found that a number of novel interactors are transcriptional targets of XBP1s and/or ATF6.[11, 21] Such findings suggest that modulation of UPR signaling not only may play a role in

16

collagen-I biosynthesis, but also may have positive effects in the collagenopathies; we discuss these possibilities in more detail below.

IRE1-XBP1s

IRE1 is the oldest stress sensor of the UPR, and fulfills dual roles in response to ER stress. Accumulation of misfolded proteins titrates the ER chaperone BiP away from the luminal portion of IRE1, causing oligomerization and auto-phosphorylation by the IRE1 kinase domain.[22] The IRE1 RNase domain subsequently splices XBP1 mRNA to generate active XBP1s (spliced) mRNA, which is translated to give the XBP1s transcription factor. In addition to generating active XBP1s, IRE1 also degrades selective subsets of gene transcripts in a process known as regulated IRE1-dependent decay.[23] XBP1s was previously shown to enable antibody production during B-cell differentiation by expanding the ER and increasing secretory pathway capacity.[24] Given that many fibrillar collagens are too large to fit into standard COP-II vesicles,[25] and that collagen-producing tissues are responsible for secreting large amounts of ECM proteins in addition to collagen, it is plausible that XBP1s facilitates a similar change during cartilage or bone development. Indeed, in addition to regulating the expression of genes involved in collagen-I biosynthesis, the IRE1/XBP1s axis has numerous physiological roles in collagen-producing tissues. IRE1 and XBP1s inhibit ER stress-induced apoptosis in differentiating cartilage.[26] XBP1s also controls the expression of the transcription factor Osterix, required for osteoblast mineralization, in addition to chondrocyte proliferation and hypertrophic growth plate shortening in cartilage maturation.[27, 28]

The effects of IRE1-XBP1s signaling, though, are multi-faceted, complicating efforts to evaluate XBP1s as a therapeutic target for the collagenopathies. A transgenic mouse model in which a 13-base pair deletion mutant of Col10α(I) accumulates in hypertrophic chondrocytes shows XBP1s splicing and increased BiP expression, suggestive of a protective UPR that allows chondrocytes to survive ER stress. Chondrocyte differentiation in these mice, however, is disrupted, leading to delayed endochondral bone formation and chondrodysplasia in a process that may involve XBP1s-mediated increases in miRNA-214.[29-31] Stress-induced IRE1 activity has been proposed to exacerbate fibrosis through degradation of miR-150 and production of XBP1s.[32] Moreover, expression of the Schmid metaphyseal chondrodysplasia (MCDS) mutant Col10α(I) N617K in XBP1s-depleted cartilage revealed that UPR activation suppresses chondrocyte differentiation through CHOP, not XBP1s.[33] Thus, while XBP1s activity can protect cells from apoptosis following collagen retention and ER stress, hyper-activation of XBP1s can also disrupt normal development,[34] or in other cases may not be related to pathology at all.[33]

Still, targeted methods to control IRE1 and/or XBP1s activity independent of ER stress may prove useful in addressing the collagenopathies. Clearance of misfolded collagen by autophagy restored bone growth in osteoblasts expressing the osteogenesis imperfecta (OI) variant Colα2(I) G610C, and autophagy has been proposed as an innate protective mechanism for dealing with collagen aggregates that accompany expression of many disease variants.[35-38] Small molecule autophagy activators such as carbamazepine (CBZ) have yielded promising results in several instances,[39-41] but may require high doses that could produce off-target effects. A more selective approach might be XBP1s-mediated induction of autophagy,[42] which could reduce disease burden by removing aggregated collagen deposits and alleviating pathological ER stress. Doxycycline-controlled expression of XBP1s has been used to profile UPR transcriptional targets and achieve sustained, stress-independent XBP1s activity;[43, 44] small molecule activators of IRE1 have also been developed that are well-suited for short (~8 h) periods of IRE1 activity.[45] Application of similar strategies to collagen disease models may identify a therapeutic window for IRE1 and/or XBP1s activation that can resolve collagen misfolding in the absence of long-term effects on differentiation (Figure 1.1).
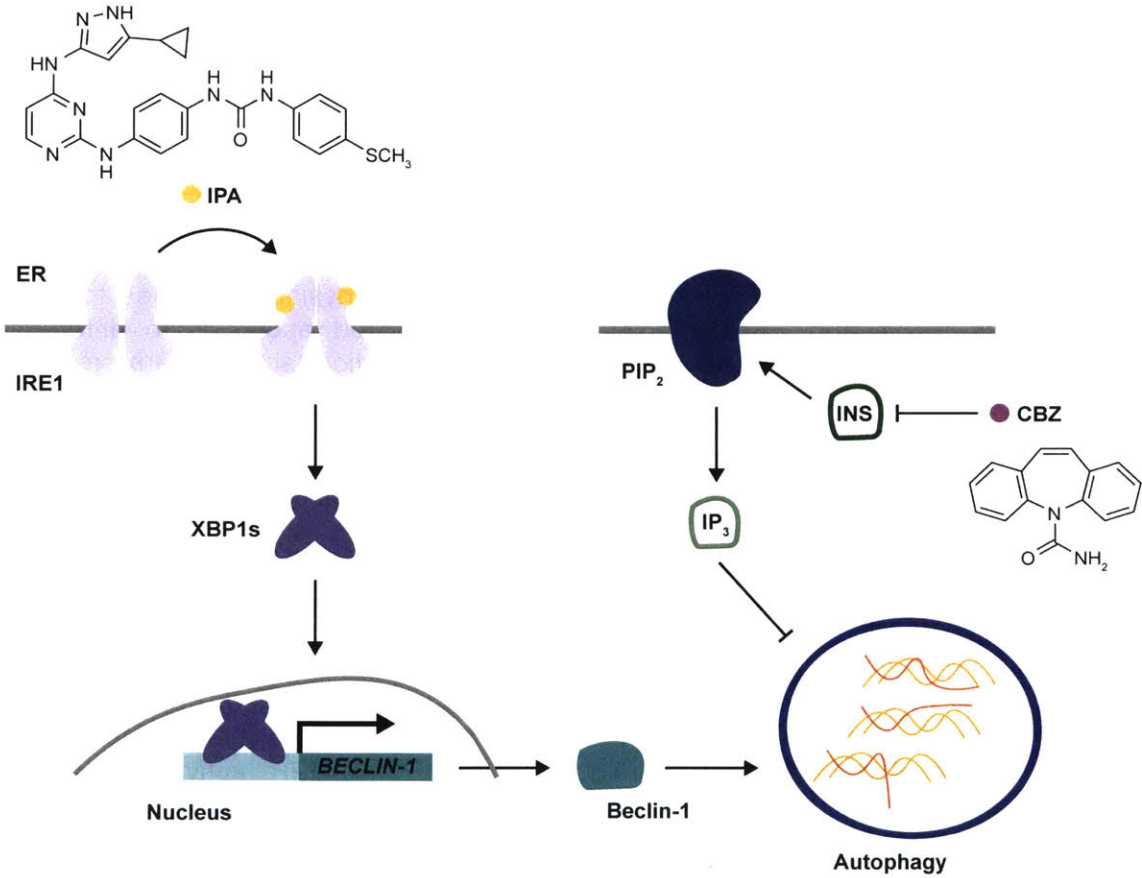
Figure 1.1: *Methods to regulate IRE1/XBP1s activity for addressing the collagenopathies.*
The small molecule IPA activates IRE1 RNase activity independent of its kinase domain,
leading to splicing of XBP1 mRNA and production of the active transcription factor XBP1s.
Among XBP1s transcriptional targets is Beclin-1, a positive regulator of autophagy that could
help clear misfolded collagen aggregates. Carbamazepine (CBZ) also activates autophagy,
but through IRE1 and XBP1s-independent signaling pathways. INS = inositol; IP_3 = inositol
1,4,5-trisphosphate; PIP_2 = Phosphatidylinositol 4,5-bisphosphate. Phosphoinositol signaling
pathway adapted from Vidal et al.[45]

ATF6 and ATF6-like transcription factors

Like XBP1s, ATF6 has physiological roles in the development of collagen-producing tissues. Overexpression of active or dominant-negative ATF6 by adenoviral transduction linked ATF6 expression to matrix mineralization downstream of bone morphogenic protein 2 (BMP2) and Runx2 signaling.[46] Activation of ATF6 also conferred protection against experimental inflammatory arthritis by C/EBPβ and mTOR-mediated inhibition of NF-κB signaling.[47] ATF6 regulates a partially overlapping set of genes with IRE1, but is activated by a distinct mechanism, in which the transcription factor is transported to the Golgi and processed by site 1 (S1) and site 2 (S2) proteases to give an active N-terminal fragment (ATF6f).

Engineered cell lines, in which ATF6 is fused to a metastable, small molecule-responsive domain, have been used to examine effects of selective, stress-independent ATF6 activation.[11, 18] An attractive alternate strategy is direct, small-molecule activation of ATF6 by compounds identified in a recent high-throughput screen.[48] However, further studies have revealed that the mechanism of action for one of these compounds involves covalent modification of ER chaperones PDIA3, 4, and 5, as well as ERP29.[49] The extent of modification was incomplete (~29%), but several of these chaperones have potential roles in collagen-I biosynthesis,[21] and so thorough characterization of ATF6-activating compounds will be needed to avoid counterproductive effects.

Beyond ATF6, a family of ATF6-like transcription factors has been identified with strong connections to collagen physiology. Despite low sequence homology, the osteoblast-specific OASIS and chondrocyte-specific BBF2H7 share key features of ATF6 structure, including a transcriptional activation domain, transmembrane domain, and motif for site 1 (S1) protease processing.[50] BBF2H7 increases expression of Sec23a, a gene required for collagen secretion, to promote chondrogenesis and protect chondrocytes from UPR-induced apoptosis; BBF2H7 knockout mice accordingly show chondrodysplasia along with a distended ER.[27, 51] In osteoblasts, OASIS increases expression of Colα1(I) to facilitate bone formation.[52] OASIS and BBF2H7 are regulated differently than ATF6, through stabilization of protein levels as opposed to disulfide reduction-mediated monomerization, and may not respond to the identified ATF6 activators.[53] Still, clarification of OASIS and UPRE binding motifs could allow for similar reporter-based screening approaches to discover compound leads for modulating these ATF6-like transcription factors in ways that could prove beneficial in the collagenopathies (Figure 1.2).

PERK

As the UPR branch responsible for initiating apoptotic signaling cascades, PERK is often treated more as part of the problem than the solution in protein misfolding diseases. ER stress-induced PERK signaling, however, is required for osteogenesis[54] and is activated in chondrocytes to facilitate collagen secretion.[55] But while PERK/ATF4 signaling protects chondrocytes from apoptosis, it also promotes chondrodysplasia by reverting chondrocyte differentiation.[56] Notably, the small molecule ISRIB,[57] which blocks PERK signaling by rendering cells insensitive to eIF2α phosphorylation, restored differentiation and lessened skeletal deformities in a mouse model of MCDS.[56] A similar approach using salubrinal showed positive effects in an osteoarthritis model.[58] However, the related compound guanbenz had no therapeutic effect, and the classification of these two compounds as eIF2α phosphorylation inhibitors has been challenged.[59-61] Obtaining a clearer understanding of the mode of action of PERK signaling inhibitors will thus be critical for understanding PERK pathophysiology in the context of collagen misfolding (Figure 1.3).
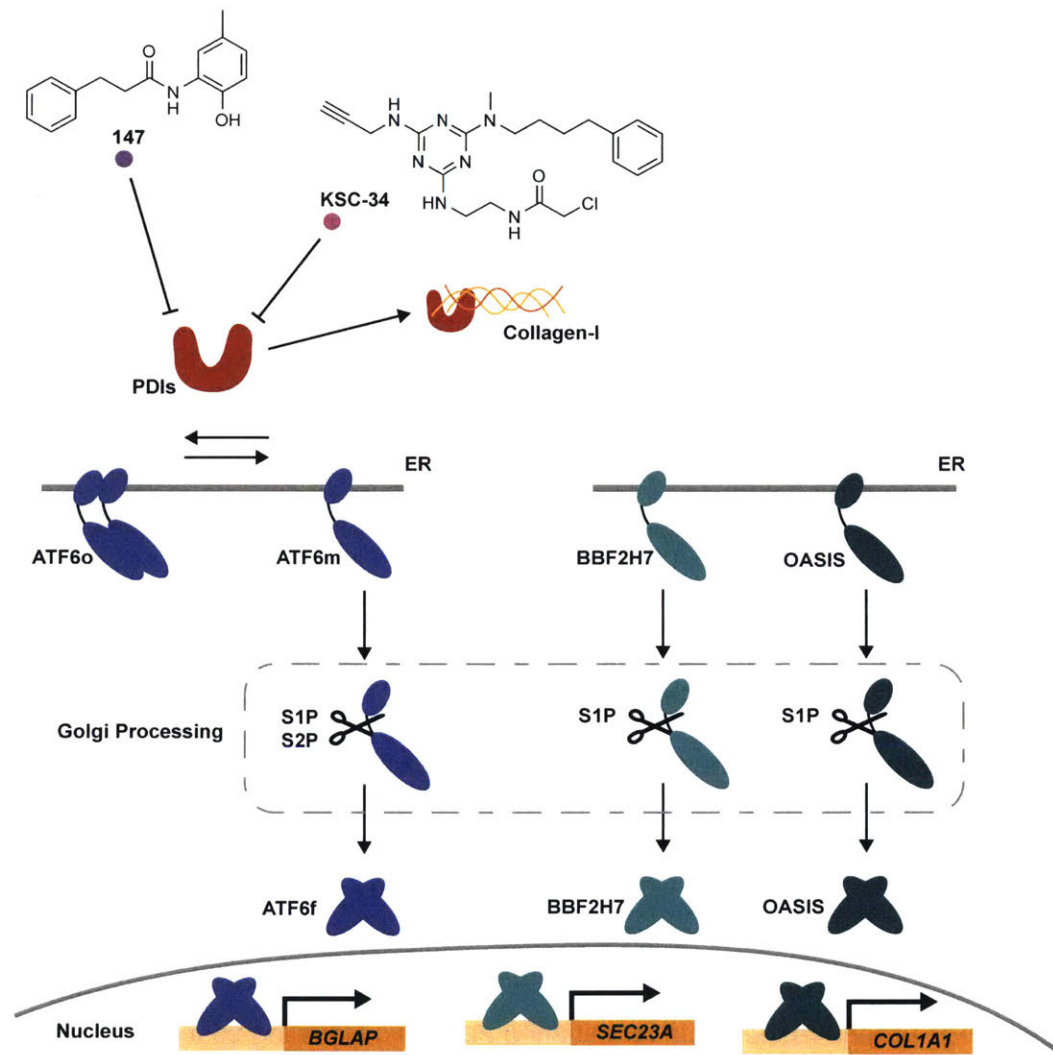
Figure 1.2: *Methods to regulate ATF6 activity for addressing the collagenopathies.* The small molecule 147 activates ATF6 by inhibiting the prolyl disulfide isomerases (PDIs) that promote ATF6 oligomerization (ATF6o); inhibition of PDIA3, PDIA4, and/or PDIA5 promotes formation of monomeric ATF6 (ATF6m), which is free to traffic to the Golgi. Processing by site 1 (S1P) and site 2 (S2P) proteases yields the active transcription factor ATF6f. Among ATF6 transcriptional targets is *BGLAP* (Osteocalcin), a positive regulator of osteoblast differentiation. The a-site specific PDIA1 inhibitor KSC-34 does not target PDIA3 or PDIA4, but could potentially prevent PDIA1-mediated hydroxylation of collagen-I in the absence of UPR activation. BBF2H7 and OASIS are two members of a family of ATF6-like transcription factors whose target genes are important for chondrogenesis and osteoblast differentiation, respectively. Methods for regulating the activity of these transcription factors may therefore prove useful for chondro- or skeletal dysplasias.
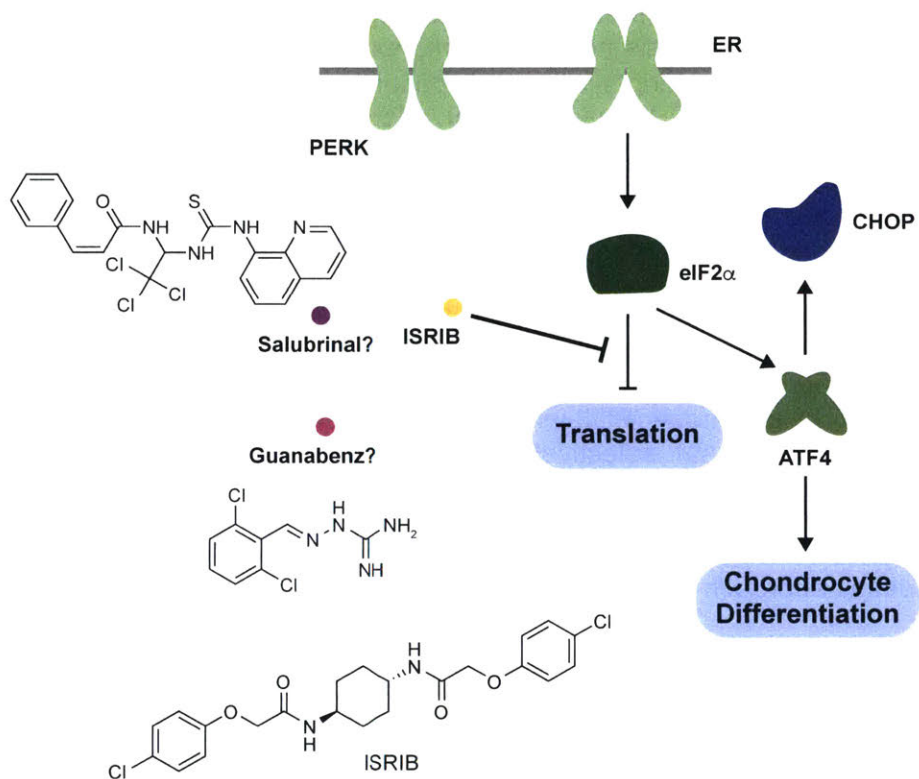
Figure 1.3: *Methods to regulate PERK activity for addressing the collagenopathies.* The small molecule ISRIB desensitizes cells to downstream signaling of phosphorylated eIF2α. Salubrinal and guanabenz have also been proposed to block effects of eIF2α phosphorylation, but their mechanism of action is not yet confirmed. PERK/ATF4 signaling can increase levels of the protein CHOP, which initiates apoptotic signaling cascades, but PERK/ATF4 signaling has also been implicated in chondrocyte differentiation. More work is therefore needed to determine therapeutic opportunities for modulating PERK activity in the context of the collagenopathies.

Targeting downstream proteostasis components

The class of compounds that has been most widely used to date in addressing the collagenopathies is chemical chaperones. 4-Phenylbutyric acid (PBA) reduced cell death in OI patient fibroblasts by promoting autophagic clearance of mutant Colα1(I) and Colα2(I) chains and has been used to reduce ER stress in Col2α(IV) mutant lines.[62, 63] PBA also reduced ER stress and autophagy, but increased Col5α(IV) expression in fibroblasts from patients with Alport Syndrome (AS).[64] Another chemical chaperone, trimethylamine *N*-oxide (TMAO), had a larger effect on collagen-I triple helix thermostability when incubated with purified wild type, R789C Col2α(I), or R992C Col2α(I) molecules than when added to cell culture,[65] suggesting that such compounds may act through direct binding to the target molecule rather than by modulating the cellular proteostasis network.[66] This explanation is consistent with the ability of PBA to increase Col5α(IV) transcripts in both AS patient lines, but improve secretion of only one mutant. In a separate study, PBA had minimal effects on retention of aggregated cartilage oligomeric matrix protein (COMP), underscoring the need to consider misfolding, disease-related proteins in the context of the proteostasis network.[67]

More promising than chemical chaperones, though still untested, are therapeutic strategies that target individual components of the collagen proteostasis network. The existence of at least nine genes linked to recessive forms of OI demonstrates that individual collagen chaperones can (and do) influence collagen production.[68] Key advantages of this approach—in contrast to the chemical chaperones—include disease-specific protein targets, improved mechanistic understanding, and access to existing medicinal chemistry knowledge. Extensive work on prolyl disulfide isomerase (PDI) inhibitors, for instance, enabled the recent development of a compound, KSC-34, that selectively binds the a-site catalytic domain.[69] Importantly, compound treatment has minimal effects on the UPR, indicating that PDIA1 can be targeted independent of induced ER stress. Similar improvements to the selectivity of inhibitors for BiP and other collagen-I chaperones will thus provide new opportunities for testing a proteostasis-targeted therapeutic approach in collagen disease models.

Conclusions

The UPR is integrally linked with development and disease in collagen-producing tissues. Connections between IRE1, ATF6, PERK, and multiple collagenopathies provide compelling evidence that regulating the different arms of the UPR could address diseases with varied causes and phenotypes. Due to the dual roles of UPR activation in development and disease, therapeutic approaches that target individual components of the proteostasis network will likely improve selectivity. However, many collagen-misfolding diseases result from inactivating mutations in collagen chaperones, and small molecule activators for these proteins are rare. Still, we note that significant advances in pharmacologic methods to tune UPR activity have come from the development of new screening platforms. Continued studies of disease-relevant signaling pathways can thus inform the design of screens for new compound leads. In the following chapters, we apply a similar approach to collagen-I proteostasis. We report the design and application of a high-throughput assay to identify selective small molecule modulators of collagen-I secretion (Chapter II). We then describe the results of mechanistic studies on a validated screening hit, 17-allylaminogeldanamycin (17-AAG), and the insights we obtain regarding its cellular target (Chapter III). Our results demonstrate that elements of the proteostasis network not traditionally considered to be collagen biosynthetic enzymes may also be important players in the collagenopathies. In addition, our findings emphasize the importance of the continued study of collagen biosynthesis and the physiology of collagen-producing tissues for discovery of new therapeutic targets.

References

[1] Ricard-Blum, S. (2011) The collagen family, *Cold Spring Harbor Perspect. Biol. 3*, a004978.

[2] Shoulders, M. D., and Raines, R. T. (2009) Collagen structure and stability, *Annu. Rev. Biochem. 78*, 929-958.

[3] Ishikawa, Y., and Bachinger, H. P. (2013) A molecular ensemble in the rER for procollagen maturation, *Biochim. Biophys. Acta 1833*, 2479-2491.

[4] Brinckmann, J. (2005) Collagens at a Glance, In *Collagen: Primer in Structure, Processing and Assembly* (Brinckmann, J., Notbohm, H., and Müller, P. K., Eds.), pp 1-6, Springer Berlin Heidelberg, Berlin, Heidelberg.

[5] Jobling, R., D'Souza, R., Baker, N., Lara-Corrales, I., Mendoza-Londono, R., Dupuis, L., Savarirayan, R., Ala-Kokko, L., and Kannu, P. (2014) The collagenopathies: review of clinical phenotypes and molecular correlations, *Curr. Rheumatol. Rep. 16*, 394.

[6] Makareeva, E., Mertz, E. L., Kuznetsova, N. V., Sutter, M. B., DeRidder, A. M., Cabral, W. A., Barnes, A. M., McBride, D. J., Marini, J. C., and Leikin, S. (2008) Structural heterogeneity of type I collagen triple helix and its role in osteogenesis imperfecta, *J. Biol. Chem. 283*, 4787-4798.

[7] Pihlajaniemi, T., Myllyla, R., and Kivirikko, K. I. (1991) Prolyl 4-hydroxylase and its role in collagen synthesis, *J. Hepatol. 13 Suppl. 3*, S2-7.

[8] Vasta, J. D., Andersen, K. A., Deck, K. M., Nizzi, C. P., Eisenstein, R. S., and Raines, R. T. (2016) Selective Inhibition of Collagen Prolyl 4-Hydroxylase in Human Cells, *ACS Chem. Biol. 11*, 193-199.

[9] Vasta, J. D., and Raines, R. T. (2016) Human Collagen Prolyl 4-Hydroxylase Is Activated by Ligands for Its Iron Center, *Biochemistry 55*, 3224-3233.

[10] Wong, M. Y., DiChiara, A. S., Suen, P. H., Chen, K., Doan, N.-D., and Shoulders, M. D. (2018) Adapting Secretory Proteostasis and Function Through the Unfolded Protein Response, In *Coordinating Organismal Physiology Through the Unfolded Protein Response* (Wiseman, R. L., and Haynes, C. M., Eds.), pp 1-25, Springer International Publishing, Cham.

[11] Shoulders, M. D., Ryno, L. M., Genereux, J. C., Moresco, J. J., Tu, P. G., Wu, C., Yates, J. R., 3rd, Su, A. I., Kelly, J. W., and Wiseman, R. L. (2013) Stress-independent activation of XBP1s and/or ATF6 reveals three functionally diverse ER proteostasis environments, *Cell Rep. 3*, 1279-1292.

[12] Plate, L., and Wiseman, R. L. (2017) Regulating Secretory Proteostasis through the Unfolded Protein Response: From Function to Therapy, *Trends Cell Biol. 27*, 722-737.

[13] Martinon, F., Chen, X., Lee, A. H., and Glimcher, L. H. (2010) TLR activation of the transcription factor XBP1 regulates innate immune responses in macrophages, *Nat. Immunol. 11*, 411-418.

[14] Martínez, G., Vidal, René L., Mardones, P., Serrano, Felipe G., Ardiles, Alvaro O., Wirth, C., Valdés, P., Thielen, P., Schneider, Bernard L., Kerr, B., Valdés, Jose L., Palacios, Adrian G., Inestrosa, Nibaldo C., Glimcher, Laurie H., and Hetz, C. (2016) Regulation of Memory Formation by the Transcription Factor XBP1, *Cell Rep. 14*, 1382-1394.

[15] Taylor, R. C., and Dillin, A. (2013) XBP-1 is a cell-nonautonomous regulator of stress resistance and longevity, *Cell 153*, 1435-1447.

[16] Maiuolo, J., Bulotta, S., Verderio, C., Benfante, R., and Borgese, N. (2011) Selective activation of the transcription factor ATF6 mediates endoplasmic reticulum proliferation triggered by a membrane protein, *Proc. Natl. Acad. Sci. U.S.A. 108*, 7832-7837.

[17] Kroeger, H., Grimsey, N., Paxman, R., Chiang, W.-C., Plate, L., Jones, Y., Shaw, P. X., Trejo, J., Tsang, S. H., Powers, E., Kelly, J. W., Wiseman, R. L., and Lin, J. H. (2018) The unfolded protein response regulator ATF6 promotes mesodermal differentiation, *Sci. Signaling 11*, eaan5785.

[18] Chen, J. J., Genereux, J. C., Qu, S., Hulleman, J. D., Shoulders, M. D., and Wiseman, R. L. (2014) ATF6 Activation Reduces the Secretion and Extracellular Aggregation of Destabilized Variants of an Amyloidogenic Protein, *Chem. Biol. 21*, 1564-1574.

[19] Cooley, C. B., Ryno, L. M., Plate, L., Morgan, G. J., Hulleman, J. D., Kelly, J. W., and Wiseman, R. L. (2014) Unfolded protein response activation reduces secretion and extracellular aggregation of amyloidogenic immunoglobulin light chain, *Proc. Natl. Acad. Sci. U.S.A. 111*, 13046-13051.

[20] Smith, S. E., Granell, S., Salcedo-Sicilia, L., Baldini, G., Egea, G., Teckman, J. H., and Baldini, G. (2011) Activating transcription factor 6 limits intracellular accumulation of mutant alpha(1)-antitrypsin Z and mitochondrial damage in hepatoma cells, *J. Biol. Chem. 286*, 41563-41577.

[21] DiChiara, A. S., Taylor, R. J., Wong, M. Y., Doan, N. D., Rosario, A. M., and Shoulders, M. D. (2016) Mapping and Exploring the Collagen-I Proteostasis Network, *ACS Chem. Biol. 11*, 1408-1421.

[22] Walter, P., and Ron, D. (2011) The unfolded protein response: from stress pathway to homeostatic regulation, *Science 334*, 1081-1086.

[23] Hollien, J., and Weissman, J. S. (2006) Decay of Endoplasmic Reticulum-Localized mRNAs During the Unfolded Protein Response, *Science 313*, 104-107.

[24] Shaffer, A. L., Shapiro-Shelef, M., Iwakoshi, N. N., Lee, A.-H., Qian, S.-B., Zhao, H., Yu, X., Yang, L., Tan, B. K., Rosenwald, A., Hurt, E. M., Petroulakis, E., Sonenberg, N., Yewdell, J. W., Calame, K., Glimcher, L. H., and Staudt, L. M. (2004) XBP1, Downstream of Blimp-1, Expands the Secretory Apparatus and Other Organelles, and Increases Protein Synthesis in Plasma Cell Differentiation, *Immunity 21*, 81-93.

[25] Malhotra, V., and Erlmann, P. (2015) The pathway of collagen secretion, *Annu. Rev. Cell Dev. Biol. 31*, 109-124.

[26] Han, X., Zhou, J., Zhang, P., Song, F., Jiang, R., Li, M., Xia, F., and Guo, F. J. (2013) IRE1alpha dissociates with BiP and inhibits ER stress-mediated apoptosis in cartilage development, *Cell Signal. 25*, 2136-2146.

[27] Hughes, A., Oxford, A. E., Tawara, K., Jorcyk, C. L., and Oxford, J. T. (2017) Endoplasmic Reticulum Stress and Unfolded Protein Response in Cartilage Pathophysiology; Contributing Factors to Apoptosis and Osteoarthritis, *Int. J. Mol. Sci. 18*, e665.

[28] Cameron, T. L., Gresshoff, I. L., Bell, K. M., Pirog, K. A., Sampurno, L., Hartley, C. L., Sanford, E. M., Wilson, R., Ermann, J., Boot-Handford, R. P., Glimcher, L. H., Briggs, M. D., and Bateman, J. F. (2015) Cartilage-specific ablation of XBP1 signaling in mouse results in a chondrodysplasia characterized by reduced chondrocyte proliferation and delayed cartilage maturation and mineralization, *Osteoarthritis Cartilage 23*, 661-670.

[29] Tsang, K. Y., Chan, D., Cheslett, D., Chan, W. C. W., So, C. L., Melhado, I. G., Chan, T. W. Y., Kwan, K. M., Hunziker, E. B., Yamada, Y., Bateman, J. F., Cheung, K. M. C., and Cheah, K. S. E. (2007) Surviving Endoplasmic Reticulum Stress Is Coupled to Altered Chondrocyte Differentiation and Function, *PLoS Biol. 5*, e44.

[30] Roberto, V. P., Gavaia, P., Nunes, M. J., Rodrigues, E., Cancela, M. L., and Tiago, D. M. (2018) Evidences for a New Role of miR-214 in Chondrogenesis, *Sci. Rep. 8*, 3704.

[31] Dong, L., Jiang, C. C., Thorne, R. F., Croft, A., Yang, F., Liu, H., de Bock, C. E., Hersey, P., and Zhang, X. D. (2011) Ets-1 mediates upregulation of Mcl-1 downstream of XBP-1 in human melanoma cells upon ER stress, *Oncogene 30*, 3716-3726.

[32] Heindryckx, F., Binet, F., Ponticos, M., Rombouts, K., Lau, J., Kreuger, J., and Gerwins, P. (2016) Endoplasmic reticulum stress enhances fibrosis through IRE1alpha-mediated degradation of miR-150 and XBP-1 splicing, *EMBO Mol. Med. 8*, 729-744.

[33] Cameron, T. L., Bell, K. M., Gresshoff, I. L., Sampurno, L., Mullan, L., Ermann, J., Glimcher, L. H., Boot-Handford, R. P., and Bateman, J. F. (2015) XBP1-Independent UPR

Pathways Suppress C/EBP-beta Mediated Chondrocyte Differentiation in ER-Stress Related Skeletal Disease, *PLoS Genet. 11*, e1005505.

[34] Ishikawa, T., Kashima, M., Nagano, A. J., Ishikawa-Fujiwara, T., Kamei, Y., Todo, T., and Mori, K. (2017) Unfolded protein response transducer IRE1-mediated signaling independent of XBP1 mRNA splicing is not required for growth and development of medaka fish, *eLife 6*, e26845.

[35] Mirigian, L. S., Makareeva, E., Mertz, E. L., Omari, S., Roberts-Pilgrim, A. M., Oestreich, A. K., Phillips, C. L., and Leikin, S. (2016) Osteoblast Malfunction Caused by Cell Stress Response to Procollagen Misfolding in α2(I)-G610C Mouse Model of Osteogenesis Imperfecta, *J. Bone Miner. Res. 31*, 1608-1616.

[36] Ishida, Y., Yamamoto, A., Kitamura, A., Lamande, S. R., Yoshimori, T., Bateman, J. F., Kubota, H., and Nagata, K. (2009) Autophagic elimination of misfolded procollagen aggregates in the endoplasmic reticulum as a means of cell protection, *Mol. Biol. Cell 20*, 2744-2754.

[37] Makareeva, E., Sun, G., Mirigian, L. S., Mertz, E. L., Vera, J. C., Espinoza, N. A., Yang, K., Chen, D., Klein, T. E., Byers, P. H., and Leikin, S. (2018) Substitutions for arginine at position 780 in triple helical domain of the alpha1(I) chain alter folding of the type I procollagen molecule and cause osteogenesis imperfecta, *PLoS One 13*, e0200264.

[38] Lisse, T. S., Thiele, F., Fuchs, H., Hans, W., Przemeck, G. K., Abe, K., Rathkolb, B., Quintanilla-Martinez, L., Hoelzlwimmer, G., Helfrich, M., Wolf, E., Ralston, S. H., and Hrabe de Angelis, M. (2008) ER stress-mediated apoptosis in a new mouse model of osteogenesis imperfecta, *PLoS Genet. 4*, e7.

[39] Mullan, L. A., Mularczyk, E. J., Kung, L. H., Forouhan, M., Wragg, J. M. A., Goodacre, R., Bateman, J. F., Swanton, E., Briggs, M. D., and Boot-Handford, R. P. (2017) Increased intracellular proteolysis reduces disease severity in an ER stress–associated dwarfism, *J. Clin. Invest. 127*, 3861-3865.

[40] Forouhan, M., Sonntag, S., and Boot-Handford, R. P. (2018) Carbamazepine reduces disease severity in a mouse model of metaphyseal chondrodysplasia type Schmid caused by a premature stop codon (Y632X) in the Col10a1 gene, *Hum. Mol. Genet.* ddy253.

[41] Vidal, R. L., Matus, S., Bargsted, L., and Hetz, C. (2014) Targeting autophagy in neurodegenerative diseases, *Trends Pharmacol. Sci. 35*, 583-591.

[42] Margariti, A., Li, H., Chen, T., Martin, D., Vizcay-Barrena, G., Alam, S., Karamariti, E., Xiao, Q., Zampetaki, A., Zhang, Z., Wang, W., Jiang, Z., Gao, C., Ma, B., Chen, Y. G., Cockerill, G., Hu, Y., Xu, Q., and Zeng, L. (2013) XBP1 mRNA splicing triggers an autophagic response in endothelial cells through BECLIN-1 transcriptional activation, *J. Biol. Chem. 288*, 859-872.

[43] Lee, A. H., Iwakoshi, N. N., and Glimcher, L. H. (2003) XBP-1 regulates a subset of endoplasmic reticulum resident chaperone genes in the unfolded protein response, *Mol. Cell Biol. 23*, 7448-7459.

[44] Dewal, M. B., DiChiara, A. S., Antonopoulos, A., Taylor, R. J., Harmon, C. J., Haslam, S. M., Dell, A., and Shoulders, M. D. (2015) XBP1s Links the Unfolded Protein Response to the Molecular Architecture of Mature N-Glycans, *Chem. Biol. 22*, 1301-1312.

[45] Mendez, A. S., Alfaro, J., Morales-Soto, M. A., Dar, A. C., McCullagh, E., Gotthardt, K., Li, H., Acosta-Alvear, D., Sidrauski, C., Korennykh, A. V., Bernales, S., Shokat, K. M., and Walter, P. (2015) Endoplasmic reticulum stress-independent activation of unfolded protein response kinases by a small molecule ATP-mimic, *eLife 4*, e05434.

[46] Jang, W. G., Kim, E. J., Kim, D. K., Ryoo, H. M., Lee, K. B., Kim, S. H., Choi, H. S., and Koh, J. T. (2012) BMP2 protein regulates osteocalcin expression via Runx2-mediated Atf6 gene transcription, *J. Biol. Chem. 287*, 905-915.

[47] Nakajima, S., Hiramatsu, N., Hayakawa, K., Saito, Y., Kato, H., Huang, T., Yao, J., Paton, A. W., Paton, J. C., and Kitamura, M. (2011) Selective Abrogation of BiP/GRP78 Blunts Activation of NF-κB through the ATF6 Branch of the UPR: Involvement of C/EBPβ and mTOR-Dependent Dephosphorylation of Akt, *Mol. Cell Biol. 31*, 1710-1718.

[48] Plate, L., Cooley, C. B., Chen, J. J., Paxman, R. J., Gallagher, C. M., Madoux, F., Genereux, J. C., Dobbs, W., Garza, D., Spicer, T. P., Scampavia, L., Brown, S. J., Rosen, H., Powers, E. T., Walter, P., Hodder, P., Wiseman, R. L., and Kelly, J. W. (2016) Small molecule proteostasis regulators that reprogram the ER to reduce extracellular protein aggregation, *eLife 5*, e15550.

[49] Paxman, R., Plate, L., Blackwood, E. A., Glembotski, C., Powers, E. T., Wiseman, R. L., and Kelly, J. W. (2018) Pharmacologic ATF6 activating compounds are metabolically activated to selectively modify endoplasmic reticulum proteins, *eLife 7*, e37168.

[50] Kondo, S., Saito, A., Asada, R., Kanemoto, S., and Imaizumi, K. (2011) Physiological unfolded protein response regulated by OASIS family members, transmembrane bZIP transcription factors, *IUBMB Life 63*, 233-239.

[51] Saito, A., Hino, S.-i., Murakami, T., Kanemoto, S., Kondo, S., Saitoh, M., Nishimura, R., Yoneda, T., Furuichi, T., Ikegawa, S., Ikawa, M., Okabe, M., and Imaizumi, K. (2009) Regulation of endoplasmic reticulum stress response by a BBF2H7-mediated Sec23a pathway is essential for chondrogenesis, *Nat. Cell Biol. 11*, 1197-1204.

[52] Murakami, T., Saito, A., Hino, S., Kondo, S., Kanemoto, S., Chihara, K., Sekiya, H., Tsumagari, K., Ochiai, K., Yoshinaga, K., Saitoh, M., Nishimura, R., Yoneda, T., Kou, I., Furuichi, T., Ikegawa, S., Ikawa, M., Okabe, M., Wanaka, A., and Imaizumi, K. (2009) Signalling mediated by the endoplasmic reticulum stress transducer OASIS is involved in bone formation, *Nat. Cell Biol. 11*, 1205-1211.

[53] Kondo, S., Hino, S. I., Saito, A., Kanemoto, S., Kawasaki, N., Asada, R., Izumi, S., Iwamoto, H., Oki, M., Miyagi, H., Kaneko, M., Nomura, Y., Urano, F., and Imaizumi, K. (2012) Activation of OASIS family, ER stress transducers, is dependent on its stabilization, *Cell Death Differ. 19*, 1939-1949.

[54] Saito, A., Ochiai, K., Kondo, S., Tsumagari, K., Murakami, T., Cavener, D. R., and Imaizumi, K. (2011) Endoplasmic reticulum stress response mediated by the PERK-eIF2(alpha)-ATF4 pathway is involved in osteoblast differentiation induced by BMP2, *J. Biol. Chem. 286*, 4809-4818.

[55] Hisanaga, S., Miyake, M., Taniuchi, S., Oyadomari, M., Morimoto, M., Sato, R., Hirose, J., Mizuta, H., and Oyadomari, S. (2018) PERK-mediated translational control is required for collagen secretion in chondrocytes, *Sci. Rep. 8*, 773.

[56] Wang, C., Tan, Z., Niu, B., Tsang, K. Y., Tai, A., Chan, W. C. W., Lo, R. L. K., Leung, K. K. H., Dung, N. W. F., Itoh, N., Zhang, M. Q., Chan, D., and Cheah, K. S. E. (2018) Inhibiting the integrated stress response pathway prevents aberrant chondrocyte differentiation thereby alleviating chondrodysplasia, *eLife 7*, e37673.

[57] Sidrauski, C., Acosta-Alvear, D., Khoutorsky, A., Vedantham, P., Hearn, B. R., Li, H., Gamache, K., Gallagher, C. M., Ang, K. K., Wilson, C., Okreglak, V., Ashkenazi, A., Hann, B., Nader, K., Arkin, M. R., Renslo, A. R., Sonenberg, N., and Walter, P. (2013) Pharmacological brake-release of mRNA translation enhances cognitive memory, *eLife 2*, e00498.

[58] Hamamura, K., Nishimura, A., Iino, T., Takigawa, S., Sudo, A., and Yokota, H. (2015) Chondroprotective effects of Salubrinal in a mouse model of osteoarthritis, *Bone Joint Res. 4*, 84-92.

[59] Crespillo-Casado, A., Chambers, J. E., Fischer, P. M., Marciniak, S. J., and Ron, D. (2017) PPP1R15A-mediated dephosphorylation of eIF2alpha is unaffected by Sephin1 or Guanabenz, *eLife 6*, e26109.

[60] Das, I., Krzyzosiak, A., Schneider, K., Wrabetz, L., D'Antonio, M., Barry, N., Sigurdardottir, A., and Bertolotti, A. (2015) Preventing proteostasis diseases by selective inhibition of a phosphatase regulatory subunit, *Science 348*, 239-242.

[61] Tsaytler, P., Harding, H. P., Ron, D., and Bertolotti, A. (2011) Selective Inhibition of a Regulatory Subunit of Protein Phosphatase 1 Restores Proteostasis, *Science 332*, 91-94.

[62] Besio, R., Iula, G., Garibaldi, N., Cipolla, L., Sabbioneda, S., Biggiogera, M., Marini, J. C., Rossi, A., and Forlino, A. (2018) 4-PBA ameliorates cellular homeostasis in fibroblasts from osteogenesis imperfecta patients by enhancing autophagy and stimulating protein secretion, *Biochim. Biophys. Acta 1864*, 1642-1652.

[63] Murray, L. S., Lu, Y., Taggart, A., Van Regemorter, N., Vilain, C., Abramowicz, M., Kadler, K. E., and Van Agtmael, T. (2014) Chemical chaperone treatment reduces intracellular accumulation of mutant collagen IV and ameliorates the cellular phenotype of a COL4A2 mutation that causes haemorrhagic stroke, *Hum. Mol. Genet. 23*, 283-292.

[64] Wang, D., Mohammad, M., Wang, Y., Tan, R., Murray, L. S., Ricardo, S., Dagher, H., van Agtmael, T., and Savige, J. (2017) The Chemical Chaperone, PBA, Reduces ER Stress and Autophagy and Increases Collagen IV alpha5 Expression in Cultured Fibroblasts From Men With X-Linked Alport Syndrome and Missense Mutations, *Kidney Int. Rep. 2*, 739-748.

[65] Gawron, K., Jensen, D. A., Steplewski, A., and Fertala, A. (2010) Reducing the effects of intracellular accumulation of thermolabile collagen II mutants by increasing their thermostability in cell culture conditions, *Biochem. Biophys. Res. Commun. 396*, 213-218.

[66] Omachi, K., Kamura, M., Teramoto, K., Kojima, H., Yokota, T., Kaseda, S., Kuwazuru, J., Fukuda, R., Koyama, K., Matsuyama, S., Motomura, K., Shuto, T., Suico, M. A., and Kai, H. (2018) A Split-Luciferase-Based Trimer Formation Assay as a High-throughput Screening Platform for Therapeutics in Alport Syndrome, *Cell Chem. Biol. 25*, 634-643.

[67] Posey, K. L., Coustry, F., Veerisetty, A. C., Liu, P., Alcorn, J. L., and Hecht, J. T. (2014) Chondrocyte-specific pathology during skeletal growth and therapeutics in a murine model of pseudoachondroplasia, *J. Bone Miner. Res. 29*, 1258-1268.

[68] Marini, J. C., and Blissett, A. R. (2013) New Genes in Bone Development: What's New in Osteogenesis Imperfecta, *J. Clin. Endocrinol. Metab. 98*, 3095-3103.

[69] Cole, K. S., Grandjean, J. M. D., Chen, K., Witt, C. H., O'Day, J., Shoulders, M. D., Wiseman, R. L., and Weerapana, E. (2018) Characterization of an A-Site Selective Protein Disulfide Isomerase A1 Inhibitor, *Biochemistry 57*, 2035-2043.

CHAPTER II
Development of a High-Throughput Assay for Collagen-I Secretion

Summary

Collagen overproduction is a feature of fibrosis and cancer, while insufficient deposition of functional collagen molecules and/or the secretion of malformed collagen are common in genetic disorders like osteogenesis imperfecta. Collagen secretion is an appealing therapeutic target in these and other diseases, as secretion directly connects intracellular biosynthesis to collagen deposition and biological function in the extracellular matrix. However, small molecule and biological methods to tune collagen secretion are severely lacking. Their discovery could prove useful not only in the treatment of disease, but also in providing tools for better elucidating mechanisms of collagen biosynthesis. We developed a cell-based, high-throughput luminescent assay of collagen type I secretion and used it to screen for small molecules that selectively enhance or inhibit that process. Among several validated hits, the Hsp90 inhibitor 17-allylaminogeldanamycin (17-AAG) robustly decreases the secretion of collagen-I by our model cell line and by human primary cells. In these systems, 17-AAG reduces collagen-I secretion post-translationally and is not a global inhibitor of protein secretion. Beyond 17-AAG, we also identify the broad-spectrum antiparasitic compound nitazoxanide (NTZ) as a post-translational inhibitor of collagen-I secretion. Our results highlight the potential of a cell-based high-throughput screen for selective modulators of collagen secretion.

Introduction

By both function and sheer mass percentage, collagen constitutes the major component of animal tissue.[1] Twenty-eight distinct types of collagen play important roles in architecturally diverse extracellular matrices, ranging from skin and bone to cartilage and basement membranes.[2] In addition to providing the structural framework for these tissues, the collagens have dynamic functions in numerous biological processes.[3, 4] For example, collagens engage integrins on cell surfaces, influence wound-healing responses and inflammation, and play critical roles in cell differentiation, organ development, and tissue maintenance.

Collagen biosynthesis is a complex process, encompassing extensive post-translational modifications, folding and assembly, propeptide cleavage, secretion, and extracellular fibril formation.[5] It is not surprising, then, that dysregulated collagen homeostasis is closely related to numerous pathologies.[6] Fibrosis is characterized by collagen overproduction and often leads to organ damage or failure.[7] Many cancers also feature high levels of collagen secretion and matrix remodeling, which promote metastasis.[8-10] Conversely, insufficient deposition of properly structured collagen and/or excessive accumulation of misfolded collagen molecules give rise to an array of phenotypically diverse pathologies.[11] These genetic disorders, ranging from osteogenesis imperfecta to Ehlers-Danlos syndrome, are typically caused by mutations in collagen genes or chaperones.[12]

Despite substantial work characterizing fibrosis[13] and the collagenopathies,[11, 14] many of the underlying disease mechanisms remain poorly defined, especially with respect to why and how collagen homeostasis fails. Biological and small molecule tools that selectively target the processes involved in the biosynthesis of collagen type I, the most abundant collagen type,[15] could yield new mechanistic insights, in addition to providing leads for therapeutic strategies. However, relatively few small molecules that can selectively alter either collagen-I folding or secretion have been identified to date.[16-18] Furthermore, high-throughput methods to facilitate the discovery and design of new compounds are limited, in large part because assaying the folding and secretion of a protein that not only lacks enzymatic activity but also requires a complex cellular folding environment[5] is inherently challenging.

To address this need, we developed a high throughput cell-based, luminescent assay for modulators of collagen-I secretion. We screened 1228 known bioactive and FDA-approved compounds, validating a number of hit molecules that selectively affect collagen-I secretion but not the secretion of a control protein. Of particular interest, we find that the Hsp90 inhibitor 17-allylaminogeldanamycin (17-AAG) selectively reduces the secretion of our collagen-I reporter. We confirmed the ability of 17-AAG to reduce the secretion of endogenous collagen-I from human primary cells independent of effects on collagen-I transcripts or synthesis. Beyond 17-AAG, we also identify the broad-spectrum antiparasitic compound nitazoxanide (NTZ) as another potential inhibitor of collagen-I secretion. Follow-up studies using these and other screening hits are thus expected to provide additional insights into mechanisms of collagen-I biosynthesis and secretion, as well as new potential opportunities for therapeutic intervention.

Results

Luminescent Assay Design and Validation. Because collagen-I lacks enzymatic activity, current methods for assaying collagen-I levels in moderate- to high-throughput fashion primarily rely on the binding of hydrophobic dyes to deposited collagen fibers.[19, 20] Other assays are conducted *in vitro* with purified collagen-I and recombinant collagen binding proteins[18, 21] or measure collagen transcription,[22] rather than the many post-transcriptional processes key to collagen biosynthesis.

We reasoned that appending a luciferase enzyme to collagen-I would allow for direct monitoring of fusion protein secretion, with increased sensitivity from the enzymatic signal amplification. With this approach, the screen would be suitable for discovering both activators and inhibitors and would not require prior knowledge of collagen folding, quality control, and secretion pathways. For our reporter molecule, we selected the engineered, or "enhanced",

variant of Gaussia luciferase (eGLuc2). eGLuc2 is natively secreted, small, bright, ATP-independent, and highly stable, and has been used successfully for high-throughput screening.[23-27] As native collagen-I is a 2:1 Colα1(I):Colα2(I) heterotrimer,[28] we attached eGLuc2 to Colα2(I) to control the stoichiometry of fusion protein incorporation. We positioned eGLuc2 at the extreme N-terminus to avoid interfering with triple-helix nucleation by collagen's C-terminal propeptide (Figure 2.1A).[29] We also included the eGLuc2 signal sequence to direct eGLuc2.Colα2(I) to the secretory pathway. Finally, we engineered the fusion protein gene to be under control of a tetracycline repressor-regulated, doxycycline (Dox)-inducible promoter both to simplify the long-term propagation of stable cell lines and to provide temporal control of eGLuc2.Colα2(I) expression.

We next engineered Saos-2 osteosarcoma cells, which are osteoblast-like and produce endogenous collagen-I,[30, 31] to stably and inducibly express eGLuc2.Colα2(I). We had previously shown that Saos-2 cells stably expressing the tetracycline repressor (termed Saos-2-TREx cells) can be used to control the heterologous expression of antibody epitope-tagged collagen-I variants.[32] Building on that work, we transduced Saos-2-TREx cells with an eGLuc2.Colα2(I)-encoding lentivirus, selected for stable gene incorporation, validated protein expression by immunoblotting (Figure 2.2A), and isolated a genetically homogeneous single-colony line for assay development. eGLuc2.Colα2(I) secretion can be assayed in the resulting Saos-2$^{GLuc.Col}$ cells by inducing transcription with Dox and then measuring the extent of secretion via a luminescence assay on the conditioned media (Figure 2.1A). We verified that the intensity of the extracellular luminescent signal is increased in response to sodium ascorbate (Figure 2.2B), which induces collagen-I translation and also facilitates collagen-I secretion by recycling the prolyl-4-hydroxylase metallocofactor.[28, 33, 34] The level of eGLuc2.Colα2(I) secretion is also strongly reduced by treatment with the ER-to-Golgi transport inhibitor Brefeldin A (Figure 2.2C).[35] Thus, the secretion of our eGLuc2.Colα2(I) fusion protein is dependent on mechanisms known to be involved in collagen export. In tandem, we also prepared a Saos-2^{GLuc} cell line, which inducibly expresses unfused eGLuc2 under the tetracycline repressor, as a control for our high-throughput screen.

Although we and others[32, 36] have generated N-terminal fusions to collagen-I and shown that such proteins are well-behaved, we sought to directly evaluate whether the presence of eGLuc2 disrupts collagen-I folding or assembly. We first determined whether eGLuc2.Colα2(I) associates with Colα1(I) intracellularly. Immunoprecipitating eGLuc2.Colα2(I) with an α-GLuc antibody from Saos-2$^{GLuc.Col}$ cell lysate revealed that both eGLuc2.Colα2(I) and Colα1(I) are enriched in the eluted fraction (Figure 2.1B), indicating formation of a stable heteromeric complex, as expected, and suggesting that eGLuc2.Colα2(I) can successfully assemble with Colα1(I).

We next sought to determine if secreted eGLuc2.Colα2(I) also formed stable heteromers with Colα1(I). As disulfide-bond formation in the C-terminal propeptide is needed to initiate collagen-I triple-helix formation,[5] we precipitated collagen-I from conditioned media of Saos-2$^{GLuc.Col}$ cells and assayed for disulfide-dependent assembly by SDS–PAGE (Figure 2.1C). We observed that secreted eGLuc2.Colα2(I) associates with Colα1(I), as detected by the co-staining of α-GLuc and Colα1(I) antibodies under nonreducing conditions. While we observe monomeric (~190 kDa) and dimeric (~245 kDa) species for both Colα1(I) and eGLuc2.Colα2(I) under non-reducing conditions, likely owing to disulfide shuffling during SDS–PAGE separation, the banding pattern is consistent with previous observations by us and others for native collagen-I.[37, 38] Importantly, in the presence of the reducing agent 1,4-dithiothreitol (DTT), both eGLuc2.Colα2(I) and Colα1(I) collapse to the monomeric molecular weight, as expected for dissociation of disulfide-linked trimers.
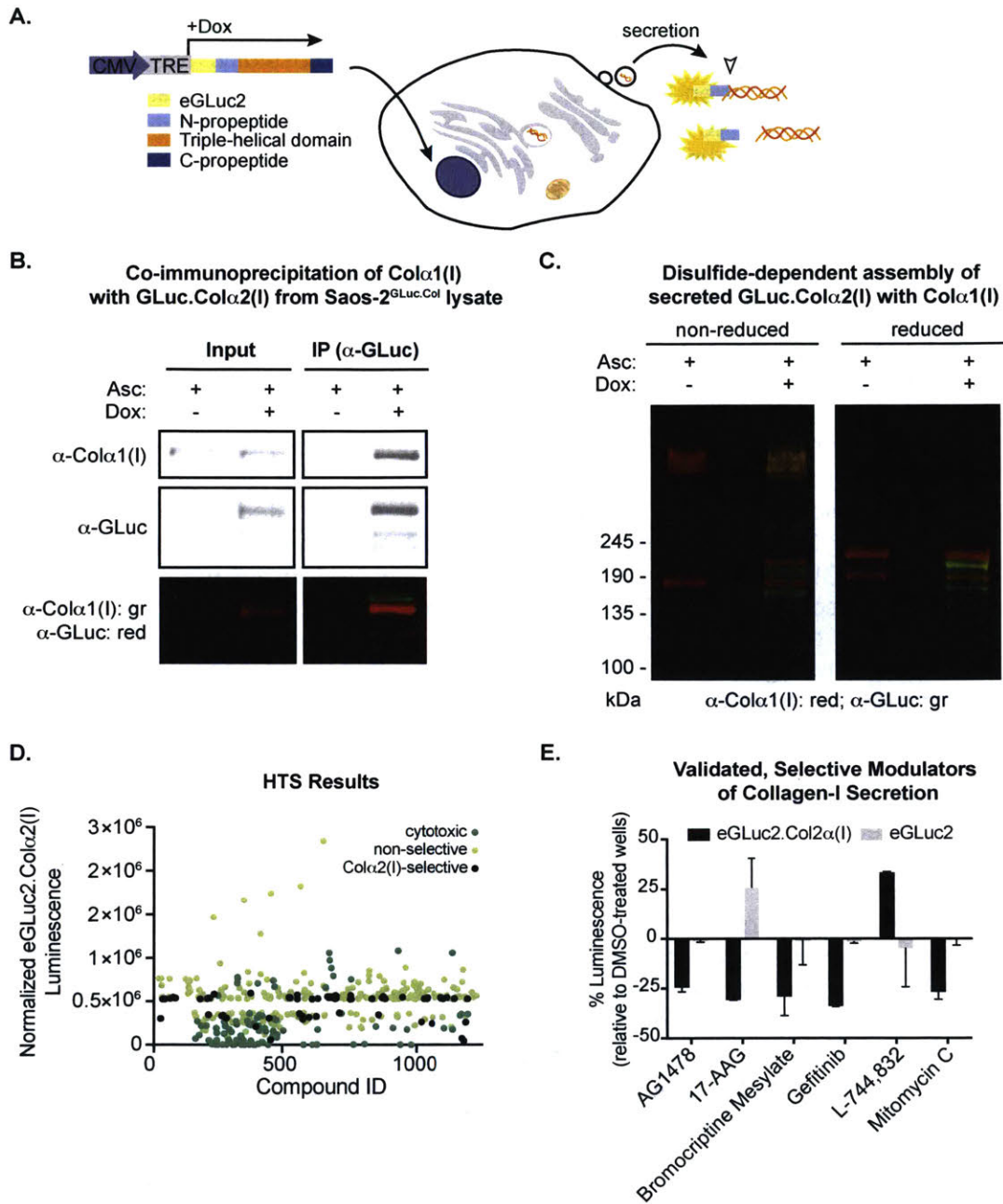
Figure 2.1: *High-throughput screening strategy and results.* (A) Schematic of screen design. Addition of 1 μg/mL doxycycline (Dox) to the stable single-colony Saos-2$^{GLuc.Col}$ cell line induces expression of the eGLuc2.Colα2(I) fusion protein, which can be proteolytically processed postsecretion. Luminescence generated by the fusion protein may be assayed upon addition of the eGLuc2 substrate. (B) eGLuc2.Colα2(I) from Saos-2$^{GLuc.Col}$ lysate was immunoprecipitated with an α-GLuc antibody and analyzed by immunoblotting, revealing that eGLuc2.Colα2(I) associates intracellularly with Colα1(I). (C) Disulfide-dependent assembly of eGLuc2.Colα2(I) with Colα1(I). Collagen-I precipitated from conditioned media of Saos-2$^{GLuc.Col}$ cells was treated with iodoacetamide and heated in Laemmli buffer with or without DTT. Samples were separated by SDS–PAGE and analyzed by immunoblotting. (D) Plot of filtered screening results. (E) Validated small molecule modulators of collagen-I secretion.

A.

eGLuc2.Colα2(I) Expression in Saos-2$^{GLuc.Col}$ Cells

B.

Ascorbate-Responsive Secretion of eGLuc2.Colα2(I)

C.

Brefeldin A Inhibits eGLuc2.Colα2(I) Secretion
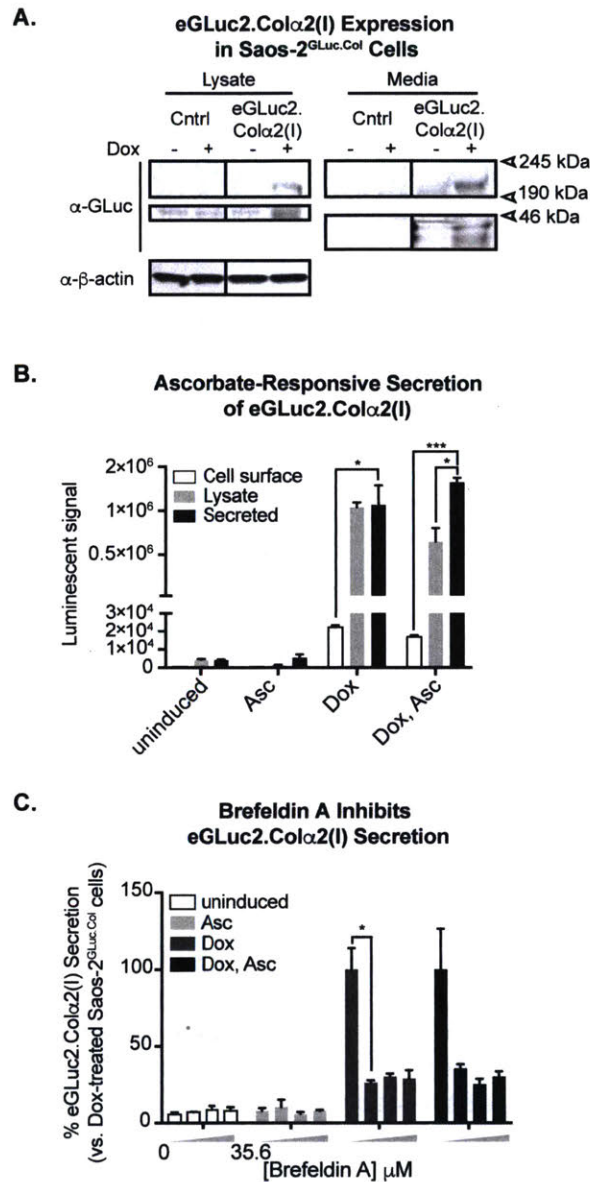
Figure 2.2: *Validation of Saos-2$^{GLuc.Col}$ cells to assay collagen-I secretion.* (A) Western blot showing Dox-inducible expression of the eGLuc2.Colα2(I) fusion protein in a Saos-2-TREx heterostable cell line. Cells were treated for 48 h with 2 μg/mL Dox prior to harvesting and analysis; conditioned media or 50 μg lysate from cells was boiled with Laemmli buffer + 167 mM dithiothreitol. Samples were then separated by SDS-PAGE on homemade 4/8/12% Tris-Gly gels, transferred to nitrocellulose membranes, and developed with a rabbit anti-GLuc antibody. The band at ~200 kDa corresponds to intact fusion protein; post-secretion, collagen-I produced by Saos-2-TREx cells can be proteolytically processed by proteinases that cleave the N-propeptide from the triple-helical domain, leaving a GLuc-reactive band at ~40 kDa. (B) eGLuc2.Colα2(I) luminescent activity was measured in conditioned media from intact cells (treated briefly with trypsin and washed with media), or from cells lysed by repeated freeze-thaw cycles. Significant luminescence is detected only from media and cell lysates. Cells were treated with 1 μg/mL Dox for 24 h and/or 50 μM sodium ascorbate, as indicated, prior to a GLuc assay. * = $p < 0.05$; *** = $p < 0.001$ ($n = 3$). (C) GLuc assay performed on cells induced for 24 h with 1 μg/mL Dox and treated with DMSO or increasing concentrations of Brefeldin A (L to R: 0, 3.56, 17.8, or 35.6 μM). * = $p < 0.05$ ($n = 3$).

We also performed pepsin digests of precipitated collagen-I from Saos-2$^{GLuc.Col}$ cells. Properly folded collagen-I triple helices are known to be pepsin-resistant[39] and, indeed, we find that upon Dox treatment, a second pepsin-resistant band corresponding to propeptide-cleaved eGLuc2.Colα2(I) appears in addition to propeptide-cleaved Colα1(I) (Figure 2.3). Together, these results suggest that the eGLuc2.Colα2(I) fusion protein is incorporated into stable triple helices with Colα1(I) and that eGLuc2 tagging does not significantly disrupt collagen-I assembly or folding.
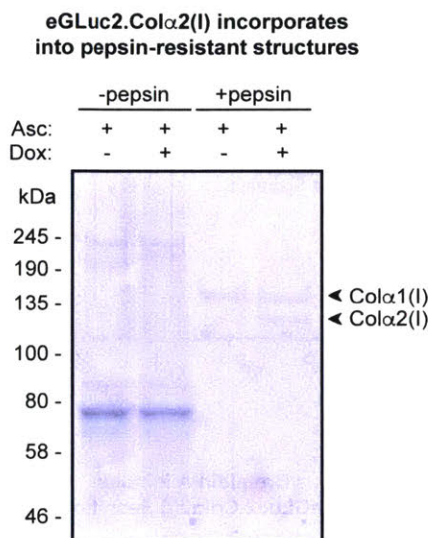
Figure 2.3: *eGLuc2.Colα2(I) incorporates into pepsin-resistant structures.* Collagen-I was selectively precipitated from conditioned media of Saos-2$^{GLuc.Col}$ cells with ammonium sulfate. Precipitated collagen-I pellets were then resuspended in 0.5 M acetic acid and incubated with or without 0.2 mg/mL pepsin for 48 h at 4 °C. While pepsin-resistant Colα1(I) is clearly visible for both digested lanes, treatment with Dox gives rise to a second band in the pepsin-digested lane corresponding to pepsin-stable eGLuc2.Colα2(I).

High-Throughput Screening for Collagen-I Secretion Modulators. We next miniaturized our collagen secretion assay for high-throughput screening in 384-well plates, tested for linearity, and optimized a timeline for induction of collagen expression and compound addition (detailed assay parameters are presented in Materials and Methods; full, normalized screening data are presented in Table S1 of ref. [75]). For our initial screen of eGLuc2.Colα2(I) secretion, we employed the Harvard Medical School Institute of Chemistry and Cell Biology (ICCB) Bioactives and the FDA-Approved small molecule libraries, commercially available from Enzo Life Sciences. Together, the libraries contained 1228 unique compounds spanning at least 22 drug target classes. Our results from analyzing conditioned media after eGLuc2.Colα2(I) secretion for 24 h were quite reproducible, as indicated by a low coefficient of variation (% CV, vehicle-treated wells) and a Pearson's correlation coefficient of 0.9611 between eGLuc2.Colα2(I) replicates for compound treatments across both libraries (Figure 2.4A, C).

A.

	Assay Reproducibility		
	% CV		Pearson's Coefficient
Cell Line	Avg (%)	Std Dev (%)	r, Confidence Interval
Saos-2[GLuc]	8.79	1.84	0.9214 (0.9126-0.9294)
Saos-2[GLuc.Col]	5.84	0.89	0.9611 (0.9567-0.9651)

B.

HTS Filtering Schematic

eGLuc2.Colα2(I) Luminescence — 1228 cmpds — 1° screen

Keep if ≥1.25 or ≤0.80 fold-change

CellTiter-Glo Luminescence — 302 hits — 2° screen

Keep if <3SD from DMSO Average

eGLuc2 Luminescence — 199 hits — counter screen

Keep if <1SD from DMSO Average

Follow-up Testing in Saos-2[GLuc.Col] cells — 17 hits

same result as screen?

6 Validated, Selective Collagen Secretion Modulators

C.

Correlation of eGLuc2.Colα2(I) Screening Replicates

Normalized Luminescence (Replicate 2) vs Normalized Luminescence (Replicate 1)

Figure 2.4: Assay reproducibility. (A) Average coefficient of variation and Pearson's correlation coefficient. (B) Overview of counterscreening and filtering strategies and results. Conditions for cell plating and compound addition in all screens are described in the Materials and Methods. (C) Correlation between screening replicates of unfiltered, normalized eGLuc2.Colα2(I) data for all compounds.

A number of factors, including compound toxicity or effects on fusion protein transcription/translation, general protein secretion, or luciferase enzyme activity, might produce a response in this high-throughput assay, despite being unrelated to our desired phenotype of selective effects on collagen-I secretion. We performed two counterscreens to filter out such false positives: (1) we used a CellTiter-Glo assay to remove cytotoxic compounds, and (2) we employed an eGLuc2 secretion assay to remove compounds that shift the eGLuc2-derived signal in the same direction as the eGLuc2.Colα2(I)-derived signal (Figure 2.4B). The results are presented in Figure 2.1D. Notably, many of the compounds that strongly decreased eGLuc2.Colα2(I) luminescence were also flagged in our CellTiter-Glo secondary screen as cytotoxic (signal more than three standard deviations below the average DMSO signal) at the concentrations used and were therefore triaged. Meanwhile, many of the compounds that yielded the largest changes in normalized luminescence in our eGLuc2.Colα2(I) secretion assay, some on the order of ~5-fold increases, were triaged because they had similar effects on the luminescence signal obtained in media from our control Saos-2[GLuc] cell line. The striking effects on protein secretion yielded by some of these compounds (see Table S1 of ref. [75]) suggest that they may prove useful as broad-spectrum modulators of secretory pathway function, although the effects may also be due merely to modification of eGLuc2 luminescence. Regardless, the aim of our screening protocol was to identify compounds that selectively target collagen-I secretion. Those with obviously non-specific effects on protein secretion were not pursued further in this study.

After filtering the primary screening results as described above, we obtained 17 putative hit compounds that selectively modulate secretion of eGLuc2.Colα2(I) from Saos-2 cells. A list

of these compounds, their effects on eGLuc2.Colα2(I) secretion, and their annotated functions is provided in Table 2.1, and structures of the six validated hits are presented in Figure 2.5.

Table 2.1. Candidate Hit Compounds.

Compound	Effect on eGLuc2.Colα2(I) (% Change vs DMSO-treated cells)	Effect on eGLuc2 (% Change vs DMSO-treated cells)	Annotated Function/Mechanism
17-AAG	−30.7 ± 0.2	25.9 ± 14.7	Hsp90 inhibitor
Ac-Leu-Leu-Nle-CHO	30.5 ± 1.3	7.6 ± 3.4	Calpain inhibitor
AG1478	−24.4 ± 2.3	−1.0 ± 0.6	Tyrosine kinase inhibitor
Benazepril HCl	25.0 ± 6.7	3.2 ± 2.5	ACE inhibitor
Bromocriptine Mesylate	−29.0 ± 9.7	−0.9 ± 12.3	Dopaminergic receptor agonist
Ceftazidime	27.1 ± 15.4	0.6 ± 0.6	Antibiotic
Chlorphenirami ne Maleate	25.7 ± 4.3	6.9 ± 2.5	Antihistamine
15-deoxy-prostaglandin J2	−32.3 ± 7.7	−5.0 ± 7.4	Bioactive prostaglandin
Gefitinib	−34.0 ± 0.4	−1.6 ± 0.7	EGFR inhibitor
L-744, 832	33.3 ± 0.5	−4.7 ± 19.5	Ras farnesyltransferase inhibitor
Losartan Potassium	27.4 ± 3.3	7.2 ± 0.6	Type-1 angiotensin II receptor (AT1) antagonist
Mitomycin C	−26.8 ± 3.8	−0.3 ± 3.0	DNA crosslinker
MY-5445	−25.3 ± 0.1	8.5 ± 22.7	Phosphodiesterase (PDE5) inhibitor
Oxaprozin	26.5 ± 5.9	−17.1 ± 44.1	NSAID
RG-14620	−31.2 ± 3.6	−8.2 ± 4.4	EGFR inhibitor
SDZ-201106	−22.6 ± 1.5	−5.7 ± 12.3	Sodium channel ligand
Tolcapone	28.5 ± 4.5	5.6 ± 1.7	Catechothyltransferase (COMT) inhibitor
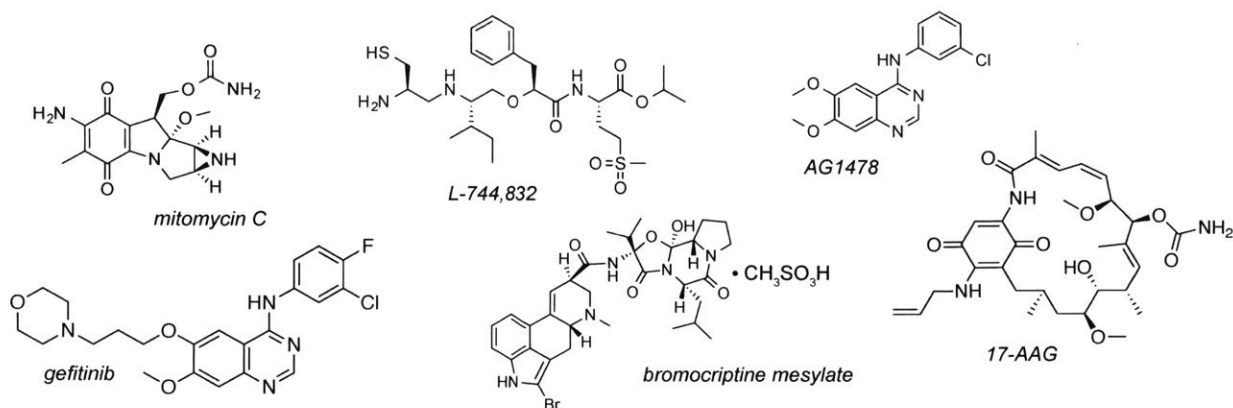
Figure 2.5: *Structures of validated compounds that selectively modulate collagen-I secretion.* Associated screening results and annotated functions are given in Table 2.1.

The Hsp90 Inhibitor 17-AAG Reduces the Secretion of Endogenous Collagen-I. We repurchased and retested 15 of the 17 compounds that passed our screening filters, yielding six validated small molecules that selectively alter eGLuc2.Colα2(I), but not eGLuc2, secretion and are not cytotoxic in Saos-2 cells at the tested concentrations (Figure 2.1E and 2.5). 17-Allylaminogeldanamycin (17-AAG), a particularly well-characterized inhibitor of Hsp90-type chaperones,[40, 41] stood out to us from these validated compound hits because it has opposite effects on eGLuc2.Colα2(I) and eGLuc2 secretion (Figure 2.1E). Furthermore, 17-AAG and its analogue 17-dimethylaminoethylamino-17-demethoxygeldanamycin (17-DMAG) have been reported to display beneficial effects in fibrosis models, by a mechanism involving reduction of inflammatory signaling but not by direct inhibition of collagen biosynthesis, folding, or secretion.[42-45] We, however, hypothesized that Hsp90 inhibition might directly impede collagen-I secretion. To explore this possibility further, we chose to follow up on the 17-AAG results from our high-throughput screen and investigate the mechanism of action.

To begin, we treated Saos-2$^{GLuc.Col}$ cells with a range of 17-AAG concentrations and observed a dose-dependent reduction in the intensity of the luminescent signal (Figure 2.6A). Next, we focused on testing whether these results would be recapitulated in primary cells expressing endogenous collagen-I, instead of an eGLuc2.Colα2(I) fusion protein with a non-native promoter and signal sequence. In this regard, we chose to employ primary human fibroblasts, which can be readily cultured *in vitro* and are one of the major producers of collagen-I in humans. After a 24 h treatment with Hsp90 inhibitor, immunoblotting revealed significant decreases in the level of endogenous collagen-I secretion relative to that of vehicle-treated cells (Figure 2.6B and D). We observed no changes in intracellular levels of collagen-I (Figure 2.6C). We also confirmed functional inhibition of cytosolic Hsp90 at the concentrations used, indicated both by the expected increase in Hsp70 levels due to induction of a compensatory heat shock response (HSR) and by the loss of signal from the cytosolic Hsp90 client Akt (Figure 2.6C).

To test whether the effects of Hsp90 inhibitor treatment on secretion of endogenous collagen-I by primary fibroblasts were cell line-specific, we obtained a second primary fibroblast line. As expected, 17-AAG treatment caused similar reductions in the level of secretion of collagen-I from these cells (Figure 2.6E), indicating that our results are not dependent on a particular primary fibroblast genotype.

17-AAG Reduces Collagen-I Secretion via a Post-Translational Effect. We next investigated the mechanism by which inhibition of Hsp90-type chaperones reduces the level of collagen-I secretion. Possible trivial explanations we considered were (1) that a reduced level of collagen-I secretion in primary fibroblasts is caused by compound cytotoxicity, (2) that Hsp90 inhibitors reduce the levels of collagen-I transcripts, and/or (3) that Hsp90 inhibitors lower the level of collagen-I translation. Our Saos-2 data (Figure 2.1) suggested that these explanations were not applicable, as we had already counterscreened for cytotoxicity and we employed a non-native promoter, but it was nonetheless important to test all three possibilities in primary cells.

We treated primary fibroblasts with 17-AAG for 24 h and then tested for cytotoxicity by staining cells with Annexin-V/Alexa Fluor 488 and propidium iodide followed by FACS analysis. Treatment with 17-AAG was not toxic by this assay, confirming that Hsp90 inhibitors do not decrease the level of secretion of collagen-I from primary fibroblasts simply by promoting cell death (Figure 2.7A). Next, we performed qPCR analysis of collagen-I transcripts from Hsp90
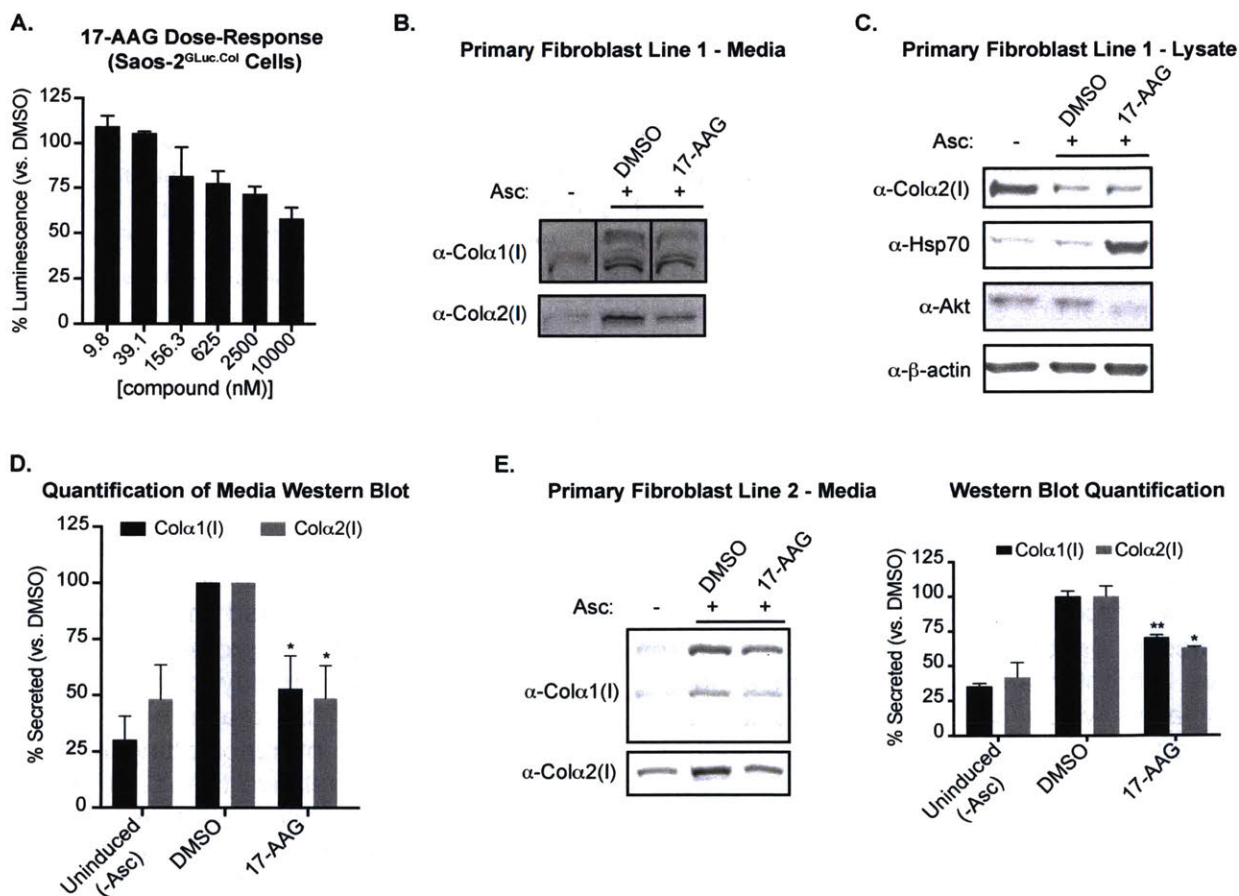
Figure 2.6: *17-AAG reduces endogenous collagen-I secretion.* (A) Dose–response curve for the screening hit 17-AAG in Saos-2$^{GLuc.Col}$ cells. Cells were plated in 384-well plates and co-treated 5–6 h after being plated with 1 μg/mL Dox and 17-AAG. After 23 h, plates were cooled at room temperature for 1–1.5 h, substrate was added, and luciferase activity was assayed. Western blot of conditioned media (B) and lysate (C) from primary dermal fibroblasts. Cells were plated and allowed to adhere overnight before receiving fresh media with 200 μM ascorbate and 250 nM 17-AAG. After being treated for 24 h, conditioned media and cells were harvested for immunoblotting. (D) Quantification of Western blot results. *$p <$ 0.05 ($n = 3$). (E) Western blot of conditioned media from a second healthy fibroblast line. *$p <$ 0.05; **$p < 0.01$ ($n = 3$).

A. Annexin V/PI Staining   **B. Collagen-I Transcript Levels**   **C. [³H]-Labeled Collagen Synthesis**
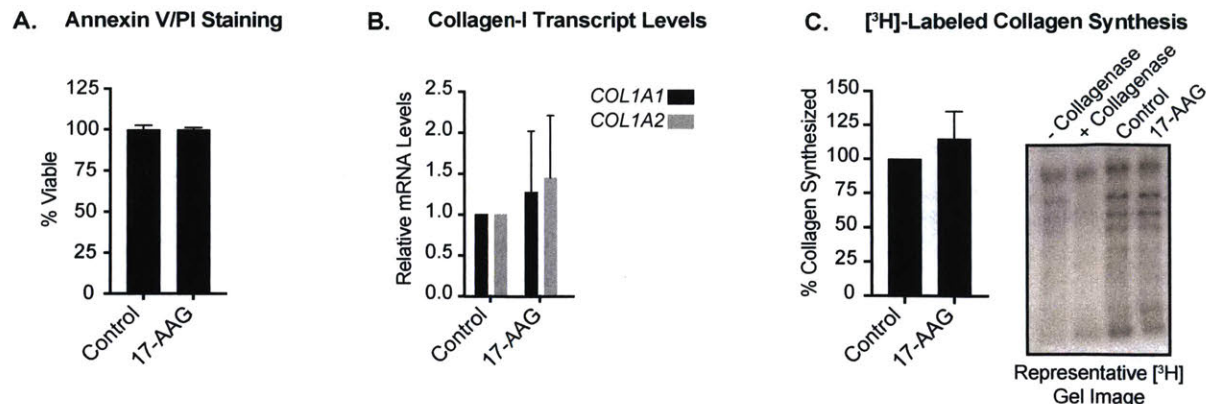
Figure 2.7: *Effects of Hsp90 inhibition on collagen-I secretion are post-translational.* (A) FACS analysis of Annexin-V/Alexa Fluor 488 and propidium iodide staining to detect cell death. Cells were plated in triplicate for each condition and allowed to adhere overnight before being treated as in Figure 2.6B. After 24 h, cells were trypsinized and labeled for FACS analysis. (B) qPCR analysis of collagen-I transcript levels from cells treated as in Figure 2.6B. Average values are shown with the standard deviation for each condition (n = 3). (C) Measurement of collagen-I synthesis by [2,3-³H]proline pulse. Cells treated for 6 h with ascorbate and 17-AAG (250 nM) were pulsed with [2,3-³H]proline for 30 min prior to being harvested and analyzed. Quantitation of newly synthesized collagen-I (mean ± standard deviation) is shown along with a representative gel image (n = 3).

inhibitor-treated primary fibroblasts. We found that *COL1A1* and *COL1A2* transcript levels are not significantly altered, indicating that the mechanism of action is post-transcriptional (Figure 2.7B). Finally, we used a tritium pulse assay to test whether collagen-I translation is altered by 17-AAG treatment. Briefly, we pulsed treated fibroblasts with [2,3-³H]proline for 30 min to label the proline-rich collagen-I protein, followed by collection of cell lysate and auto-radiography analysis of the collagen-I bands (identified by bacterial collagenase treatment). We observed no significant alterations in levels of newly synthesized, radiolabeled collagen-I upon 17-AAG treatment (Figure 2.7C). Cumulatively, these data indicate that 17-AAG reduces collagen-I secretion via a post-translational mechanism.

17-AAG Selectively Reduces Collagen-I, but Not General Protein, Secretion. Another potential explanation for our results is that Hsp90 inhibitors non-selectively reduce the level of protein secretion. We had established that 17-AAG treatment does not inhibit eGLuc2 secretion in Saos-2 cells (Figure 2.1), but it remained unclear whether effects on endogenous collagen-I secretion in primary fibroblasts would be selective. Indeed, prior work suggests that Hsp90 inhibitors can disrupt protein secretion in yeast[46] or inhibit ER-to-Golgi transport by perturbing Rab1 GTPase recycling in mammalian cells,[47] so we wondered if the effects of Hsp90 inhibitor treatment on collagen-I secretion might extend generally to other proteins.

To examine this question, we transduced primary fibroblasts with adenoviruses expressing FLAG-tagged transthyretin (TTR) or fibulin-3 (two well-characterized and soluble secreted proteins), treated the cells with 17-AAG at a concentration that reduces the level of collagen-I secretion, and assayed for altered secretion of these additional ER client proteins (Figure 2.8A). Treatment with 17-AAG did not alter TTR secretion, while fibulin-3 secretion was actually enhanced ~3-fold by 17-AAG treatment (Figure 2.8B). We next tested the secretion of fibronectin, an endogenous secreted protein that interacts with collagen-I but is secreted in globular form.[48, 49] Again, 17-AAG robustly decreased the level of collagen-I secretion but did

A. Experimental Workflow

B. TTR and Fibulin-3 Secretion

C. Fibronectin (FN) Secretion

D. Western Blot Quantification

E. Global Protein Secretion

Representative ^{35}S gel
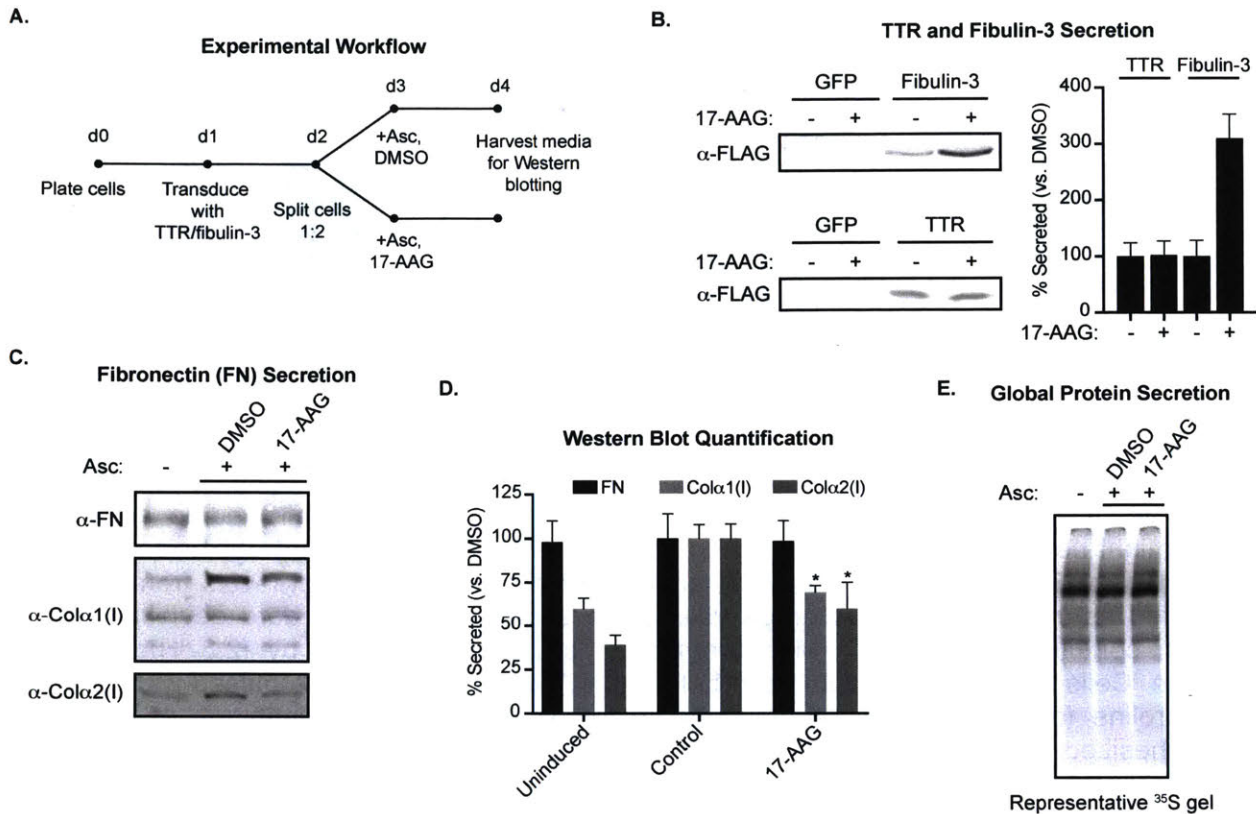
Figure 2.8: *17-AAG does not affect general protein secretion.* (A) Experimental workflow. Cells were plated and allowed to adhere overnight before being transduced with either GFP, transthyretin (TTR), or fibulin-3-expressing adenoviruses. The next day, transductions were split 1:2 and allowed to adhere overnight; cells were then treated with ascorbate and DMSO or 17-AAG (250 nM) for 24 h prior to being harvested. (B) Western blot of conditioned media from primary dermal fibroblasts transduced with FLAG-tagged transthyretin (TTR) or fibulin-3. (C) Immunoblot analysis of the secretion of endogenous fibronectin after treatment with 17-AAG. Cells were plated and treated as described in Figure 2.6B. (D) Quantification of Western blot results. *$p < 0.05$ ($n = 3$). (E) Phosphorimage of ^{35}S-labeled secreted proteins. Cells were plated in triplicate for each condition and allowed to adhere overnight before being pretreated with ascorbate and DMSO or 17-AAG (250 nM) for 2 h. Cells were then washed and cultured in radiolabeled media containing ascorbate, [^{35}S]-Met/Cys, and DMSO/17-AAG for 6 h. A representative radiograph is shown.

not alter fibronectin export (Figure 2.8C, D). Although these proteins are not representative of all possible secretory pathway cargo, their response to 17-AAG treatment, in contrast to what we observe for collagen-I from the same cells, strongly suggests that 17-AAG at the concentrations employed reduces the level of collagen-I secretion with considerable selectivity. However, because high concentrations of Hsp90 inhibitors had been reported to induce Golgi fragmentation,[47] we sought to more rigorously test whether global protein secretion was inhibited under our experimental conditions.

We first evaluated Golgi integrity using imaging techniques. We treated primary fibroblasts with ascorbate and DMSO or 17-AAG for 20 h and then stained cells with an antibody raised against the *cis*-Golgi marker GM-130.[47] Confocal microscopy showed that treatment with Brefeldin A resulted in a punctate α-GM-130 distribution, as expected (Figure

40
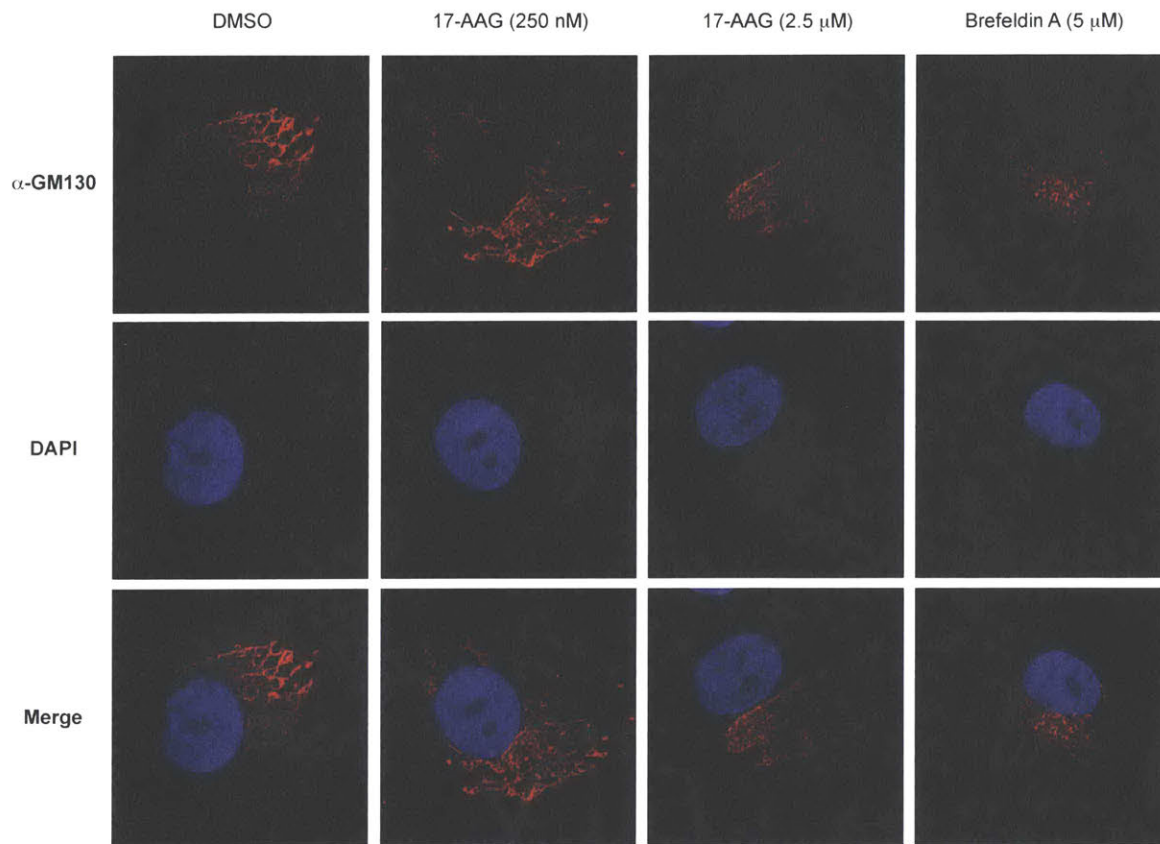
Figure 2.9: *Confocal microscopy analysis of Golgi structure with 17-AAG treatment.* GM05659 fibroblasts were plated and treated with ascorbate and DMSO or the indicated compounds for 20 h. Cells were then fixed and stained with α-GM130, a *cis*-Golgi marker, before imaging on a Zeiss spinning disk confocal microscope.

2.9). By contrast, 17-AAG did not significantly alter Golgi structure relative to that seen upon DMSO treatment, either at the concentration used in our collagen-I secretion experiments (250 nM) or at a 10-fold larger dose. We also performed electron microscopy on fibroblasts treated with ascorbate and 17-AAG for 6 or 18 h. In addition to the characteristic Golgi stacks, some regions of vesicular tubular networks were visible in all samples; however, neither 6 nor 18 h of 17-AAG treatment resulted in Golgi collapse or a loss of morphology relative to that seen upon DMSO treatment (Figure 2.10). Next, we employed radiolabeling to globally assess protein secretion. We pretreated fibroblasts with ascorbate and 17-AAG for 2 h and then incubated the cells with [^{35}S]Met and Cys-labeled medium, ascorbate, and 17-AAG for 6 h. The resulting phosphorimages, shown in Figure 2.8E, reveal that treatment with 17-AAG does not broadly affect the secretion of radiolabeled proteins in our system. On the basis of these results, we conclude that inhibiting Hsp90-type chaperones can reduce the level of secretion of collagen-I from primary fibroblasts without globally inhibiting protein trafficking.

17-AAG Treatment Does Not Result in Intracellular Collagen-I Deposits. Because we had observed significant reductions in collagen-I secretion in the absence of an apparent increase in intracellular collagen-I levels (Figure 2.6), we wondered what could account for the decreased collagen-I levels. To address this question, we treated primary fibroblasts with ascorbate and DMSO or 17-AAG for 6 h and performed confocal microscopy to visualize collagen-I (using the polyclonal antibody LF-68) and its co-localization with ER (PDI; Figure 2.11A), Golgi (GM-130; Figure 2.11B), and lysosomal (Lamp1; Figure 2.11C) markers. As expected given that we see a partial but not complete ablation of collagen-I secretion, collagen-I distribution as detected using LF-68 is broadly similar by this technique. One possibility not tested here is that collagen-I accumulates in plasma membrane-bound deposits, as has been observed for collagen IV in *Drosophila*;[50] alternatively, repeating the imaging analysis at multiple time points may reveal
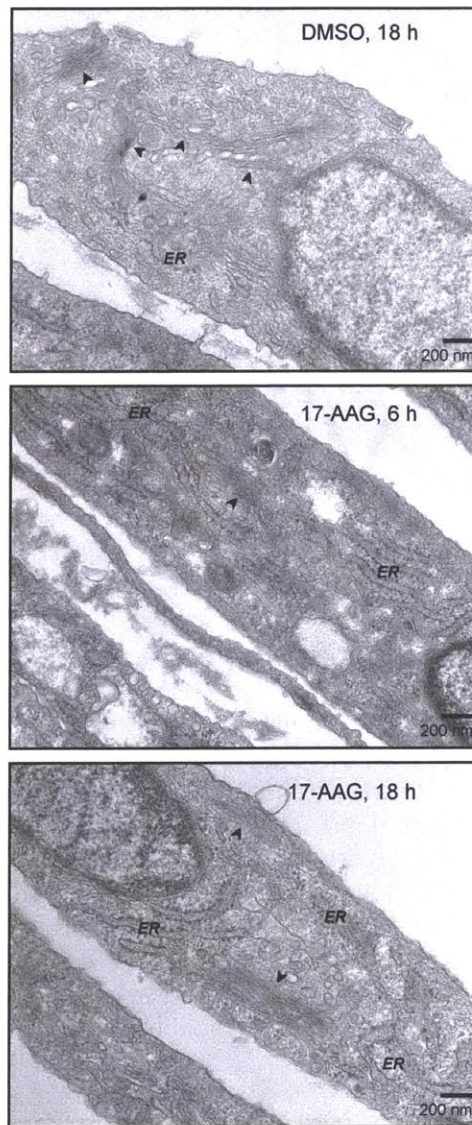
Figure 2.10: *Electron microscopy analysis of Golgi structure.* GM05659 fibroblasts were plated and treated with ascorbate and DMSO or 17-AAG (250 nM) for 18 h. A third plate was treated with ascorbate and 17-AAG for 6 h. Cells were then fixed, embedded, stained, and imaged on an FEI Tecnai Spirit equipped with an AMT CCD camera. Representative images are shown, and portions of rough ER are labeled with text; arrowheads indicate Golgi stacks.

effects of 17-AAG treatment on collagen-I localization. Yet while the mechanistic details of 17-AAG mediated decreases in collagen-I levels remain unknown, our data collectively confirm the ability of Hsp90-type chaperone inhibitors to selectively reduce collagen-I secretion independent of effects on cell viability, collagen-I synthesis, or general protein trafficking.
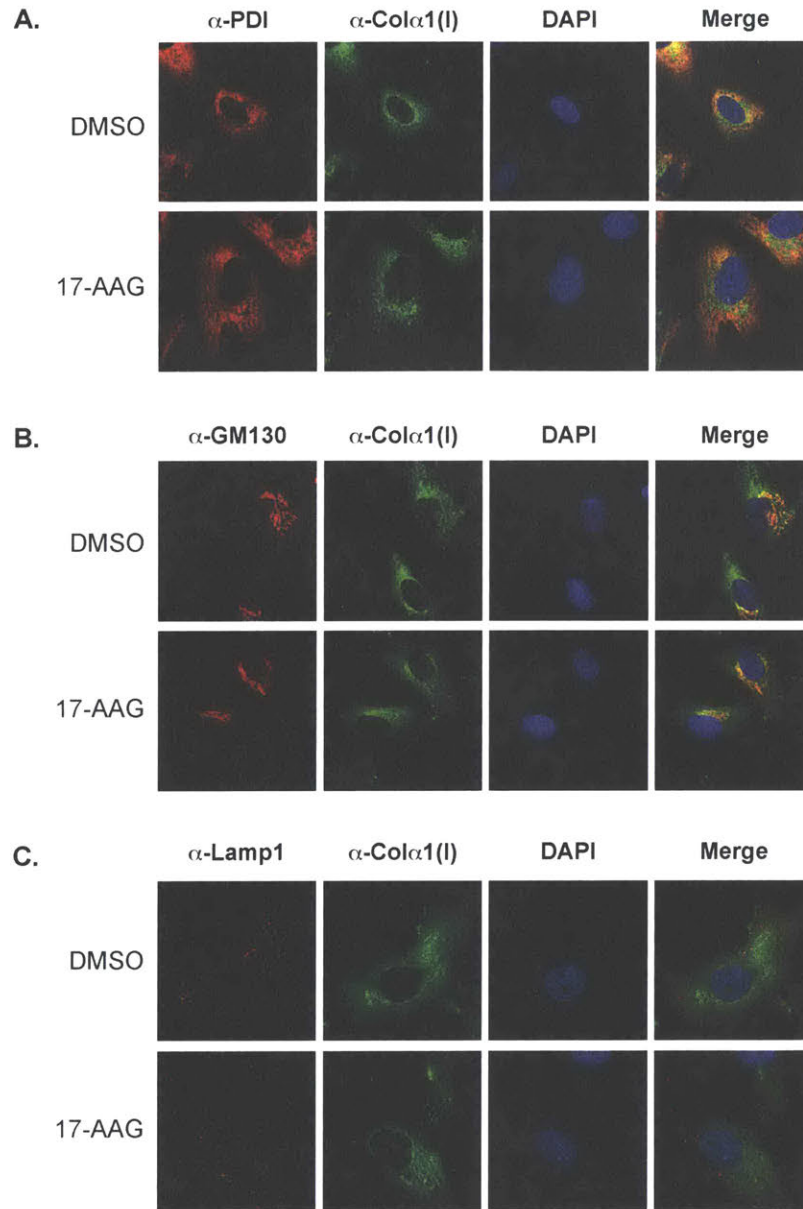
Figure 2.11: 17-AAG does not affect collagen-I localization. GM05659 fibroblasts were plated on poly-D-lysine-coated coverslips and allowed to adhere overnight. The next day, cells were treated for 6 h with 200 μM sodium ascorbate and 250 nM 17-AAG, and processed for confocal microscopy to visualize collagen-I (using the polyclonal antibody LF-68) and its co-localization with ER (PDI; (A)), Golgi (GM-130; (B)), and lysosomal (Lamp1; (C)) markers. As expected given that we see a partial but not complete ablation of collagen-I secretion, collagen-I distribution as detected using LF-68 is broadly similar by this technique.

43

The Broad-Spectrum Anti-Parasitic Nitazoxanide Also Reduces Collagen-I Secretion. Finally, we repurchased the anti-parasitic compound nitazoxanide (NTZ) and tested it for effects on endogenous collagen-I secretion by primary cells. While NTZ had not, in fact, met our fold-change or cytotoxicity cut-offs in the high-throughput screen, we nonetheless selected NTZ for retesting because it had shown opposite, if moderate, effects on eGLuc2.Colα2(I) and eGLuc2 secretion. We treated primary fibroblasts with 200 μM ascorbate and 10 μM NTZ for 24 h, and then harvested conditioned media for immunoblotting. Notably, NTZ showed a reduction in collagen-I secretion by primary fibroblasts, of equal (or larger, for Colα2(I)) magnitude to that observed in our screening cell line (Figure 2.12A and B).

We were concerned that these results might be primarily due to cytotoxicity, and so we tested NTZ for effects on cell viability using the same FACS assay employed previously. We found that NTZ did not reduce cellular viability under the conditions tested (Figure 2.12C), consistent with reports of high pharmacological tolerability.[51] We also used the tritium pulse assay to test effects of NTZ treatment on collagen-I synthesis. As shown in Figure 2.12D, we did not detect any significant change in the amount of newly synthesized collagen-I from cells treated with NTZ. Together, our results suggest that NTZ can also reduce collagen-I secretion in a post-translational manner, and motivate further studies to identify its mechanism of action.

Discussion

Although 17-AAG did not alter global protein secretion or the export of transthyretin or fibronectin, the Hsp90 inhibitor had significant and opposite effects on collagen-I (~50% decrease) and fibulin-3 (~3-fold increase) secretion. Intriguingly, 17-AAG had opposite effects on eGLuc2 and eGLuc2.Colα2(I) secretion in our screening cell line, similar to what we observe with fibulin-3 and collagen-I in primary fibroblasts. The diversity of 17-AAG effects on secreted proteins examined here suggests more nuanced rules for secretion than size or shape. The precise mechanism by which 17-AAG decreases collagen-I secretion may therefore prove useful for understanding how Hsp90-type chaperones participate in protein export;[46, 47] our efforts on this front are discussed in Chapter III.

We note that ~50% reductions in collagen-I levels, on the order of what we observe in primary fibroblasts, are sufficient to reduce the fibrotic area in tissue sections from mouse models,[42] suggesting that the effect sizes we detect upon Hsp90 inhibition could be sufficient to reduce disease burden. Moreover, the level of collagen-I secretion is reduced by 17-AAG in some cases to levels similar to that of the ascorbate-untreated control and thus likely approaches the realistic limit for collagen-I secretion inhibition in our model system. However, recent reports that collagen stiffness, rather than concentration, determines disease severity suggest that co-treatment with collagen secretion and collagen cross-linking inhibitors may prove most effective as a therapeutic strategy.[52]

In addition to the specific insights we obtained regarding the capacity of Hsp90 inhibitors to reduce the level of collagen-I secretion, we note that as a large, multimeric protein that simultaneously lacks enzymatic activity and requires highly specialized cellular folding machinery, collagen is a challenging protein to assay. We believe the cell-based, high-throughput screen for collagen-I secretion designed and validated here provides a viable solution to this challenge that will prove useful on numerous accounts. Indeed, in addition to the Hsp90 inhibitors we also identified six other functionally and structurally distinct small molecules that selectively affect collagen-I secretion. These validated hits include two structurally related EGFR tyrosine kinase inhibitors (AG1478 and gefitinib), a Ras farnesyl transferase inhibitor (L-744,832), a dopamine D2 receptor agonist (bromocriptine mesylate), and mitomycin C. While most of these compounds have not been tested in collagen-related disease models, mitomycin C was previously used as a wound healing adjuvant, where it was found to slow the healing process but also to reduce the level of fibrosis.[53-55] Follow-up studies using these compounds
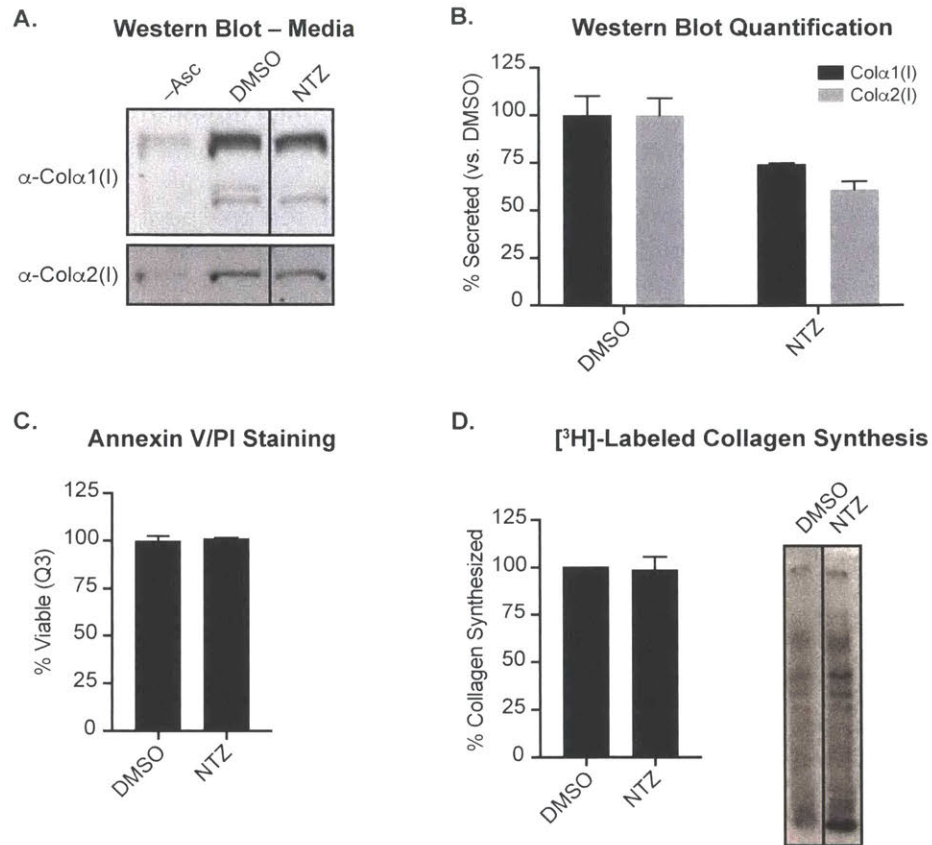
Figure 2.12: *Nitazoxanide also reduces collagen-I secretion without affecting cell viability or collagen-I synthesis.* GM05659 fibroblasts were plated and allowed to adhere overnight before receiving fresh media with 200 μM ascorbate and 10 μM nitazoxanide (NTZ; Sigma). After being treated for 24 h, conditioned media and cells were harvested for immunoblotting. Representative gel images and quantification are shown in (A) and (B), respectively. (C) FACS analysis of Annexin-V/Alexa Fluor 488 and propidium iodide staining to detect cell death. Cells were plated in triplicate and allowed to adhere overnight before being treated as in (A). After 24 h, cells were trypsinized and labeled for FACS analysis. (D) Measurement of collagen-I synthesis by [2,3-^{3}H]proline pulse. Cells treated for 6 h with ascorbate and NTZ (10 μM) were pulsed with [2,3-^{3}H]proline for 30 min prior to being harvested and analyzed. Quantitation of newly synthesized collagen-I (mean ± standard deviation) is shown along with a representative gel image (*n* = 3).

are expected to provide additional insights into mechanisms of collagen-I biosynthesis and secretion, as well as new potential opportunities for therapeutic intervention.

The anti-parasitic compound NTZ has been linked to many diseases beyond its original application, with proposed activity as an antiviral and anticancer drug.[56, 57] The breadth of reported uses might suggest that NTZ is a promiscuous small molecule with multiple off-target effects, but NTZ nonetheless appears to be well tolerated in humans and has been used for treating *Giardia lamblia* or *Cryptosporidium parvum* infection since receiving FDA approval in 2002.[51, 58] Among its proposed mechanisms of action are parasite membrane depolarization, pyruvate:ferredoxin oxidoreductase inhibition, depletion of ER- or Golgi-Ca^{2+} stores, induction of PKR and eIF2α phosphorylation, and stimulation of autophagy; inhibition of viral protein maturation; and inhibition of β-catenin signaling.[51, 57-61] In addition to these processes, the thiol

oxidoreductase ERp57 was recently identified as a possible target of NTZ.[62] Using both recombinant and endogenous ERp57, the authors showed that NTZ acts as a non-competitive inhibitor of ERp57's non-catalytic domains, blocking paramyxovirus infection by causing aggregation of a viral protein required for host cell entry. Notably, the proposed drug-binding site mediates association of ERp57 with the chaperones calnexin and calreticulin, consistent with a separate report of calreticulin as a putative NTZ binder.[57]

While the other mechanisms cannot yet be ruled out, the prospect that NTZ functions as an ERp57 inhibitor in human cells presents exciting opportunities for future studies. A proteomic study with engineered cell lines previously identified ERp57 as a novel collagen-I interactor.[32] Testing effects of NTZ treatment on secretion of the collagen-I C-terminal propeptide domain alone, which like the paramyxoviral F-protein utilizes both inter- and intra-chain disulfide bonds to stabilize a protein trimer, could thus help clarify whether the preliminary collagen-I secretion results obtained here are due to a direct effect of NTZ on ERp57.

In summary, we screened 1228 compounds against our cell-based, high-throughput assay and identified six selective small molecule modulators of collagen-I secretion. Using the compound 17-AAG identified in our screen, we further validate inhibition of Hsp90-type chaperones as a viable antifibrotic strategy[42-45, 63-65] that points to a possible role for Hsp90 in post-translationally regulating collagen-I secretion. Our screen design employs cells that natively synthesize collagen-I and uses stably engineered cell lines for improved convenience and consistency. Our high-throughput assay can also be readily adapted to study disease-causing, misfolding variants of collagen-I by mutagenizing the wild-type Colα2(I) gene in our eGLuc2.Colα2(I) fusion protein, and we expect this flexibility to provide a valuable addition to the suite of current tools to study the collagenopathies.

Materials and Methods

Cell Lines and Reagents. Saos-2^{TREx} cells[32] were grown in complete DMEM (Corning) supplemented with 15% heat-inactivated fetal bovine serum (FBS), 100 IU penicillin, 100 μg/mL streptomycin, and 2 mM L-glutamine (Corning). Healthy dermal fibroblasts (GM05659, denoted "Line 1", or GM05294, denoted "Line 2"; Coriell Cell Repositories) were cultured in complete MEM supplemented with 15% FBS, 100 IU penicillin/100 μg/mL streptomycin, and 2 mM L-glutamine (Corning). Expression of eGLuc2 and eGLuc2.Colα2(I) was induced in stable Saos-2 lines by treating cells with 1 μg/mL doxycycline (Dox; Alfa Aesar). Secretion of endogenous collagen-I was induced in Saos-2 cells or primary fibroblasts using 50 or 200 μM sodium ascorbate (Amresco), respectively. The ICCB Bioactives and FDA-approved compound decks were purchased from Enzo as DMSO stock solutions in 96-well plate format and stored at −80 °C for <2 years prior to use. For screening, the compounds were rearrayed in V-bottom 96- and 384-well plates. Compound plates were thawed overnight at room temperature before being pinned and heat-sealed before being refrozen at −20 °C. The Alexa Fluor 488 Annexin V/Dead Cell Apoptosis Kit with Alexa Fluor 488 Annexin V and propidium iodide was obtained from Life Technologies. Luciferase assays were performed either with homemade buffers[66] containing an assay stabilizer from the commercially available Biolux kit (New England Biolabs) and 10 μg/mL coelenterazine (prepared from a 1 mg/mL stock in acidified methanol; Biotium) for high-throughput screening or using the Biolux kit (with a commercial stabilizer) for dose−response curves. Immunoblots were probed with the following primary antibodies, as indicated: GLuc (New England 126 Biolabs E8023), Colα1(I) (DSHB SP1.D8), Colα2(I) (Abcam ab96723), Colα1(I) (NIH LF-68), Hsp70/Hsp72 (Cell Signaling Technologies 4873), Akt (Cell Signaling Technologies 9272), β-actin (Sigma A1978), fibronectin (Sigma F7387), FLAG (Agilent 131 200473).

Vector and Stable Cell Line Construction. The gene encoding eGLuc2 was amplified via polymerase chain reaction (PCR) from an eGLuc2.pENTR1A plasmid[23] to remove the stop codon and then reinserted into the pENTR1A plasmid. COL1A2 (UniProt entry P08123) was cut from a previously developed PPT.FLAG.Colα2(I).pENTR1A plasmid[32] and inserted after eGLuc2 using the NotI and EcoRV sites to create the eGLuc2.Colα2(I).pENTR1A plasmid. eGLuc2 and eGLuc2.Colα2(I) were then recombined into pLenti.CMV.TO.DEST Gateway destination vectors[67] using LR clonase (Life Technologies). Lentiviruses were generated as described previously[68] and used to transduce Saos-2-TREx cells. Stable cells were selected with 250 μg/mL hygromycin B and propagated in the same, with the addition of 10 μg/mL blasticidin to maintain the tetracycline repressor. Single colonies inducibly expressing moderate levels of eGLuc2.Colα2(I) (termed Saos-2$^{GLuc.Col}$ cells) were selected using immunoblotting.

High-Throughput Screening and Dose-Response Curves. Saos-2$^{GLuc.Col}$ cells were plated in flat-bottom, white 384-well plates (Corning) at a density of 7000 cells/well with an EL406 Washer Dispenser (BioTek) in a total volume of 50 μL of complete medium; 5–6 h postplating, 1 μL of a 50 μg/mL Dox stock solution was added, and cells were treated with 50 nL of compound using a Freedom Evo 150 Liquid Handler (Tecan) fitted with a floating, slotted pin tool (V&P Scientific). Plates were incubated for 23 h at 37 °C, followed by equilibration for 1–1.5 h at room temperature in a single layer before reading. Ten microliters of assay buffer was then dispensed into the wells using a plate washer; plates were mixed by orbital shaking for 5 s, and the luminescence signal was read using a Thermo Infinite M1000 Plate Reader (Tecan) with a 100 ms integration time, beginning 35 s after buffer addition. Saos-2^{GLuc} cells were plated at a density of 3000 cells/well prior to screening using an otherwise identical protocol. For cytotoxicity counterscreening, intracellular ATP was quantified after the same time courses of Dox and compound treatments using the CellTiter-Glo Assay (CTG; Promega). Briefly, 10 μL of CTG buffer was added to the cells; the plate was incubated for 10 min at room temperature, and then the luminescence was read using an integration time of 100 ms. All screens were performed in biological duplicate. Screening results were filtered for cytotoxicity using three standard deviations below the average of DMSO-treated cells as a cutoff. A Z' factor of 0.72 was approximated for the assay by calculating the separation between the highest and lowest compound-treated wells, according to Zhang et al.[69] The coefficient of variation (% CV) was calculated by taking the ratio of the SD to the average DMSO signal for each cell line. Complete raw screening results are presented in Table S1 of ref. [75].

Co-Immunoprecipitation Experiments. Saos-2$^{GLuc.Col}$ cells were plated in 10 cm dishes at a density of 1.25×10^6 cells/plate and allowed to adhere overnight. The next day, the medium was changed and cells were treated with ascorbate with or without 1 μg/mL Dox. After 24 h, cells received fresh medium and were retreated with fresh ascorbate and Dox. After an additional 24 h, the medium was removed, and cells were trypsinized, pelleted, washed with 1× phosphate-buffered saline (PBS), and then lysed in Triton-X lysis buffer [200 mM NaCl, 1% Triton X-100, 50 mM Tris (pH 7.5), 1 mM EDTA, and 1.5 mM MgCl$_2$] with 1.5 mM phenylmethanesulfonyl fluoride (PMSF; Sigma) and a protease inhibitor tablet (Thermo Fisher). Lysates were cleared by centrifugation at 21.1×10^3 ×g and 4 °C for 15 min; total protein was quantified using the bicinchoninic acid assay, and 2 mg of lysate was incubated with 5 μL of α-GLuc antibody (NEB) in a 500 μL volume end over end at 4 °C. After 4 h, 30 μL of a Protein A/G PLUS-Agarose bead slurry (Santa Cruz) was added and samples were incubated end over end at 4 °C overnight. The next day, samples were washed three times with Triton-X lysis buffer, dried with a 30.5 gauge needle, and eluted by being boiled in Laemmli buffer supplemented with 167 mM 1,4-dithiothreitol (DTT; Sigma). Samples were separated on 4/8/12% Tris-Gly sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) gels, with 2.5% of unprecipitated

47

lysate as input and 50% of the elution, before being transferred to nitrocellulose membranes (Bio-Rad). Blots were probed using an anti-GLuc antibody or SP1.D8, a mouse monoclonal antibody against the N-terminal propeptide of Colα1(I) (DSHB).

Precipitation of Collagen-I from Conditioned Media. Saos-2$^{GLuc.Col}$ cells were plated in 10 cm dishes at a density of 1.25×10^6 cells/plate and allowed to adhere overnight. The next day, the medium was changed and cells were treated with ascorbate either with or without 1 μg/mL Dox. After 24 h, cells were retreated with fresh ascorbate and Dox. After an additional 24 h, the conditioned medium was transferred to falcon tubes and spun at 228 ×g for 5 min to pellet cell debris. The supernatants were then transferred to new tubes, and collagen-I was precipitated according to the procedure described by Makareeva and co-workers.[70]

Disulfide-Dependent Assembly Assays. Precipitated collagen-I pellets were thawed on ice, resuspended in 80 μL of 0.1 M NaHCO$_3$ (pH 8.3), and split into Eppendorf tubes (35 μL each); 3.5 μL of 200 mM DTT in 0.1 M NaHCO$_3$ (pH 8.3) was then added (to give a final concentration of 20 mM), and samples were heated at 60 °C for 15 min. Nonreduced samples were treated with an equal volume of 0.1 M NaHCO$_3$ (pH 8.3) before being heated. After reduction, samples were cooled at room temperature for 15 min, and then 4.3 μL of 600 mM iodoacetamide (IAA; VWR) in NaHCO$_3$ buffer was added to yield a final concentration of 60 mM. Samples were incubated in the dark with IAA at room temperature for 1 h, mixed with 8.6 μL of Laemmli buffer, heated, and separated on 4/7% Tris-Gly SDS−PAGE gels. Proteins were then transferred to nitrocellulose membranes, which were developed with anti-GLuc and SP1.D8.

Pepsin Digestion. Pepsin digestions were performed as described by Makareeva et al.[70] Precipitated collagen-I pellets were thawed on ice and resuspended in 1 mL of pre-chilled (4 °C) 0.5 M acetic acid. Samples were then split into Eppendorf tubes (450 μL each) and combined with 50 μL of 0.5 M acetic acid with or without 2 mg/mL pepsin (Amresco). Samples were inverted gently to mix and incubated at 4 °C for 36 h. Pepsin-stable collagen-I was then precipitated by adding 110 μL of pre-chilled (4 °C) 5 M NaCl, to give a final concentration of 0.9 M, and incubating end-over-end at 4 °C for 2 h. Tubes were then centrifuged at 21.1×10^3 ×g at 4 °C for 10 min, inverted, and centrifuged again for 10 min. The supernatant was removed, pellets were dissolved by vortexing in Laemmli buffer containing 1 M DTT, heated, and separated on 4/7% Tris-Gly SDS-PAGE gels. Bands were visualized by Coomassie staining.

Western Blotting Analysis. GM05659 primary fibroblasts were plated in 6-well plates at a density of 2×10^5 cells/well and allowed to adhere overnight. The next day, the medium was changed to full MEM with 0.5× penicillin/streptomycin, collagen expression was induced with 200 μM sodium ascorbate (Amresco), and cells were treated with DMSO or the indicated compound concentrations for 24 h. Samples of medium were denatured by being boiled in Laemmli buffer (supplemented with 167 mM DTT), separated on homemade 4/8% Tris-Gly SDS−PAGE gels, and transferred to nitrocellulose membranes. Because of cross-reactivity between the collagen-I antibodies, samples of medium for collagen-I blots were run on duplicate gels, transferred, and then probed separately for either rabbit anti-Colα2(I) or rabbit anti-Colα1(I).

 For Western blots of cell lysates, cells were trypsinized, washed with PBS, and lysed in Triton-X lysis buffer with 1.5 mM PMSF and a protease inhibitor tablet (Pierce). Total protein was quantified using the bicinchoninic acid assay, and equal amounts of total protein were analyzed for each sample. Each experiment was performed in biological triplicate. Blots were imaged after incubation with appropriate primary and 800CW or 680LT secondary antibodies (LI-COR) by scanning on an Odyssey infrared imager (LI-COR), followed by quantification using ImageJ.

Fluorescence-Assisted Cell Sorting. GM05659 primary fibroblasts were plated at a density of 1 × 10^6 cells/plate in 6-cm dishes and allowed to adhere overnight. The next day, cells were treated as described for Western blotting. After being treated with the compound for 24 h, cells were harvested and labeled with an Annexin V-Alexa Fluor 488 conjugate and propidium iodide (Life Technologies) according to the manufacturer's instructions and analyzed on a FACSAria sorter (BD Biosciences) using a flow rate of 2.7 mL/h with 2 × 10^4 events gated to exclude cell debris.

Quantitative PCR (qPCR). GM05659 primary fibroblasts were plated and treated as described above for Western blotting. After being treated with the compound for 24 h, cells were washed with PBS, and RNA was extracted using the Omega E.Z.N.A. Total RNA extraction kit. cDNA was prepared from 500 ng of RNA, normalized for all samples in each run, using an Applied Biosystems Reverse Transcriptase cDNA Kit in a Bio-Rad Thermocycler. Samples were run on a Light Cycler 480 II Real Time PCR Instrument in the MIT BioMicro Center using previously described primers[32, 71] and analysis methods.[68]

Pulse Labeling. Collagen translation was assessed by [2,3-^{3}H]proline labeling. GM05659 primary fibroblasts were plated at a density of 1.2 × 10^5 cells/well in 12-well plates and allowed to adhere overnight. The next day, cells were treated as described for Western blotting. After a 6 h treatment, the medium was removed, and cells were washed three times with PBS and then serum-starved in MEM with 15% dialyzed FBS and 2 mM L-glutamine. Medium was changed to radiolabeled medium (100 μCi of [2,3-^{3}H]proline/mL; PerkinElmer) for 30 min, after which cells were washed three times with full, nonradiolabeled MEM and harvested. Both ascorbate and 17-AAG (250 nM) were included during the 30 min serum starvation and 30 min pulse. Harvested samples were washed with PBS, lysed in standard RIPA buffer containing protease inhibitor tablets and 1.5 mM PMSF, and run on 4/6% Tris-Gly SDS−PAGE gels. Gels were dehydrated, incubated with a scintillant solution [20% (w/w) 2,5-diphenyloxazole in DMSO; Sigma], washed, and then dried as previously described.[72] Results were quantified in ImageJ.

Adenoviral Production. Adenoviral vectors encoding FLAG-tagged TTR or FLAG-tagged fibulin-3[73] were used to produce adenovirus according to the manufacturer's instructions (LifeTechnologies ViraPower Adenoviral Expression System). The FLAG-tagged TTR virus was constructed using lambda red recombineering and ccdB[74] counterselection to insert the TTR gene from a pcDNA3.1 vector into the E1 region of adenovirus. Viruses were subsequently amplified in 293A cells and titrated with GM05659 cells to select an appropriate volume to use for experiments.

Transthyretin and Fibulin-3 Secretion. GM05659 primary fibroblasts were plated at a density of 2 × 10^5 cells/well in 6-well plates the day before transduction. After being allowed to adhere overnight, cells received fresh media and were transduced with adenoviruses encoding GFP, transthyretin (TTR), or fibulin-3. The next day, cells were split 1:2 into 12-well plates. Collagen-I expression was induced the following day as described for Western blotting, and cells were treated with either DMSO or 250 nM 17-AAG. After 24 h, the media was collected, separated on 4/8% (collagen-I), 4/10% (fibulin-3), or 4/15% (TTR) Tris-Gly SDS−PAGE gels, and probed for collagen-I or FLAG.

Metabolic Labeling of Secreted Proteins. GM05659 primary fibroblasts were plated at a density of 2 × 10^5 cells/well in 6-well plates and allowed to adhere overnight. The next day, cells received fresh medium containing ascorbate and DMSO or 17-AAG (250 nM) and were returned to the incubator for 2 h. After pretreatment, the medium was then replaced with [^{35}S]Met and

Cys-containing medium: 100 μCi of EXPRESS Protein Labeling Mix [^{35}S] (Sigma) per mL of Met- and Cys-free DMEM (Corning) supplemented with 15% dialyzed FBS and 2 mM L-glutamine. Cells were retreated with ascorbate and DMSO/17-AAG and incubated for 6 h. Plates were then removed to ice, and then samples of medium were transferred to Eppendorf tubes, spun at 207g and 4 °C for 10 min to pellet cell debris, and transferred to new tubes. Samples of medium were separated on 4/8% Tris-Gly SDS–PAGE gels and dried on a gel slab dryer for 75 min at 80 °C. Dried gels were developed with a GE Healthcare PhosphoScreen and imaged the next day on a Typhoon FLA 7000 instrument.

Confocal Microscopy. GM05659 primary fibroblasts were plated in 24-well plates on poly-D-lysine-coated coverslips (Chemglass Life Sciences, NJ) at a density of 4 × 10^4 cells/well and allowed to adhere overnight. The next day, cells were treated as described for Western blotting and incubated with DMSO, 17-AAG (250 nM or 2.5 μM), or Brefeldin A (5 μM) for 20 h. Media was then removed and cells were fixed in 4% paraformaldehyde / PBS for 30 min at rt. Cells were washed 4× with PBS, blocked in 2% bovine serum albumin (BSA) / PBS for 30 min at rt, and incubated with α-GM130 (BD Biosciences #610822) at a 1:100 dilution in 2% BSA/PBS overnight at 4 °C; a subset of samples received only 2% BSA/PBS to control for non- specific secondary staining. The next day, cells were washed again 4× with PBS and incubated with Alexa Fluor 568 goat-anti mouse (Life Technologies) at a 1:1000 dilution in 2% BSA/PBS for 2 h at rt in the dark. Cells were then washed 4× with PBS, incubated with DAPI at a 1:1000 dilution in PBS for 10 min at rt, washed again 5× with PBS, and mounted on slides in ProLong Gold Antifade Mountant (Life Technologies). Slides were cured overnight at rt before imaging with a 100X 1.40NA oil objective mounted onto a Zeiss RPI spinning disk confocal microscope equipped with MetaMorph acquisition software and a Hamamatsu ORCA-ER CCD camera. Representative images were chosen from 10 images per slide in biological triplicate.

Electron Microscopy. GM05659 primary fibroblasts were plated in 10-cm dishes at a density of 1.6 × 10^6 cells per plate and allowed to adhere overnight. The next day (18 h before fixation), two plates received fresh media containing ascorbate and either 0.1% DMSO or 250 nM 17-AAG. The day after (6 h before fixation), the third plate received fresh media containing ascorbate and 250 nM 17-AAG. Prior to fixation, an aliquot of media was taken from each plate to confirm 17-AAG reductions in collagen-I secretion by immunoblot. Samples were then removed from the incubator and immediately fixed in 2.5% glutaraldehyde, 3% paraformaldehyde with 5% sucrose in 0.1M sodium cacodylate buffer (pH 7.4), pelleted, and post-fixed in 1% OsO$_4$ in veronal-acetate buffer. The cells were stained en block overnight with 0.5% uranyl acetate in veronal-acetate buffer (pH 6.0), then dehydrated and embedded in Embed-812 resin. Sections were cut on a Leika EM UC7 microtome with a Diatome diamond knife at a thickness setting of 50 nm, stained with uranyl acetate, and lead citrate. The sections were examined using an FEI Tecnai Spirit at 80KV and photographed with an AMT CCD camera in the W. M. Keck Microscopy Facility at the Whitehead Institute at MIT.

Statistical Analyses. Data are presented as the mean and standard deviation from at least three biological replicates, with the exception of screening data, which were generated in biological duplicate. Statistical significance was assessed on normalized data from immunoblots using an unpaired Student's t-test (unequal variances), with the following significance thresholds: *$p < 0.05$, and **$p < 0.01$.

Funding

References

[1] Brinckmann, J. (2005) Collagens at a Glance, In *Collagen: Primer in Structure, Processing and Assembly* (Brinckmann, J., Notbohm, H., and Müller, P. K., Eds.), pp 1-6, Springer Berlin Heidelberg, Berlin, Heidelberg.

[2] Ricard-Blum, S. (2011) The collagen family, *Cold Spring Harbor Perspect. Biol. 3*, a004978.

[3] An, B., Lin, Y. S., and Brodsky, B. (2016) Collagen interactions: Drug design and delivery, *Adv. Drug Deliv. Rev. 97*, 69-84.

[4] Heino, J. (2007) The collagen family members as cell adhesion proteins, *Bioessays 29*, 1001-1010.

[5] Ishikawa, Y., and Bachinger, H. P. (2013) A molecular ensemble in the rER for procollagen maturation, *Biochim. Biophys. Acta 1833*, 2479-2491.

[6] Jobling, R., D'Souza, R., Baker, N., Lara-Corrales, I., Mendoza-Londono, R., Dupuis, L., Savarirayan, R., Ala-Kokko, L., and Kannu, P. (2014) The collagenopathies: Review of clinical phenotypes and molecular correlations, *Curr. Rheum. Rep. 16*, 394.

[7] Rosenbloom, J., Castro, S. V., and Jimenez, S. A. (2010) Fibrotic diseases: Cellular and molecular mechanisms and novel therapies, *Ann. Int. Med. 152*, 159-166.

[8] Provenzano, P. P., Inman, D. R., Eliceiri, K. W., Knittel, J. G., Yan, L., Rueden, C. T., White, J. G., and Keely, P. J. (2008) Collagen density promotes mammary tumor initiation and progression, *BMC Med. 6*, 11.

[9] Gilkes, D. M., Chaturvedi, P., Bajpai, S., Wong, C. C., Wei, H., Pitcairn, S., Hubbi, M. E., Wirtz, D., and Semenza, G. L. (2013) Collagen prolyl hydroxylases are essential for breast cancer metastasis, *Cancer Res. 73*, 3285-3296.

[10] Lu, P. F., Takai, K., Weaver, V. M., and Werb, Z. (2011) Extracellular matrix degradation and remodeling in development and disease, *Cold Spring Harbor Perspect. Biol. 3*, a005058.

[11] Bateman, J. F., Boot-Handford, R. P., and Lamande, S. R. (2009) Genetic diseases of connective tissues: Cellular and extracellular effects of ECM mutations, *Nat. Rev. Genet. 10*, 173-183.

[12] Forlino, A., and Marini, J. C. (2016) Osteogenesis imperfecta, *Lancet 387*, 1657-1671.

[13] Wynn, T. A. (2008) Cellular and molecular mechanisms of fibrosis, *J. Pathol. 214*, 199-210.

[14] Marini, J. C., Forlino, A., Cabral, W. A., Barnes, A. M., San Antonio, J. D., Milgrom, S., Hyland, J. C., Korkko, J., Prockop, D. J., De Paepe, A., Coucke, P., Symoens, S., Glorieux, F. H., Roughley, P. J., Lund, A. M., Kuurila-Svahn, K., Hartikka, H., Cohn, D. H., Krakow, D., Mottes, M., Schwarze, U., Chen, D., Yang, K., Kuslich, C., Troendle, J., Dalgleish, R., and Byers, P. H. (2007) Consortium for osteogenesis imperfecta mutations in the helical domain of type I collagen: regions rich in lethal mutations align with collagen binding sites for integrins and proteoglycans, *Hum. Mutat. 28*, 209-221.

[15] Di Lullo, G. A., Sweeney, S. M., Korkko, J., Ala-Kokko, L., and San Antonio, J. D. (2002) Mapping the ligand-binding sites and disease-associated mutations on the most abundant protein in the human, type I collagen, *J. Biol. Chem. 277*, 4223-4231.

[16] Vasta, J. D., Andersen, K. A., Deck, K. M., Nizzi, C. P., Eisenstein, R. S., and Raines, R. T. (2016) Selective inhibition of collagen prolyl 4-hydroxylase in human cells, *ACS Chem. Biol. 11*, 193-199.

[17] Harwood, R., Grant, M. E., and Jackson, D. S. (1976) The route of secretion of procollagen. The influence of alphaalpha'-bipyridyl, colchicine and antimycin A on the secretory process in embryonic-chick tendon and cartilage cells, *Biochem. J. 156*, 81-90.

[18] Ito, S., Ogawa, K., Takeuchi, K., Takagi, M., Yoshida, M., Hirokawa, T., Hirayama, S., Shin-Ya, K., Shimada, I., Doi, T., Goshima, N., Natsume, T., and Nagata, K. (2017) A small-molecule compound inhibits a collagen-specific molecular chaperone and could represent a potential remedy for fibrosis, *J. Biol. Chem. 292*, 20076-20085.

[19] Tullberg-Reinert, H., and Jundt, G. (1999) In situ measurement of collagen synthesis by human bone cells with a Sirius Red-based colorimetric microassay: effects of transforming growth factor beta 2 and ascorbic acid 2-phosphate, *Histochem. Cell Biol. 112*, 271-276.

[20] Le, B. Q., Fernandes, H., Bouten, C. V. C., Karperien, M., van Blitterswijk, C., and de Boer, J. (2015) High-throughput screening assay for the identification of compounds enhancing collagenous extracellular matrix production by ATDC5 cells, *Tissue Eng. Part C - Methods 21*, 726-736.

[21] Okano-Kosugi, H., Matsushita, O., Asada, S., Herr, A. B., Kitagawa, K., and Koide, T. (2009) Development of a high-throughput screening system for the compounds that inhibit collagen-protein interactions, *Anal. Biochem. 394*, 125-131.

[22] Bagchi, R. A., Mozolevska, V., Abrenica, B., and Czubryt, M. P. (2015) Development of a high throughput luciferase reporter gene system for screening activators and repressors of human collagen I alpha 2 gene expression, *Can. J. Physiol. Pharm. 93*, 887-892.

[23] Hulleman, J. D., Brown, S. J., Rosen, H., and Kelly, J. W. (2013) A high-throughput cell-based Gaussia luciferase reporter assay for identifying modulators of fibulin-3 secretion, *J. Biomol. Screen. 18*, 647-658.

[24] Tannous, B. A. (2009) Gaussia luciferase reporter assay for monitoring biological processes in culture and in vivo, *Nat. Protoc. 4*, 582-591.

[25] Maguire, C. A., Deliolanis, N. C., Pike, L., Niers, J. M., Tjon-Kon-Fat, L. A., Sena-Esteves, M., and Tannous, B. A. (2009) Gaussia luciferase variant for high-throughput functional screening applications, *Anal. Chem. 81*, 7102-7106.

[26] Welsh, J. P., Patel, K. G., Manthiram, K., and Swartz, J. R. (2009) Multiply mutated Gaussia luciferases provide prolonged and intense bioluminescence, *Biochem. Biophys. Res. Commun. 389*, 563-568.

[27] Zadoo, S., Nguyen, A., Zode, G., and Hulleman, J. D. (2016) A novel luciferase assay for sensitively monitoring myocilin variants in cell culture, *Invest. Ophth. Vis. Sci. 57*, 1939-1950.

[28] Shoulders, M. D., and Raines, R. T. (2009) Collagen structure and stability, *Annu. Rev. Biochem. 78*, 929-958.

[29] Boudko, S. P., Engel, J., and Bächinger, H. P. (2012) The crucial role of trimerization domains in collagen folding, *Int. J. Biochem. Cell Biol. 44*, 21-32.

[30] Rodan, S. B., Imai, Y., Thiede, M. A., Wesolowski, G., Thompson, D., Barshavit, Z., Shull, S., Mann, K., and Rodan, G. A. (1987) Characterization of a human osteosarcoma cell line (Saos-2) with osteoblastic properties, *Cancer Res. 47*, 4961-4966.

[31] Pautke, C., Schieker, M., Tischer, T., Kolk, A., Neth, P., Mutschler, W., and Milz, S. (2004) Characterization of osteosarcoma cell lines MG-63, Saos-2 and U-2OS in comparison to human osteoblasts, *Anticancer Res. 24*, 3743-3748.

[32] DiChiara, A. S., Taylor, R. J., Wong, M. Y., Doan, N. D., Rosario, A. M., and Shoulders, M. D. (2016) Mapping and exploring the collagen-I proteostasis network, *ACS Chem. Biol. 11*, 1408-1421.

[33] Blanck, T. J. J., and Peterkofsky, B. (1975) The stimulation of collagen secretion by ascorbate as a result of increased proline hydroxylation in chick embryo fibroblasts, *Arch. Biochem. Biophys. 171*, 259-267.

[34] Peterkofsky, B. (1991) Ascorbate requirement for hydroxylation and secretion of procollagen - Relationship to inhibition of collagen-synthesis in scurvy, *Am. J. Clin. Nutr. 54*, S1135-S1140.

[35] Ripley, C. R., Fant, J., and Bienkowski, R. S. (1993) Brefeldin A inhibits degradation as well as production and secretion of collagen in human lung fibroblasts, *J. Biol. Chem. 268*, 3677-3682.

[36] Kamel-ElSayed, S. A., Tiede-Lewis, L. M., Lu, Y., Veno, P. A., and Dallas, S. L. (2015) Novel approaches for two and three dimensional multiplexed imaging of osteocytes, *Bone 76*, 129-140.

[37] Heard, M. E., Besio, R., Weis, M., Rai, J., Hudson, D. M., Dimori, M., Zimmerman, S. M., Kamykowski, J. A., Hogue, W. R., Swain, F. L., Burdine, M. S., Mackintosh, S. G., Tackett, A. J., Suva, L. J., Eyre, D. R., and Morello, R. (2016) Sc65-Null Mice Provide Evidence for a Novel Endoplasmic Reticulum Complex Regulating Collagen Lysyl Hydroxylation, *PLoS Genet. 12*, e1006002.

[38] Canty-Laird, E. G., Lu, Y., and Kadler, K. E. (2012) Stepwise proteolytic activation of type I procollagen to collagen within the secretory pathway of tendon fibroblasts in situ, *Biochem. J. 441*, 707-717.

[39] Fitzgerald, J., Lamandé, S. R., and Bateman, J. F. (1999) Proteasomal Degradation of Unassembled Mutant Type I Collagen Pro-α1(I) Chains, *J. Biol. Chem. 274*, 27392-27398.

[40] Sidera, K., and Patsavoudi, E. (2014) HSP90 inhibitors: Current development and potential in cancer therapy, *Recent Pat. Anti-Canc. 9*, 1-20.

[41] Trepel, J., Mollapour, M., Giaccone, G., and Neckers, L. (2010) Targeting the dynamic HSP90 complex in cancer, *Nat. Rev. Cancer 10*, 537-549.

[42] Noh, H., Kim, H. J., Yu, M. R., Kim, W. Y., Kim, J., Ryu, J. H., Kwon, S. H., Jeon, J. S., Han, D. C., and Ziyadeh, F. (2012) Heat shock protein 90 inhibitor attenuates renal fibrosis through degradation of transforming growth factor-beta type II receptor, *Lab. Invest. 92*, 1583-1596.

[43] Ambade, A., Catalano, D., Lim, A., Kopoyan, A., Shaffer, S. A., and Mandrekar, P. (2014) Inhibition of heat shock protein 90 alleviates steatosis and macrophage activation in murine alcoholic liver injury, *J. Hepatol. 61*, 903-911.

[44] Tomcik, M., Zerr, P., Pitkowski, J., Palumbo-Zerr, K., Avouac, J., Distler, O., Becvar, R., Senolt, L., Schett, G., and Distler, J. H. (2014) Heat shock protein 90 (Hsp90) inhibition targets canonical TGF-beta signalling to prevent fibrosis, *Ann. Rheum. Dis. 73*, 1215-1222.

[45] Radovanac, K., Morgner, J., Schulz, J. N., Blumbach, K., Patterson, C., Geiger, T., Mann, M., Krieg, T., Eckes, B., Fassler, R., and Wickstrom, S. A. (2013) Stabilization of integrin-linked kinase by the Hsp90-CHIP axis impacts cellular force generation, migration and the fibrotic response, *EMBO J. 32*, 1409-1424.

[46] McClellan, A. J., Xia, Y., Deutschbauer, A. M., Davis, R. W., Gerstein, M., and Frydman, J. (2007) Diverse cellular functions of the Hsp90 molecular chaperone uncovered using systems approaches, *Cell 131*, 121-135.

[47] Chen, C. Y., and Balch, W. E. (2006) The Hsp90 chaperone complex regulates GDI-dependent Rab recycling, *Mol. Biol. Cell 17*, 3494-3507.

[48] Schwarzbauer, J. E., Spencer, C. S., and Wilson, C. L. (1989) Selective secretion of alternatively spliced fibronectin variants, *J. Cell Biol. 109*, 3445-3453.

[49] Singh, P., Carraher, C., and Schwarzbauer, J. E. (2010) Assembly of Fibronectin Extracellular Matrix, *Annu. Rev. Cell Dev. Biol. 26*, 397-419.

[50] Zang, Y., Wan, M., Liu, M., Ke, H., Ma, S., Liu, L. P., Ni, J. Q., and Pastor-Pareja, J. C. (2015) Plasma membrane overgrowth causes fibrotic collagen accumulation and immune activation in Drosophila adipocytes, *eLife 4*, e07187.

[51] Lam, K. K., Zheng, X., Forestieri, R., Balgi, A. D., Nodwell, M., Vollett, S., Anderson, H. J., Andersen, R. J., Av-Gay, Y., and Roberge, M. (2012) Nitazoxanide stimulates autophagy and inhibits mTORC1 signaling and intracellular proliferation of Mycobacterium tuberculosis, *PLoS Pathog. 8*, e1002691.

[52] Jones, M. G., Andriotis, O. G., Roberts, J. J., Lunn, K., Tear, V. J., Cao, L., Ask, K., Smart, D. E., Bonfanti, A., Johnson, P., Alzetani, A., Conforti, F., Doherty, R., Lai, C. Y.,

Johnson, B., Bourdakos, K. N., Fletcher, S. V., Marshall, B. G., Jogai, S., Brereton, C. J., Chee, S. J., Ottensmeier, C. H., Sime, P., Gauldie, J., Kolb, M., Mahajan, S., Fabre, A., Bhaskar, A., Jarolimek, W., Richeldi, L., O'Reilly, K. M., Monk, P. D., Thurner, P. J., and Davies, D. E. (2018) Nanoscale dysregulation of collagen structure-function disrupts mechano-homeostasis and mediates pulmonary fibrosis, *eLife 7*, e36354.

[53] Porter, G. T., Gadre, S. A., and Calhoun, K. H. (2006) The effects of intradermal and topical mitomycin C on wound healing, *Otolaryng. Head Neck 135*, 56-60.

[54] Ribeiro, F. D. Q., Guaraldo, L., Borges, J. D., Zacchi, F. F. S., and Eckley, C. A. (2004) Clinical and histological healing of surgical wounds treated with mitomycin C, *Laryngoscope 114*, 148-152.

[55] Talamo, J. H., Gollamudi, S., Green, W. R., Delacruz, Z., Filatov, V., and Stark, W. J. (1991) Modulation of corneal wound-healing after excimer laser keratomileusis using topical mitomycin-C and steroids, *Arch. Ophthalmol.-Chic. 109*, 1141-1146.

[56] Trabattoni, D., Gnudi, F., Ibba, S. V., Saulle, I., Agostini, S., Masetti, M., Biasin, M., Rossignol, J.-F., and Clerici, M. (2016) Thiazolides Elicit Anti-Viral Innate Immunity and Reduce HIV Replication, *Sci. Rep. 6*, 27148.

[57] Qu, Y., Olsen, J. R., Yuan, X., Cheng, P. F., Levesque, M. P., Brokstad, K. A., Hoffman, P. S., Oyan, A. M., Zhang, W., Kalland, K. H., and Ke, X. (2018) Small molecule promotes beta-catenin citrullination and inhibits Wnt signaling in cancer, *Nat. Chem. Biol. 14*, 94-101.

[58] de Carvalho, L. P., Darby, C. M., Rhee, K. Y., and Nathan, C. (2011) Nitazoxanide Disrupts Membrane Potential and Intrabacterial pH Homeostasis of Mycobacterium tuberculosis, *ACS Med. Chem. Lett. 2*, 849-854.

[59] Muller, J., Naguleswaran, A., Muller, N., and Hemphill, A. (2008) Neospora caninum: functional inhibition of protein disulfide isomerase by the broad-spectrum anti-parasitic drug nitazoxanide and other thiazolides, *Exp. Parasitol. 118*, 80-88.

[60] Ashiru, O., Howe, J. D., and Butters, T. D. (2014) Nitazoxanide, an antiviral thiazolide, depletes ATP-sensitive intracellular Ca(2+) stores, *Virology 462-463*, 135-148.

[61] Rossignol, J. F., La Frazia, S., Chiappa, L., Ciucci, A., and Santoro, M. G. (2009) Thiazolides, a new class of anti-influenza molecules targeting viral hemagglutinin at the post-translational level, *J. Biol. Chem. 284*, 29798-29808.

[62] Piacentini, S., La Frazia, S., Riccio, A., Pedersen, J. Z., Topai, A., Nicolotti, O., Rossignol, J. F., and Santoro, M. G. (2018) Nitazoxanide inhibits paramyxovirus replication by targeting the Fusion protein folding: role of glycoprotein-specific thiol oxidoreductase ERp57, *Sci. Rep. 8*, 10425.

[63] Lee, W. J., Lee, J. H., Ahn, H. M., Song, S. Y., Kim, Y. O., Lew, D. H., and Yun, C. O. (2015) Heat Shock Protein 90 Inhibitor Decreases Collagen Synthesis of Keloid Fibroblasts and Attenuates the Extracellular Matrix on the Keloid Spheroid Model, *Plast. Reconstr. Surg. 136*, 328e-337e.

[64] Ambade, A., Catalano, D., Lim, A., and Mandrekar, P. (2012) Inhibition of heat shock protein (molecular weight 90 kDa) attenuates proinflammatory cytokines and prevents lipopolysaccharide-induced liver injury in mice, *Hepatology 55*, 1585-1595.

[65] Myung, S. J., Yoon, J. H., Kim, B. H., Lee, J. H., Jung, E. U., and Lee, H. S. (2009) Heat shock protein 90 inhibitor induces apoptosis and attenuates activation of hepatic stellate cells, *J. Pharmacol. Exp. Ther. 330*, 276-282.

[66] Luft, C., Freeman, J., Elliott, D., Al-Tamimi, N., Kriston-Vizi, J., Heintze, J., Lindenschmidt, I., Seed, B., and Ketteler, R. (2014) Application of Gaussia luciferase in bicistronic and non-conventional secretion reporter constructs, *BMC Biochem. 15*.

[67] Campeau, E., Ruhl, V. E., Rodier, F., Smith, C. L., Rahmberg, B. L., Fuss, J. O., Campisi, J., Yaswen, P., Cooper, P. K., and Kaufman, P. D. (2009) A versatile viral system for expression and depletion of proteins in mammalian cells, *PLoS One 4*, e0006529.

[68] Dewal, M. B., DiChiara, A. S., Antonopoulos, A., Taylor, R. J., Harmon, C. J., Haslam, S. M., Dell, A., and Shoulders, M. D. (2015) XBP1s Links the Unfolded Protein Response to the Molecular Architecture of Mature N-Glycans, *Chem. Biol. 22*, 1301-1312.

[69] Zhang, J. H., Chung, T. D., and Oldenburg, K. R. (1999) A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays, *J. Biomol. Screen 4*, 67-73.

[70] Makareeva, E., Mertz, E. L., Kuznetsova, N. V., Sutter, M. B., DeRidder, A. M., Cabral, W. A., Barnes, A. M., McBride, D. J., Marini, J. C., and Leikin, S. (2008) Structural heterogeneity of type I collagen triple helix and its role in osteogenesis imperfecta, *J. Biol. Chem. 283*, 4787-4798.

[71] Moore, C. L., Dewal, M. B., Nekongo, E. E., Santiago, S., Lu, N. B., Levine, S. S., and Shoulders, M. D. (2016) Transportable, chemical genetic methodology for the small molecule-mediated inhibition of heat shock factor 1, *ACS Chem. Biol. 11*, 200-210.

[72] Bonner, W. M., and Laskey, R. A. (1974) A film detection method for tritium-labelled proteins and nucleic acids in polyacrylamide gels, *Eur. J. Biochem. 46*, 83-88.

[73] Hulleman, J. D., and Kelly, J. W. (2015) Genetic ablation of N-linked glycosylation reveals two key folding pathways for R345W fibulin-3, a secreted protein associated with retinal degeneration, *FASEB J. 29*, 565-575.

[74] Wang, H., Bian, X., Xia, L., Ding, X., Muller, R., Zhang, Y., Fu, J., and Stewart, A. F. (2014) Improved seamless mutagenesis by recombineering using ccdB for counterselection, *Nucleic Acids Res. 42*, e37.

[75] Wong, M. Y., Doan, N. D., DiChiara, A. S., Papa, L. J., Cheah, J. H., Soule, C. K., Watson, N., Hulleman, J. D., and Shoulders, M. D. (2018) A High-Throughput Assay for Collagen Secretion Suggests an Unanticipated Role for Hsp90 in Collagen Production, *Biochemistry 57*, 2814-2827.

CHAPTER III
A Role for Cytosolic Hsp90β in Collagen-I Secretion

Summary

In previous work, we found that the Hsp90 inhibitor 17-AAG reduces collagen-I secretion post-translationally and does not disrupt global protein secretion. This chapter describes our efforts to determine the mechanism of action of pan-isoform Hsp90 inhibitors. We show that Hsp90 inhibitors do not reduce collagen secretion by activating the heat shock response. Surprisingly, we also show that the consequences of Hsp90 inhibition cannot be attributed to inhibition of the endoplasmic reticulum's Hsp90 isoform, Grp94. Instead, collagen-I secretion likely depends on the activity of cytosolic Hsp90 chaperones, even though these chaperones cannot directly engage nascent collagen molecules. Results obtained with isoform-selective inhibitors and genetic knockdown experiments suggest that the constitutively expressed, cytosolic Hsp90β is specifically required for collagen-I secretion. Finally, we present preliminary evidence consistent with a role for the cytosolic protein NudCL and propose a possible model for our observations.

Contributions

Introduction

As one of the most abundant components of human cells, comprising 1–2% of total protein, it is hardly surprising that Hsp90 participates in a large number of cellular processes.[1] Though first discovered as part of the chromatin "puffs" induced by heat shock,[2] the repertoire of Hsp90 functions has since expanded to encompass organismal development, evolution, and disease.[3-6] In addition to regulating cellular responses to hypoxia, nitric oxide, and viral infection,[1] Hsp90 assists in protein secretion in both yeast and possibly mammalian cells.[7, 8] Specifically, the chaperone has been linked to function of vesicle-tethering proteins such as the COG and TRAPP families, Rab GTPase recycling, and export of model secretory cargo.[7, 8]

The Hsp90 chaperone cycle has been studied extensively, and involves rounds of ATP hydrolysis coupled to still unknown conformational changes in Hsp90 that promote client maturation.[1] Improvements to early natural product-based Hsp90 inhibitors have enabled identification of additional clients, as well as structural and biochemical studies of Hsp90-client interactions.[9, 10] However, the molecular details of how Hsp90 interacts with its co-chaperones and clients, or how these interactions contribute to diseases like cancer, are still largely unknown.

Due to the abundance of Hsp90 and the presence of multiple isoforms and transcript variants,[11] small-molecule inhibitors provide a key alternative to genetic approaches to investigate chaperone function. One of the challenges to using such compounds, however, is the lack of isoform selectivity. Hsp90 exists as not one, but two cytosolic isoforms (the stress-inducible α and constitutively-expressed β), in addition to isoforms that reside in the mitochondria (TRAP1) and endoplasmic reticulum (Grp94). While certain processes and clients seem to prefer Hsp90α over Hsp90β,[11, 12] others appear to require only sufficient levels of total Hsp90.[13] Many Hsp90 inhibitors, including those tested in this study, target the N-terminal domain of Hsp90, functioning as competitive inhibitors of ATP binding.[14, 15] However, the high similarity of this region, with 95% sequence identity among the cytosolic isoforms, limits the utility of such compounds for parsing isoform-specific functions.[16]

Among the proposed Hsp90 clients with isoform-specific requirements is collagen type I, hereafter referred to as collagen-I. A key structural component of skin and bone, collagen-I is synthesized and then exported from the ER as rigid, ~300 nm triple helical molecules, which are deposited in the extracellular space and then self-assemble into fibrils.[17] Collagen-I molecules are too large to fit into conventional COP-II vesicles, which average 60–90 nm in diameter, for export from the ER.[18] Other collagen types have varying lengths and spatial constraints: type II, for instance, is similar to collagen-I, with 360 uninterrupted repeats of the [Gly-X-Y] amino acid triplet; by contrast, type VIII is less than half the length of type II, with multiple breaks that allow it to fold in on itself.[19]

Across collagen types, insufficient collagen deposition and/or intracellular accumulation lead to cellular stress and tissue dysfunction.[20, 21] How collagen-I is exported from the ER to the extracellular space, and whether other collagen types or proteins utilize the same secretion pathway, thus constitute key questions for the field of collagen biosynthesis. Collagen-I has been shown to associate with Grp94 in the ER, although the functional importance of the interaction was not determined.[22-24] However, it seemed likely that Grp94 association with collagen-I reflected a role for this ER-resident Hsp90 isoform in collagen folding and/or secretion.

We previously designed a high-throughput luminescent assay for collagen-I secretion and used it to screen two libraries of FDA-approved or bioactive small-molecules.[25] In that work, we identified six validated, small molecule modulators that selectively inhibit collagen-I secretion by primary dermal fibroblasts. Of those compounds, the Hsp90 inhibitor 17-AAG had shown efficacy as an anti-fibrotic compound, through a mechanism proposed to involve altered inflammatory signaling.[26-31] We, however, showed that 17-AAG also operates post-

translationally to directly and selectively affect collagen-I secretion (Chapter II). Based on that work, we wanted to explore the mechanism of action of Hsp90 inhibition. Here, we present results from a panel of pan-isoform Hsp90 inhibitors, as well as recently developed isoform-selective small molecule inhibitors.[16] We find that while Grp94 is dispensable for collagen-I secretion, selective inhibition or knockdown of cytosolic Hsp90β recapitulates our initial results with 17-AAG. Genetic knockdown studies also suggest a potential role for the Hsp90β adapter and co-chaperone NudCL in facilitating collagen-I secretion, perhaps by regulating dynein motor dynamics.

Results

Structurally Diverse Hsp90 Inhibitors Reduce the Secretion of Endogenous Collagen-I. We had previously found that the Hsp90 inhibitor 17-AAG was capable of reducing collagen-I secretion both in our screening cell line and in primary human fibroblasts. Furthermore, we had shown that 17-AAG acted in a post-translational manner, and that the observed reduction in collagen-I secretion was independent of adverse effects on cell viability or Golgi morphology. We next wanted to confirm that the response to 17-AAG was a result of Hsp90 inhibition, and not an unappreciated off-target effect. We therefore obtained four additional Hsp90 inhibitors spanning diverse molecular scaffolds (Figure 3.1A) and tested the entire Hsp90 inhibitor set for effects on collagen-I secretion in primary human fibroblasts.

After a 24 h treatment with pan-isoform Hsp90 inhibitors, immunoblotting revealed significant decreases in endogenous collagen-I secretion relative to that of vehicle-treated cells for 17-AAG, 17-DMAG, EC 144, and STA-9090 (Figure 3.1B, C). We observed no changes in intracellular levels of collagen-I. We also confirmed functional inhibition of cytosolic Hsp90 at the concentrations used, indicated both by the expected increase in Hsp70 levels due to induction of a compensatory heat shock response (HSR) and by the loss of signal from the cytosolic Hsp90 client Akt (Figure 3.1C). One pan-isoform Hsp90 inhibitor, radicicol, caused a notable but not significant decrease in collagen-I secretion (Figure 3.1 B, C). However, radicicol was later found to be mildly cytotoxic (see Figure 3.2A), and radicicol treatment induced Akt degradation less effectively than the other pan-isoform Hsp90 inhibitors, perhaps explaining its less striking effects on collagen-I secretion. Cumulatively, the data in Figure 3.1 indicate that reductions in collagen-I secretion mediated by diverse Hsp90 inhibitors do occur in primary cells and that the reduced level of collagen-I secretion is an on-target consequence of Hsp90 inhibitor treatment.

We then tested the pan-isoform Hsp90 inhibitors for effects on cell viability, collagen-I transcripts, and collagen-I translation. We first treated primary fibroblasts with pan-isoform Hsp90 inhibitors for 24 h and then tested for cytotoxicity by staining cells with Annexin-V/Alexa Fluor 488 and propidium iodide followed by FACS analysis. With the exception of radicicol, which we found to be mildly cytotoxic, none of the Hsp90 inhibitors tested were toxic at the concentrations used, confirming that Hsp90 inhibitors do not decrease the secretion of collagen-I from primary fibroblasts simply by promoting cell death (Figure 3.2A).

Next, we performed qPCR analysis of collagen-I transcripts from Hsp90 inhibitor-treated primary fibroblasts. We found that *COL1A1* and *COL1A2* transcript levels were not significantly altered, indicating that the mechanism of action is post-transcriptional (Figure 3.2B). Finally, we used a tritium pulse assay to test whether collagen-I translation is altered by Hsp90 inhibitor treatment. Briefly, we pulsed treated fibroblasts with [5-^{3}H]proline for 30 min to label the proline-rich collagen-I protein, followed by collection of cell lysate and autoradiography analysis of the collagen-I bands (identified by bacterial collagenase treatment). We observed no significant alterations in levels of newly synthesized, radiolabeled collagen-I upon treatment with any of the Hsp90 inhibitors tested (Figure 3.2C). Cumulatively, these data indicate that structurally diverse Hsp90 inhibitors reduce collagen-I secretion via a post-translational mechanism.

59
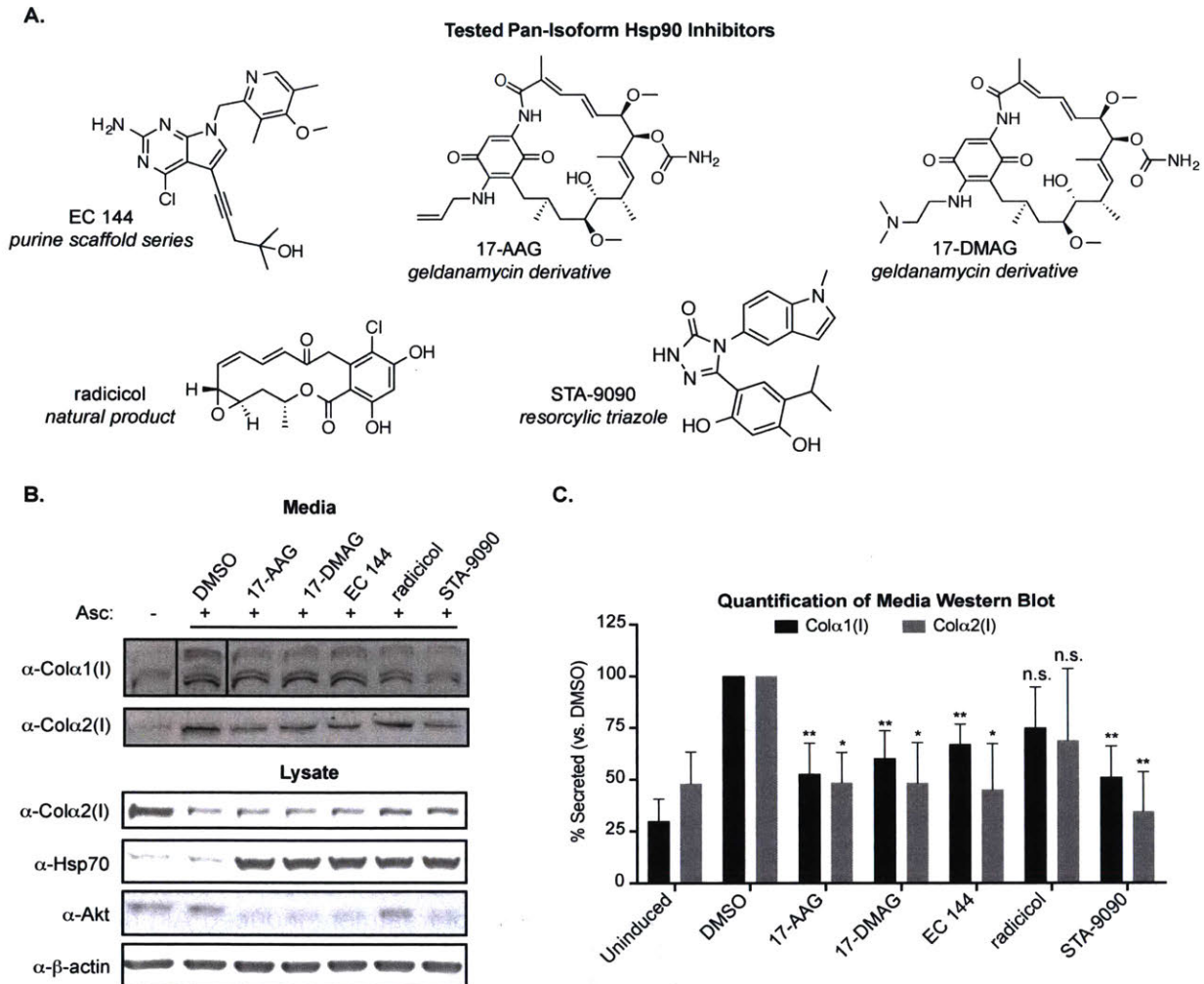
Figure 3.1: *Pan-isoform Hsp90 inhibitors reduce the level of endogenous collagen-I secretion.* (A) Structures, names, and compound classes of tested pan-isoform Hsp90 inhibitors. (B) Western blot of conditioned media (top) and lysate (bottom) from primary dermal fibroblasts. Cells were plated and allowed to adhere overnight before receiving fresh media with 200 µM ascorbate and the following concentrations of Hsp90 inhibitors, as indicated: 17-AAG (250 nM), 17-DMAG (250 nM), EC 144 (100 nM), radicicol (250 nM), and STA-9090 (300 nM). After being treated for 24 h, conditioned media and cells were harvested for immunoblotting. (C) Quantification of Western blot results. *p-value < 0.05 ($n = 3$).

Reduced Collagen-I Secretion Does Not Require the Heat Shock Response. While our data indicated that diverse Hsp90 inhibitors could reduce secretion of endogenous collagen-I, we wondered if our results might be due to secondary, rather than direct, effects of Hsp90 inhibition. Hsp90 inhibition is known to activate a protective heat shock response (HSR), which increases levels of small heat shock proteins in response to cellular stress.[32, 33] Among these proteins is Hsp70, a chaperone whose function in clathrin uncoating might suggest a more general role in vesicle-mediated transport.[34, 35] In addition to Hsp70, the heat shock protein Hsp47 is a well-established collagen-specific chaperone whose level of expression increases

A. Annexin V/PI Staining

B. Collagen-I Transcript Levels

C. [³H]-Labeled Collagen Synthesis
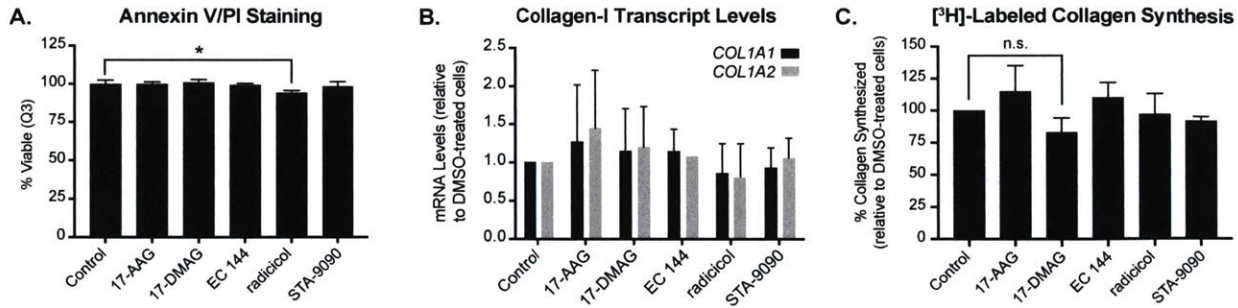
Figure 3.2: Effects of Hsp90 inhibition on collagen-I secretion are post-translational. (A) FACS analysis of Annexin-V/Alexa Fluor 488 and propidium iodide staining to detect cell death. Cells were plated in triplicate for each condition and allowed to adhere overnight before being treated as in Figure 3.1B. After 24 h, cells were trypsinized and labeled for FACS analysis. (B) qPCR analysis of collagen-I transcript levels from cells treated as in Figure 3.1B. Average values are shown with the standard deviation for each condition (n = 3). (C) Measurement of collagen-I synthesis by [5-³H]proline pulse. Cells treated for 6 h with ascorbate and 17-AAG (250 nM) were pulsed with [5-³H]proline for 30 min prior to being harvested and analyzed. Quantitation of newly synthesized collagen-I (mean ± standard deviation) is shown (n = 3).

upon HSR activation. We therefore examined the possibility that Hsp90 inhibitors reduced collagen-I secretion by increasing Hsp47 (or Hsp70) availability through the HSR.

To test this hypothesis, we transduced primary fibroblasts with lentivirus encoding a previously published, dominant negative version of Heat Shock Factor 1 (dncHSF1), the key transcriptional activator of the HSR.[33] We then treated cells with DMSO or 17-AAG and analyzed HSR induction by qPCR (Figure 3.3A). Our experimental time course was kept short to allow for expression of dncHSF1 prior to treatment with DMSO or 17-AAG, while minimizing repression of basal chaperone levels from prolonged dncHSF1 activity. Transcript levels of Hsp70 (*HSPA1A*) were clearly induced by 17-AAG treatment in a background of GFP-containing lentivirus. By contrast, co-treatment of cells with dncHSF1 lentivirus prevented upregulation of Hsp70 transcripts by 17-AAG, indicating that we are able to inhibit Hsp90 in the absence of a compensatory HSR. We found that HSR suppression did not rescue collagen-I secretion after treatment with 17-AAG (Figure 3.3B, C). Thus, our results suggest that Hsp90 itself, and not the HSR, is required for collagen-I secretion.

Grp94 Knockdown Does Not Inhibit Collagen-I Secretion. There are several Hsp90 isoforms in the cell: Hsp90α and Hsp90β in the cytosol, Grp94 in the ER, and TRAP1 in mitochondria (Figure 3.4A). Although 17-AAG and the other Hsp90 inhibitors employed can bind to and inhibit all Hsp90 isoforms, nascent collagen-I molecules are found only in the ER and Golgi. We therefore considered it likely that inhibition of Grp94, which is already known to engage collagen-I in the ER, was responsible for the observed effects of pan-isoform Hsp90 inhibitors on collagen-I secretion, possibly by slowing or preventing the folding of collagen-I. To test this hypothesis, we generated lentiviruses encoding a non-mammalian control shRNA or three distinct shRNA constructs targeting Grp94 (Table 3.1). We examined the levels of cytosolic Hsp90α and Hsp90β and the mitochondrial isoform TRAP1 and showed that they were unaffected by Grp94 knockdown, confirming the specificity of our shRNA constructs (Figure

A.

qPCR Analysis of Transcript Levels

Legend:
- DNAJB1
- HSPA1A
- HSP90AA1
- COL1A1
- COL1A2

17-AAG: − + − +
GFP dncHSF1

B.

Western Blot – Media

GFP dncHSF1
17-AAG: −   − +   − +

α-Colα1(I)

α-Colα2(I)

C.

Western Blot Quantification

Legend: Colα1(I) Colα2(I)
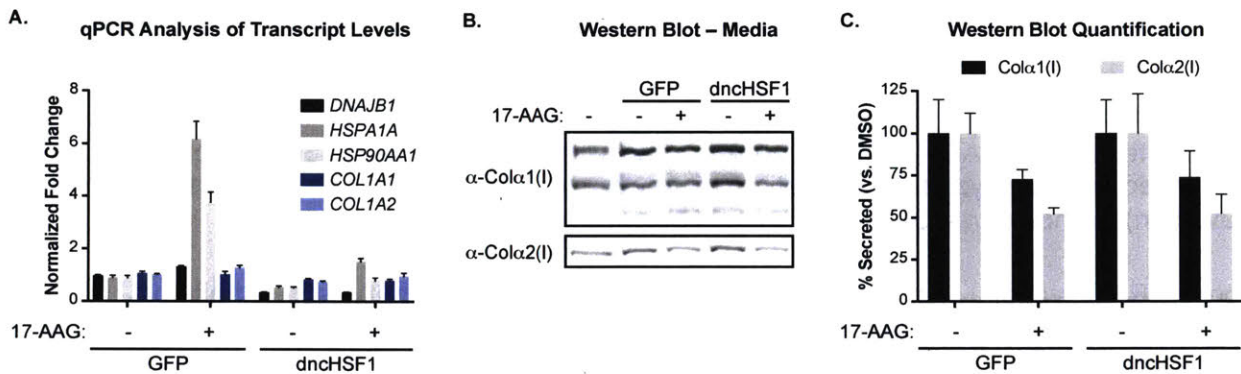
17-AAG: − + − +
GFP dncHSF1

Figure 3.3: *HSF1 is dispensable for 17-AAG mediated reductions in collagen-I secretion.* Primary dermal fibroblasts were plated and allowed to adhere overnight. Cells were then transduced with lentivirus encoding GFP or a constitutive, dominant-negative HSF1 (dncHSF1) in the presence of polybrene. The next day, cells were split and allowed to adhere for 24 h, after which cells were changed to fresh media with 200 μM ascorbate and either DMSO or 250 nM 17-AAG, as indicated. Cells and conditioned media were harvested after 24 h of treatment for qPCR analysis and immunoblotting. (A) qPCR analysis shows effective suppression of Hsp70 (*HSPA1A*) induction by dncHSF1. Collagen transcripts (blue columns) were not altered under any of the experimental conditions. (B) Western blot of conditioned media and (C) quantification showing that 17-AAG reduces collagen-I secretion even when the HSR is inhibited.

A.

B. Grp94 shRNA knockdown - Lysate

C. Grp94 shRNA knockdown - Media

D. Experimental Workflow

E. Grp94 knockdown and 17-AAG co-treatment - Media
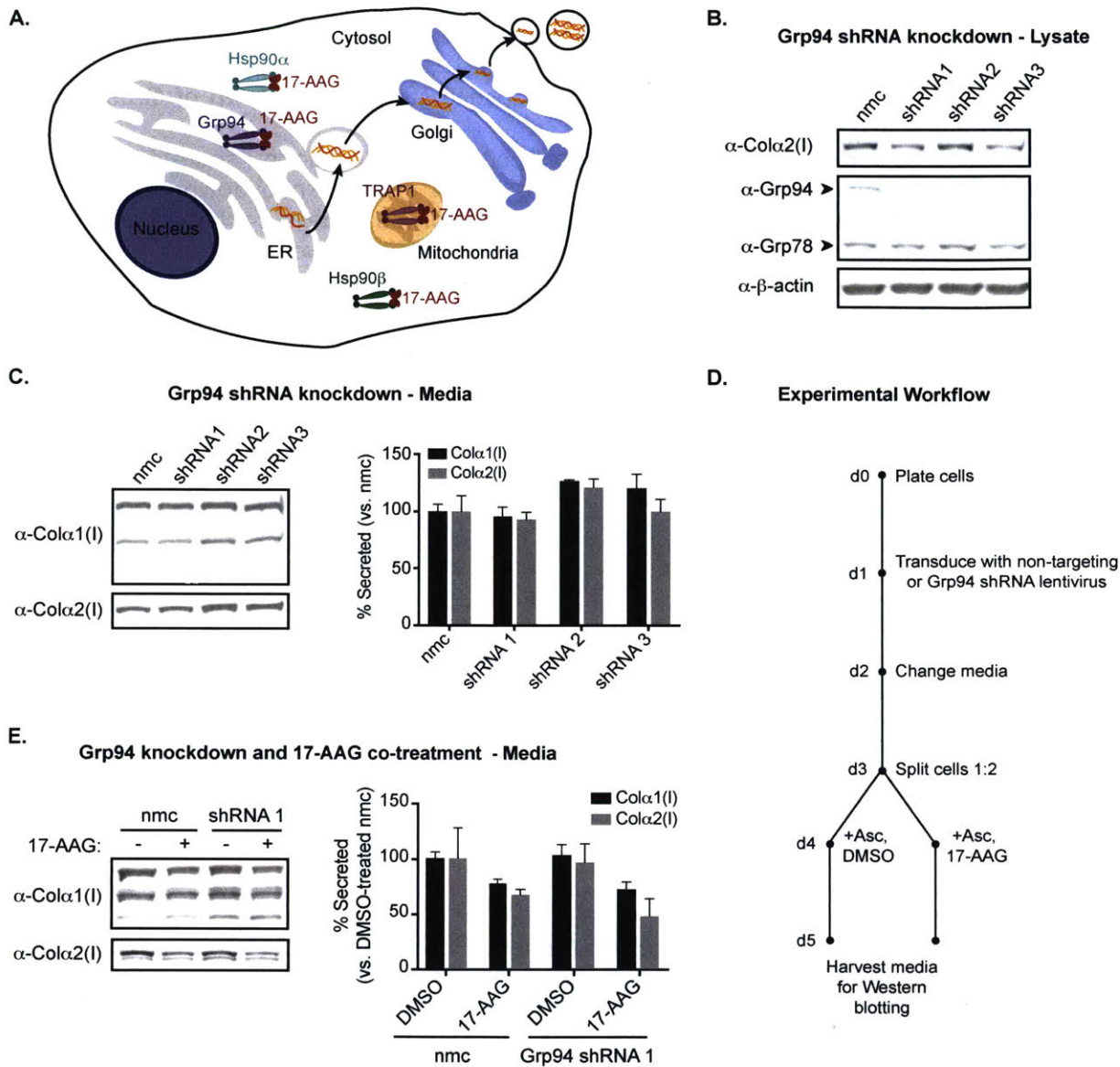
Figure 3.4: Grp94 is dispensable for secretion of collagen-I from primary fibroblasts. (A) Diagram of cellular Hsp90 isoforms along with the known collagen-I secretion pathway. Grp94, in the ER, is the only Hsp90 isoform capable of directly interacting with nascent collagen-I. (B) Western blot of lysate after lentiviral shRNA knockdown of Grp94. Cells were plated in triplicate for each construct and allowed to adhere overnight before being transduced with nontargeting control (nmc) or Grp94 shRNA lentiviruses. Media was changed the next day, and cells were left until 72 h post- transduction, when they received fresh media with ascorbate for 24 h prior to being harvested. (C) Immunoblots of corresponding media for the experiment described in panel B. (D) Experimental workflow. Cells were plated in triplicate for each construct and allowed to adhere overnight before being transduced with a nontargeting control or a Grp94 shRNA lentivirus. After 48 h, transductions were split 1:2 and allowed to adhere overnight. Cells subsequently received fresh media and were treated with ascorbate and DMSO or 17-AAG for 24 h prior to being harvested. (E) Immunoblots of conditioned media from cells co-treated with shRNA-containing lentivirus and DMSO or 17-AAG (*n* = 3).

63

3.5A). We then transduced primary fibroblasts with these lentiviruses, induced collagen-I expression with ascorbate 72 h post-transduction, collected media for 24 h, and then assayed levels of secreted collagen-I by immunoblotting (Figure 3.4C). Strikingly, we observed no reduction in collagen-I levels in conditioned media, despite strong knockdown of Grp94 at the protein level (Figure 3.4B).

We next evaluated whether Hsp90 inhibitors could reduce collagen-I secretion even in a background of Grp94 knockdown. We repeated the Grp94 shRNA knockdown and then co-treated cells with DMSO or 17-AAG (Figure 3.4D). qPCR of these cells confirmed that Grp94 transcript levels decreased upon shRNA knockdown; in addition, we observed that Hsp70 transcript levels increased upon 17-AAG treatment, indicating functional inhibition of Hsp90 (Figure 3.5B). We then collected and analyzed conditioned media by immunoblotting and found
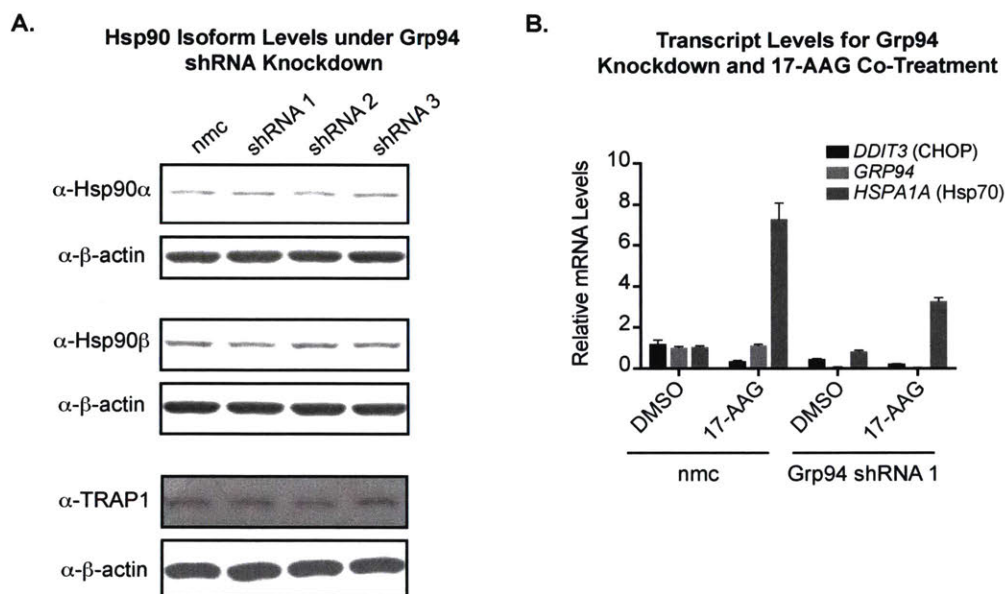
Figure 3.5: *Grp94 knockdown does not affect the levels of other Hsp90 isoforms.* (A) Levels of other Hsp90 isoforms do not change upon Grp94 shRNA knockdown. GM05659 fibroblasts were plated and transduced as described in the Materials and Methods. Equal amounts of lysate were separated on 4/8% Tris-Gly SDS-PAGE gels and probed for the indicated proteins. (B) qPCR analysis of transcript levels for CHOP (indicative of unfolded protein response activation), Grp94 (indicative of shRNA knockdown efficiency), and Hsp70 (indicative of heat shock response activation by 17-AAG treatment). GM05659 fibroblasts were plated and treated as outlined in Figure 3.4D.

that treating fibroblasts with 17-AAG decreased collagen-I secretion for both a nontargeting control and a Grp94-targeting shRNA (Figure 3.4E). These results demonstrate that Hsp90 inhibitors can reduce collagen-I secretion independent of Grp94 and suggest that Grp94 is in fact dispensable for collagen-I secretion under the conditions studied here. Instead, inhibition of another Hsp90 isoform, most likely cytosolic Hsp90, is likely to be responsible for reductions in collagen-I secretion.

Isoform-Selective Inhibitors Suggest that Cytosolic Hsp90β is Specifically Required for Collagen-I Secretion. We next sought to determine which Hsp90 isoform(s) is(are) required for collagen-I secretion. To do so, we employed a panel of isoform-selective small molecule Hsp90 inhibitors developed by the Blagg group (University of Notre Dame). The panel included one

inhibitor selective for each of the Hsp90 isoforms (Figure 3.6A; structures for the remaining unpublished compounds are not known),[36] and primary fibroblasts were treated at concentrations that did not significantly compromise cell viability, as assessed by Cell-TiterGlo and resazurin assays. After 24 h of treatment, immunoblotting revealed significant decreases in endogenous collagen-I secretion relative to that of vehicle-treated cells only for the Hsp90β-selective inhibitor (Figure 3.6B–D).
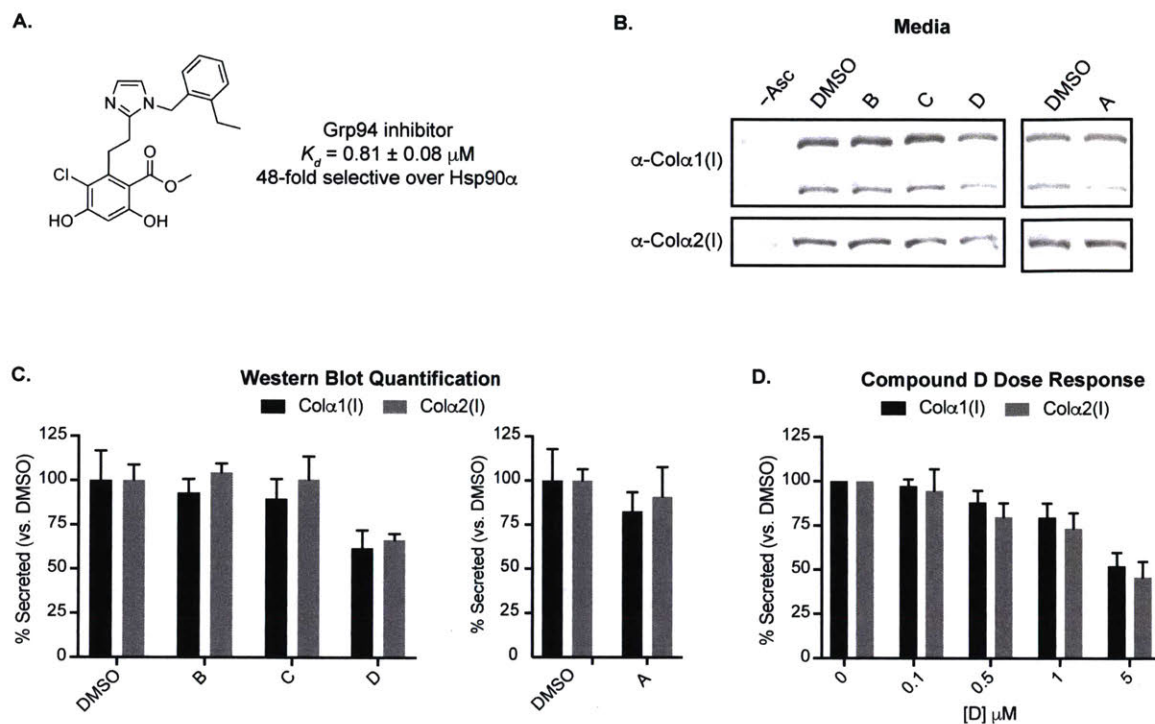
Figure 3.6: *Isoform-selective Hsp90 inhibitors suggest a role for cytosolic Hsp90β in collagen-I secretion.* Cells were plated and allowed to adhere overnight before receiving fresh media with 200 μM ascorbate and the following concentrations of Hsp90 inhibitors, as indicated: A (TRAP-1 selective, 10 μM), B (Grp94 selective, 5 μM), C (Hsp90α selective, 5 μM), or D (Hsp90β selective, 5 μM). After being treated for 24 h, conditioned media and cells were harvested for immunoblotting. (A) Structure of the Grp94-selective inhibitor, reported in Crowley, et al.[36] The remaining compounds are unpublished, and so structures are not shown. Representative Western blot images (B) are shown along with quantification of results in (C). Compound D was subsequently tested in dose at the indicated concentrations shown in (D).

Genetic Knockdown of Hsp90α and Hsp90β. Isoform selectivity remains a challenge for the development of small molecule Hsp90 inhibitors,[16, 37, 38] and so we wanted to verify our compound results by genetic knockdown. We had already shown that Grp94 shRNA had no effect on collagen-I secretion or the inhibitory action of 17-AAG (Figure 3.4). As the mitochondrial Hsp90 TRAP1 is separated from collagen-I biosynthesis and export by two membranes, and TRAP1 inhibition did not affect collagen-I secretion, we focused our efforts on the cytosolic isoforms. We obtained shRNA constructs for Hsp90α and Hsp90β (Table 3.1), generated lentivirus, and transduced primary fibroblasts with Hsp90-targeting constructs or a non-mammalian control. Collagen-I secretion was induced with ascorbate 72 h after transduction, and secreted collagen-I was collected for 24 h prior to analysis by immunoblotting.

Table 3.1: Sequences of shRNA constructs.

Target	Construct Name	Vector ID	Sequence
GFP	Non-mammalian control	SHC002	CCGGCAACAAGATGAAGAGCACCAACTCGAGTTGGTG CTCTTCATCTTGTTGTTTT
Grp94	shRNA 1	TRCN 0000029425	CCGGCCTGTGGATGAATACTGTATTCTCGAGAATACAG TATTCATCCACAGGTTTTT
	shRNA 2	TRCN 0000029426	CCGGCGTGGTCTGTTTGACGAATATCTCGAGATATTCG TCAAACAGACCACGTTTTT
	shRNA 3	TRCN 0000276248	CCGGGCGAGACTCTTCAGCAACATACTCGAGTATGTT GCTGAAGAGTCTCGCTTTTTG
Hsp90α	shRNA 1	TRCN 0000315009	CCGGTACTTGGAGGAACGAAGAATACTCGAGTATTCTT CGTTCCTCCAAGTATTTTTG
	shRNA 2	TRCN 0000314936	CCGGGTTATCCTACACCTGAAAGAACTCGAGTTCTTTC AGGTGTAGGATAACTTTTTG
Hsp90β	shRNA 1	TRCN 0000008750	CCGGCCAACTCATGTCCCTCATCATCTCGAGATGATGA GGGACATGAGTTGGTTTTT
	shRNA 2	TRCN 0000315415	CCGGCGCATGGAAGAAGTCGATTAGCTCGAGCTAATC GACTTCTTCCATGCGTTTTTG
	shRNA 3	TRCN 0000315416	CCGGCTTGTGTTGAAGGCAGTAAACCTCGAGGTTTACT GCCTTCAACACAAGTTTTTG
NudCL	shRNA 1	TRCN 0000134970	CCGGCCCTGCTTTAATAAACAGCAACTCGAGTTGCTGT TTATTAAAGCAGGGTTTTTG
	shRNA 2	TRCN 0000338520	CCGGCCCTGCTTTAATAAACAGCAACTCGAGTTGCTGT TTATTAAAGCAGGGTTTTTG
	shRNA 3	TRCN 0000138360	CCGGCCCAGCAAATTGGGATCACATCTCGAGATGTGA TCCCAATTTGCTGGGTTTTTG
USP19	shRNA 1	TRCN 0000376670	CCGGCTCCACTGCGAGCGAAGTATTCTCGAGAATACTT CGCTCGCAGTGGAGTTTTTG
	shRNA 2	TRCN 0000371020	CCGGCCGGTACTCTGTGAGTGTATTCTCGAGAATACAC TCACAGAGTACCGGTTTTTG

66

A.

Hsp90β shRNA knockdown

Western Blot Quantification
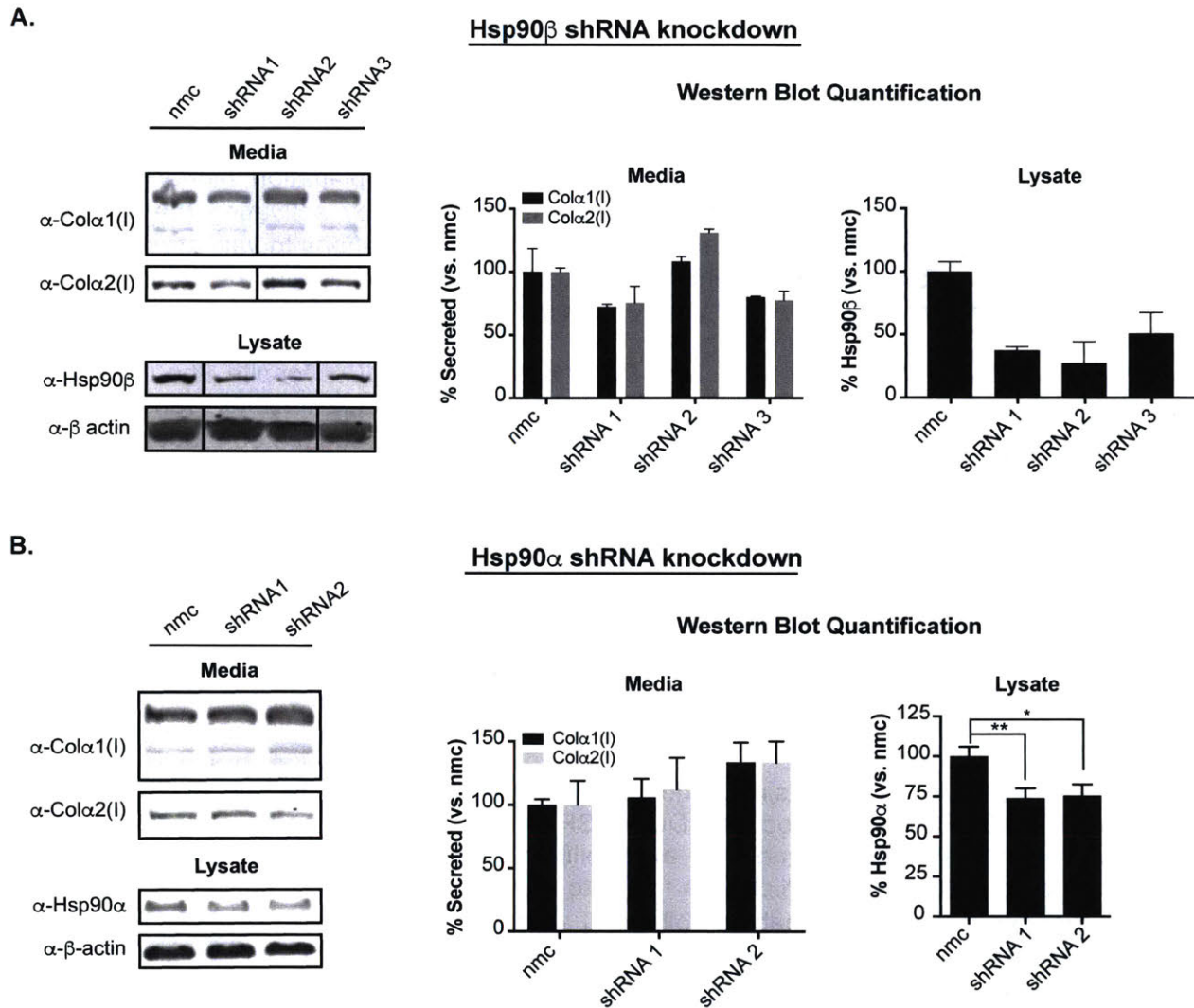
Figure 3.7: *Hsp90β, but not Hsp90α, shRNA knockdown decreases collagen-I secretion.* GM05294 cells were plated at a density of 2×10^5 cells per well in 6-well plates the day before transduction. Cells then received 500 μL of each virus in triplicate along with polybrene. The next day, cells received fresh media, and were incubated until 72 h post-transduction, when collagen-I secretion was induced with ascorbate. Conditioned media was collected for 24 h before immunoblotting analysis. Representative Western blots for shRNA knockdown of (A) Hsp90β and (B) Hsp90α are shown along with quantification.

Although we were only able to deplete Hsp90α to ~75% of control levels, Hsp90α knockdown had no effect on collagen-I secretion, as we observed with the small molecule inhibitors (Figure 3.7A).

Hsp90β knockdown proved much more challenging than knockdown of Hsp90α. We first transduced primary fibroblasts and induced collagen-I secretion 72 h after transduction. Immunoblots showed that both of our shRNA constructs resulted in 50% knockdown of the protein, with an ~20% reduction in collagen-I levels (Figure 3.7B). An additional construct also reduced Hsp90β levels, but surprisingly appeared to increase levels of secreted collagen-I (Figure 3.7B). Subsequent attempts to knockdown Hsp90β gave mixed results, including visible

cytotoxicity, no change in Hsp90β levels, and Hsp90β depletion in the absence of reduced collagen-I secretion. The high variability between experiments has been previously observed by us for other lentiviral shRNA targets, and may be at least partially attributed to an inconsistent multiplicity of infection (based on different titers of lentiviral batches or variations in the number of plated cells). However, it seems likely that an additional barrier to robust, reproducible knockdown is the essential function of Hsp90β. In support of this explanation, higher volumes of Hsp90β shRNA lentivirus induces expression of Hsp90α, likely by activating the HSR, and Hsp90β knockout cells and animals are not viable.[3, 11] Thus, future efforts to validate the Hsp90β inhibitor results by shRNA depletion will require careful attention to handling of lentiviral stocks, in addition to a thorough optimization of both viral volume and length of knockdown.

Genetic Knockdown of USP19 and the Co-Chaperone NudCL. Despite the substantial challenges of shRNA knockdown, the results of the small molecule inhibitors and failure of either Hsp90α or Grp94 knockdown to reduce collagen-I secretion suggested that Hsp90β was indeed the relevant isoform. This finding was surprising to us, as Hsp90β exists in the cytosol, and thus would not directly interact with collagen-I. Recent work, however, had identified cytosolic adapter and signaling proteins involved in protein ER export. Among these proteins was an E3 ubiquitin ligase complex composed of Cullin3 (CUL3) and the adapter protein KLHL12, which together installed a non-degradable mono-ubiquitin signal on the COP-II coat protein Sec31a;[39] ubiquitylation of Sec31a was sufficient to trigger assembly of enlarged COP-II vesicles capable of accommodating the collagen-I triple helix. Overexpression of KLHL12 led to installation of polyubiquitin chains on Sec31a and induced Sec31a degradation by the proteasome, but the mechanism for selecting mono- over poly-ubiquitination remained unknown.

One means of ensuring single modification by E3 ligases is by the action of a de-ubiquitinase (DUB). The ER-associated DUB USP19 had previously been reported to recruit misfolded substrates to endosomes for extracellular release; USP19 is also the only DUB known to interact with Hsp90, which was found to promote USP19 DUB activity *in vitro*.[40] We therefore wondered if Hsp90β enabled collagen-I secretion by enhancing USP19 DUB activity towards Sec31a in the cytosol. Selective inhibitors for USP19 do not exist, and initial

USP19 shRNA knockdown
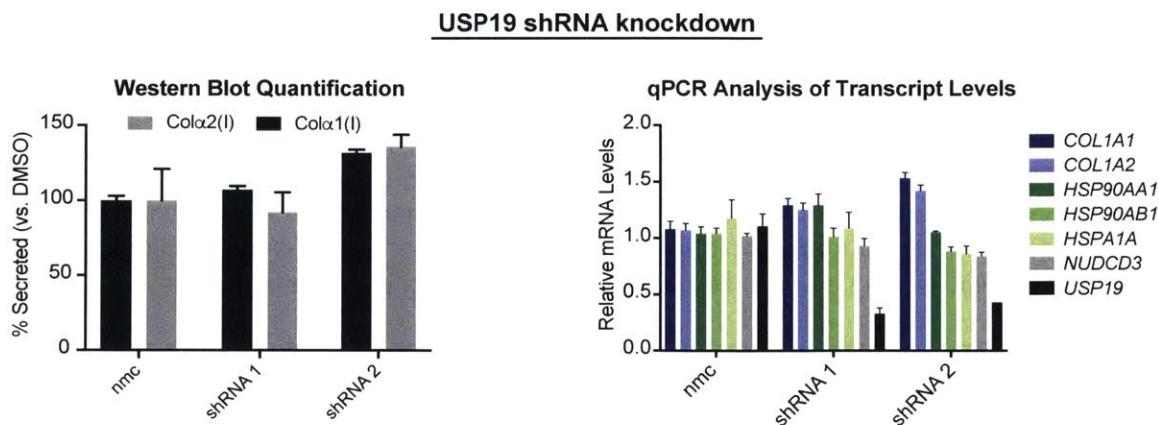
Figure 3.8: *USP19 shRNA knockdown does not affect collagen-I secretion.* GM05294 cells were plated at a density of 2×10^5 cells per well in 6-well plates the day before transduction. Cells then received 500 μL of each virus in triplicate along with polybrene. The next day, cells received fresh media, and were incubated until 72 h post-transduction, when collagen-I secretion was induced with ascorbate. Conditioned media was collected for 24 h before immunoblotting analysis.

experiments with the pan-DUB inhibitor PR-619[41] proved toxic to cells. Thus, to test our hypothesis, we obtained shRNA constructs against USP19 (Table 3.1) and transduced primary fibroblasts with generated lentiviruses. Collagen-I was induced with ascorbate 72 h post-transduction, and conditioned media was harvested for immunoblotting after 24 h; transcript levels of USP19 were analyzed by qPCR. While we were not able to detect endogenous USP19 by immunoblotting, qPCR confirmed that we were able to deplete USP19 transcripts to ~40% of control levels. However, this modest level of USP19 depletion did not affect collagen-I secretion (Figure 3.8).

Because our USP19 knockdown results did not affect collagen-I secretion, we next considered whether Hsp90 might participate in assembly or activation of the CUL3-KLHL12 complex. Hsp90 was previously found to influence assembly of the DBC2-CUL3-COP9 complex and bind CUL3 *in vitro*.[42, 43] In addition, the co-chaperone NudCL (encoded by the gene *NUDCD3*) could be co-purified from cellular lysates by KLHL12; a separate study used recombinant Hsp90 in the luminescent LUMIER assay and found that NudCL also associated with Hsp90β.[44, 45] We thus hypothesized that NudCL might serve as an adapter protein, allowing Hsp90β to bind to and promote the assembly of CUL3-KLHL12. We were able to deplete NudCL transcripts to ~50% of control levels using two lentiviral shRNA constructs (Figure 3.9A). In both cases, NudCL depletion reduced collagen-I secretion (Figure 3.9B and C). Later efforts to repeat NudCL shRNA knockdown, however, proved difficult, underscoring the problem of batch-to-batch variability in lentiviral titers and highlighting the importance of careful optimization of viral volume and length of knockdown.

Discussion

How does cytosolic Hsp90β, which should never interact with collagen-I directly inside cells, influence collagen-I secretion? As the heat shock protein Hsp47 is a well-established collagen-specific chaperone[23] whose level of expression increases upon HSR activation,[46] it is possible that Hsp90 inhibitors reduce the level of collagen-I secretion by increasing Hsp47 availability through the HSR. However, we found that suppression of the HSR via expression of a constitutive, dominant-negative HSF1 was unable to rescue collagen-I secretion in the presence of the Hsp90 inhibitor 17-AAG, suggesting that the HSR is not involved. Moreover, a recent report revealed that a small molecule inhibitor of the Hsp47−collagen interaction instead reduces the level of collagen secretion,[47] indicating that a different mechanism is likely operating here.

A more likely possibility is that cytosolic Hsp90 influences collagen-I secretion via interactions with cytosolic components of the secretory pathway. Because of its size and rigidity, collagen-I cannot be exported from cells in conventional COP-II secretory vesicles. Recent work has suggested that collagen is instead transported via a unique secretion pathway involving expanded COP-II vesicle coats.[39, 44] A combination of live-cell imaging, correlated light electron microscopy, and super-resolution microscopy techniques showed that enlarged COP-II vesicles are competent to carry collagen-I and are physically separated from the ER.[48] Because the inhibitors we tested selectively reduce the level of secretion of collagen-I, but not the level of secretion of the more compact proteins eGLuc2, TTR, fibulin-3, and fibronectin (see Chapter II), we speculate that cytosolic Hsp90 may be required for the assembly or transport of these large, collagen-specific vesicles.

Significant efforts by multiple groups have revealed how another large collagen, type VII, is exported from cells. Specifically, Tango1 concentrates collagen-VII at ER exit sites (ERES), recruits ERGIC membranes to the growing collagen carrier, and cooperates with cytosolic protein partners such as cTAGE5, Sedlin, and Sly1 to regulate vesicle growth (and possibly geometry).[49-55] A separate study found that transcription factor BBF2H7 upregulates the expression of genes involved in COP-II vesicle expansion to facilitate a switch from export of short chain collagen type VIII to extended collagen type II.[19] These results suggest that the

A.

qPCR Analysis of Transcript Levels

Legend:
- HSP90AA1
- HSP90AB1
- DNAJB1
- HSPA1A
- NUDCD3
- COL1A1
- COL1A2

B.

Western Blot – Media

α-Colα1(I)

α-Colα2(I)

nmc shRNA 1 shRNA 2 –Asc

C.

Western Blot Quantification

- Colα1(I)
- Colα2(I)

D.

NudCL and Collagen-I Secretion

(1) Hsp90β / Lis1 / NudCL
(2) Dynein / Lis1
(3)

Microtubule-mediated transport
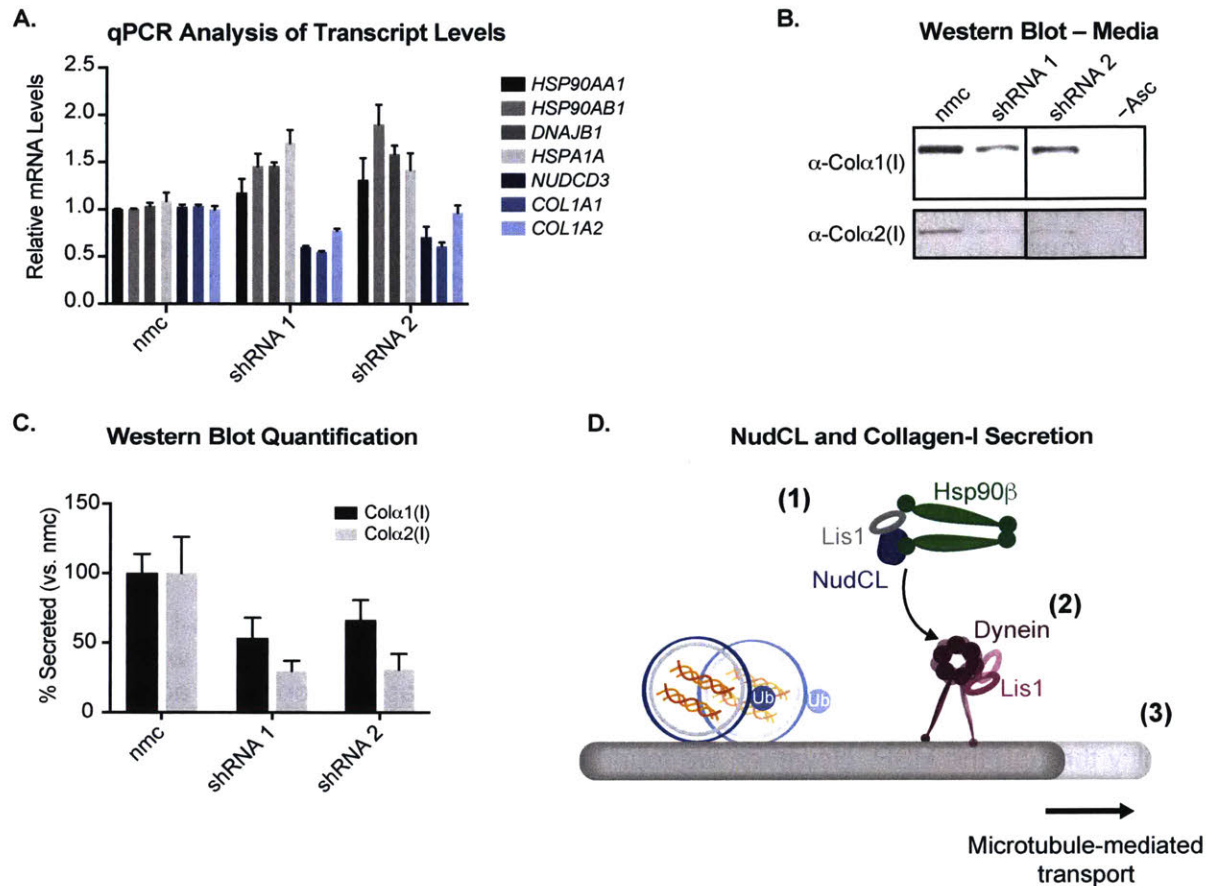
Figure 3.9: *NudCL shRNA knockdown and model.* GM05659 cells were plated at a density of 2×10^5 cells per well in 6-well plates the day before transduction. Cells then received 200 μL of each virus in triplicate along with polybrene. Media was changed the next day, and changed again after 72 h. Collagen secretion was induced with ascorbate 24 h prior to harvesting, and samples were harvested 7 d post-transduction. (A) shRNA knockdown was validated by qPCR analysis of transcript levels. (B) Immunoblots of conditioned media from cells transduced with control or NudCL shRNA-containing lentivirus ($n = 3$). (C) Quantification of Western blot data. (D) Potential model for NudCL involvement in collagen-I secretion. (1) Hsp90β and NudCL stabilize Lis1 and protect the protein from proteasomal degradation. (2) Lis1 is then able to bind the microtubule motor dynein, delaying the power-stroke and prolonging association of dynein with the microtubule. (3) Such interactions may promote collagen-I secretion by facilitating transport of collagen-containing vesicles to the plasma membrane.

mode of collagen secretion may be tailored to specific types, based on developmental signaling[19] or the directionality of collagen secretion.[18] Indeed, the dependence of collagen-I secretion on the Sec16 binding protein TFG implies the existence of analogous, but distinct, mechanisms of endoplasmic reticulum exit site organization for collagen-I and collagen-VII export.[56] Exactly how TFG, Tango1, and their partners facilitate the secretion of specific collagen types (Table 3.2) awaits further clarification.[49, 53, 57]

Table 3.2: Consequences of secretory pathway component depletion.

Protein	Method	Collagen Type/ Secretion Defect	Other Secretory Cargo/ Secretion Defect	Reference
cTAGE5	siRNA	Type VII / ER export blocked	VSVG, general protein secretion / none	Saito et al. (Ref. [51])
Sec24D	Mutational screen	Type II / blocked	Matrilin / blocked B1 integrin receptor, cadherins / none	Sarmah et al. (Ref. [71])
Sedlin	siRNA	Type I, II / ER export blocked	tsVSVG, CD8a, albumin, a1AT, general protein secretion / none	Venditti et al. (Ref. [52])
Sly1	siRNA	Type I / none Type VII / 75% reduction	General protein secretion / none ssHRP / blocked BiP, ERGIC-53 / none	Nogueira et al. (Ref. [53])
Syntaxin 5	siRNA	Type I, VII / blocked	---	Nogueira et al. (Ref. [53])
Syntaxin 18	siRNA	Type I / none Type VII / blocked	---	Nogueira et al. (Ref. [53])
Tango1	siRNA	Type VII / 90% reduction	ssHRP / 50% inhibition	Nogueira et al. (Ref. [53])
	KO	Type I–IV, VII, IX / blocked	COMP / blocked	Wilson et al. (Ref. [57])
TFG	siRNA	Type I, IV, VI / blocked	Mannosidase II, VSVG, E-cadherin, GalT / none	McCaughey et al. (Ref. [56])

Beyond these proteins, recent work has shown that calcium flux triggers localization of the Ca^{2+}-binding protein ALG2 to ERES.[58] ALG2 was previously identified as a joint interactor of KLHL12 and Sec31, and proposed to serve as a Ca^{2+}-sensing trigger for COP-II vesicle growth.[44] ALG2, in turn, interacts with TFG, and has been hypothesized to promote TFG retention at ERES as well as TFG polymerization in response to Ca^{2+}.[59] Such findings add another potential layer of Ca^{2+}-dependent regulation to collagen-I secretion, in which ALG2 not only bridges the interaction between the CUL3/KLHL12 complex and its ubiquitylation target Sec31a, but also facilitates organization of the transitional ER via TFG polymerization.[44, 56] A separate study suggested that TFG does not function at ERES, but instead migrates with COP-II vesicles to promote release of outer coat proteins in preparation for fusion with the ERGIC.[60]

NudCL had been identified in affinity purification and *in vitro* binding assays as an interactor of both KLHL12 (but not Sec31) and Hsp90β.[44, 45] Because Hsp90 had been implicated in the assembly of CUL3 E3 ligase complexes before,[43] we wondered if NudCL might

be acting as an adapter for Hsp90, allowing the chaperone to assist in the folding and/or association of KLHL12 with CUL3. However, we have thus far been unable to isolate a ternary complex between CUL3/KLHL12, NudCL, and Hsp90β. Moreover, difficulties in reproducing the initial NudCL shRNA knockdown results have further hindered our efforts to conclusively test this hypothesis.

The cytotoxicity we observed with NudCL depletion suggested that, like Hsp90, the co-chaperone might have critical, general roles in cells. Indeed, NudCL is required for stabilization of the microtubule motor protein dynein, and depletion of NudCL leads to mitotic defects and cell death.[61] NudCL also acts in concert with Hsp90 to stabilize the protein Lis1, which regulates dynein ATPase activity and has been implicated in mitotic defects as well.[62-64] But while the importance of NudCL in stabilizing dynein complicates genetic approaches to test NudCL involvement in collagen-I secretion, it also presents another, alternate mechanism by which NudCL might facilitate protein export. In addition to binding Sec31, the protein ALG2 was shown to complex with MISSL and microtubule-associated protein 1 B (MAP1B) in a Ca^{2+}-dependent manner. MAP1B binds Lis1, and shares a microtubule-binding motif with dynein, leading the authors of the study to suggest that MAP1B may compete with dynein for access to microtubules to regulate protein secretion.[65] The reductions in collagen-I secretion we observe upon Hsp90β and NudCL inhibition or depletion are thus consistent with a role for NudCL in promoting collagen-I export, not by promoting CUL3/KLHL12 complex formation, but instead by generating sufficient levels of dynein and Lis1 to allow for microtubule-mediated (and possibly Ca^{2+}-dependent) transport of COP-II vesicles (Figure 3.9D).

NudCL binds the N-terminal domain of Hsp90, and the pan-isoform inhibitor STA-9090 blocks this interaction in vitro;[66] meanwhile, the pan-isoform, N-terminal inhibitors geldanamycin and radicicol both reduced Lis1 protein levels in HeLa cells.[62] It will thus be interesting to determine if shorter treatment times (on the scale used here) are also sufficient to disrupt this interaction, and whether altered Lis1 levels can affect the MISSL-ALG2-MAP1B complex or the secretion of cargo besides secreted alkaline phosphatase.

While the precise interactions described above remain to be validated and clarified, a striking feature of our collagen-I secretion data is the importance of abundant, multi-functional proteins like NudCL and Hsp90. Adding to the puzzle is the absence of TFG, cTAGE5, and other Tango1 isoforms in Drosophila, which can nonetheless secrete collagen IV and has been used as a model for studying collagen secretion.[21, 51] The single Drosophila Tango1 isoform, however, is not only sufficient for collagen export, but also required for general protein secretion, pointing to a possible generic mechanism of ER export.[67] It will therefore be of interest to test the involvement of NudCL and Hsp90 in collagen secretion by Drosophila and other organisms, and perhaps determine if the chaperones are also part of a minimal secretion cohort that was later expanded in specialized mammalian cells.

Finally, the apparent involvement of NudCL and Hsp90β in the non-canonical secretion of collagen-I also highlights the advantages of post-translational regulation. Hsp90 is subject to phosphorylation, acetylation, S-nitrosylation, ubiquitylation, and SUMOylation, all of which can regulate chaperone activity.[68] Linking collagen-I export to an abundant protein whose activity is regulated by a post-translational modification would allow for tight temporal control, as has been proposed for the monoubiquitylation of Sec31 in response to transient Ca^{2+} influx.[44] The discovery of other modifications, such as O-GlcNAcylation of Sec23A[69] and Tango1[70] and the ubiquitylation of PEF1,[44] suggest that post-translational regulation may be a common theme of collagen-I secretion, allowing a large pool of general chaperones, enzymes, and effectors to be rapidly mobilized in response to changing developmental or physiological demands.

Materials and Methods

Cell Lines and Reagents. Healthy dermal fibroblasts (GM05659, Coriell Cell Repositories) were cultured in complete MEM supplemented with 15% FBS, 100 IU penicillin/100 μg/mL streptomycin, and 2 mM L-glutamine (Corning). Immunoblots were probed with the following primary antibodies, as indicated: Colα1(I) (Sigma HPA008405), Colα2(I) (Sigma SAB4500363), Colα2(I) (Abcam ab96723), Colα1(I) (NIH LF-68), Hsp70/Hsp72 (Cell Signaling Technologies 4873), Akt (Cell Signaling Technologies 9272), β-actin (Sigma A1978), fibronectin (Sigma F7387), KDEL (Enzo 10C3 ADI-SPA-827; recognizes Grp94 and Grp78), Hsp90α (Santa Cruz sc515081), Hsp90β (Abcam ab53497), TRAP1 (Santa Cruz sc13557), or NudCL (Sigma HPA019136). Duplicate blots from representative experiments were probed with both sets of Colα1(I) or Colα2(I) antibodies to compare antibody sensitivity (Figure 3.7).

Western Blotting Analysis. GM05659 primary fibroblasts were plated in 6-well plates at a density of 2×10^5 cells/well and allowed to adhere overnight. The next day, the medium was changed to full MEM with 0.5× penicillin/streptomycin, collagen expression was induced with 200 μM sodium ascorbate (Amresco), and cells were treated with DMSO or the indicated compound concentrations for 24 h. Samples of medium were denatured by being boiled in Laemmli buffer (supplemented with 167 mM 1,4-dithiothreitol; DTT), separated on homemade 4/8% Tris-Gly SDS–PAGE gels, and transferred to nitrocellulose membranes. Because of cross-reactivity between the collagen-I antibodies, samples of medium for collagen-I blots were run on duplicate gels, transferred, and then probed separately for either rabbit anti-Colα2(I) or rabbit anti-Colα1(I).

For Western blots of cell lysates, cells were trypsinized, washed with phosphate-buffered saline (PBS), and lysed in Triton-X lysis buffer with 1.5 mM PMSF and a protease inhibitor tablet (Pierce). Total protein was quantified using the bicinchoninic acid assay, and equal amounts of total protein were analyzed for each sample. Each experiment was performed in biological triplicate. Blots were imaged after incubation with appropriate primary and 800CW or 680LT secondary antibodies (LI-COR) by scanning on an Odyssey infrared imager (LI-COR), followed by quantification using ImageJ.

Lentivirus Production. Lentiviral vectors (Sigma) encoding shRNA constructs against Grp94, Hsp90α, Hsp90β, USP19, NudCL, or GFP (referred to in the text as a non-mammalian control) were amplified and prepped along with the lentiviral packaging vectors RRE, REV, and VSVG (Addgene). 9×10^6 293FT cells were plated on poly-D-lysine (Sigma)-coated 10-cm dishes and allowed to adhere overnight. The next day, 15 μg of each lentiviral vector was combined with 15 μg of RRE, 6 μg REV, 3 μg VSVG, and 60 μL Lipofectamine 2000 in OptiMEM (Life Technologies). Transfection mixtures were incubated for 40 min at rt and 293FT cell media was changed to OptiMEM. The transfection complexes were then added dropwise to the plate and incubated 12–16 h at 37 °C, after which the OptiMEM was removed and replaced by full medium. Lentivirus was harvested 48 h after the media change, spun at 207 ×g for 5 min to remove cell debris, and stored in single-use aliquots at −80 °C.

Quantitative PCR (qPCR). GM05659 primary fibroblasts were plated and treated as described above. Cells were then washed with PBS, and RNA was extracted using the Omega E.Z.N.A. Total RNA extraction kit. cDNA was prepared from 500 ng of RNA, normalized for all samples in each run, using an Applied Biosystems Reverse Transcriptase cDNA Kit in a Bio-Rad Thermocycler. Samples were run on a Light Cycler 480 II Real Time PCR Instrument in the MIT BioMicro Center using previously described primers and analysis methods. The following qPCR primer sequences were used for detecting transcript levels of Grp94 (*HSP90B1*): 5'-GGCCAGTTTGGTGTCGGT-3' and 5'-CGTTCCCCGTCCTAGAGTGTT-3'; Hsp90β

(*HSP90AB1*): 5'-CGAAGTTGGACAGTGGTAAAGAG-3' and 5'-TGCCCAATCATGGAGATGTCT-3'; USP19: 5'-TACCACCGCTGAACTCCCAA-3' and 5'-ACAGTCATCTTTCCCAGGGTT-3'; or NudCL (*NUDCD3*): 5'-CACCACAGAATTGGATGGGCA-3' and 5'-CACCATTGTAACTGTCGGGATTT-3'.

shRNA Knockdown. GM05659 primary fibroblasts were plated at a density of 2 × 10^5 cells/well in 6-well plates the day before transduction. After being allowed to adhere overnight, cells were given fresh medium (with 4 µg/mL Polybrene) and the indicated amounts of lentivirus. The next day, cells were subjected to a medium change and incubated until 72 h post-transduction, when the medium was changed and collagen-I expression was induced with ascorbate. Conditioned medium was collected 24 h later for Western blotting, as described above. Sequences for shRNA constructs are provided in Table 3.1. Transductions for co-treatment experiments with Grp94 shRNAs and 17-AAG were performed as described above, but using 300 µL of the indicated lentivirus. shRNA knockdown of Hsp90α and Hsp90β was performed as described above but using GM05294 fibroblasts.

References

[1] Taipale, M., Jarosz, D. F., and Lindquist, S. (2010) HSP90 at the hub of protein homeostasis: emerging mechanistic insights, *Nat. Rev. Mol. Cell Biol. 11*, 515-528.

[2] Morcillo, G., Diez, J. L., Carbajal, M. E., and Tanguay, R. M. (1993) HSP90 associates with specific heat shock puffs (hsr omega) in polytene chromosomes of Drosophila and Chironomus, *Chromosoma 102*, 648-659.

[3] Voss, A. K., Thomas, T., and Gruss, P. (2000) Mice lacking HSP90beta fail to develop a placental labyrinth, *Development 127*, 1-11.

[4] Rutherford, S. L., and Lindquist, S. (1998) Hsp90 as a capacitor for morphological evolution, *Nature 396*, 336-342.

[5] Rohner, N., Jarosz, D. F., Kowalko, J. E., Yoshizawa, M., Jeffery, W. R., Borowsky, R. L., Lindquist, S., and Tabin, C. J. (2013) Cryptic variation in morphological evolution: HSP90 as a capacitor for loss of eyes in cavefish, *Science 342*, 1372-1375.

[6] Rodina, A., Wang, T., Yan, P., Gomes, E. D., Dunphy, M. P., Pillarsetty, N., Koren, J., Gerecitano, J. F., Taldone, T., Zong, H., Caldas-Lopes, E., Alpaugh, M., Corben, A., Riolo, M., Beattie, B., Pressl, C., Peter, R. I., Xu, C., Trondl, R., Patel, H. J., Shimizu, F., Bolaender, A., Yang, C., Panchal, P., Farooq, M. F., Kishinevsky, S., Modi, S., Lin, O., Chu, F., Patil, S., Erdjument-Bromage, H., Zanzonico, P., Hudis, C., Studer, L., Roboz, G. J., Cesarman, E., Cerchietti, L., Levine, R., Melnick, A., Larson, S. M., Lewis, J. S., Guzman, M. L., and Chiosis, G. (2016) The epichaperome is an integrated chaperome network that facilitates tumour survival, *Nature 538*, 397-401.

[7] Lotz, G. P., Brychzy, A., Heinz, S., and Obermann, W. M. (2008) A novel HSP90 chaperone complex regulates intracellular vesicle transport, *J. Cell Sci. 121*, 717-723.

[8] McClellan, A. J., Xia, Y., Deutschbauer, A. M., Davis, R. W., Gerstein, M., and Frydman, J. (2007) Diverse cellular functions of the Hsp90 molecular chaperone uncovered using systems approaches, *Cell 131*, 121-135.

[9] Sidera, K., and Patsavoudi, E. (2014) HSP90 inhibitors: current development and potential in cancer therapy, *Recent Pat. Anticancer Drug Discov. 9*, 1-20.

[10] Shrestha, L., Patel, H. J., and Chiosis, G. (2016) Chemical Tools to Investigate Mechanisms Associated with HSP90 and HSP70 in Disease, *Cell Chem. Biol. 23*, 158-172.

[11] Zou, M., Bhatia, A., Dong, H., Jayaprakash, P., Guo, J., Sahu, D., Hou, Y., Tsen, F., Tong, C., O'Brien, K., Situ, A. J., Schmidt, T., Chen, M., Ying, Q., Ulmer, T. S., Woodley, D. T., and Li, W. (2017) Evolutionarily conserved dual lysine motif determines the non-chaperone function of secreted Hsp90alpha in tumour progression, *Oncogene 36*, 2160-2171.

[12] Bergmayr, C., Thurner, P., Keuerleber, S., Kudlacek, O., Nanoff, C., Freissmuth, M., and Gruber, C. W. (2013) Recruitment of a cytoplasmic chaperone relay by the A2A adenosine receptor, *J. Biol. Chem. 288*, 28831-28844.

[13] Shinozaki, F., Minami, M., Chiba, T., Suzuki, M., Yoshimatsu, K., Ichikawa, Y., Terasawa, K., Emori, Y., Matsumoto, K., Kurosaki, T., Nakai, A., Tanaka, K., and Minami, Y. (2006) Depletion of hsp90beta induces multiple defects in B cell receptor signaling, *J. Biol. Chem. 281*, 16361-16369.

[14] Gooljarsingh, L. T., Fernandes, C., Yan, K., Zhang, H., Grooms, M., Johanson, K., Sinnamon, R. H., Kirkpatrick, R. B., Kerrigan, J., Lewis, T., Arnone, M., King, A. J., Lai, Z., Copeland, R. A., and Tummino, P. J. (2006) A biochemical rationale for the anticancer effects of Hsp90 inhibitors: Slow, tight binding inhibition by geldanamycin and its analogues, *Proc. Natl. Acad. Sci. U.S.A. 103*, 7625-7630.

[15] Roe, S. M., Prodromou, C., O'Brien, R., Ladbury, J. E., Piper, P. W., and Pearl, L. H. (1999) Structural Basis for Inhibition of the Hsp90 Molecular Chaperone by the Antitumor Antibiotics Radicicol and Geldanamycin, *J. Med. Chem. 42*, 260-266.

[16] Khandelwal, A., Kent, C. N., Balch, M., Peng, S., Mishra, S. J., Deng, J., Day, V. W., Liu, W., Subramanian, C., Cohen, M., Holzbeierlein, J. M., Matts, R., and Blagg, B. S. J. (2018) Structure-guided design of an Hsp90beta N-terminal isoform-selective inhibitor, *Nat. Commun. 9*, 425.

[17] Shoulders, M. D., and Raines, R. T. (2009) Collagen structure and stability, *Annu. Rev. Biochem. 78*, 929-958.

[18] Malhotra, V., and Erlmann, P. (2015) The pathway of collagen secretion, *Annu. Rev. Cell Dev. Biol. 31*, 109-124.

[19] Ishikawa, T., Toyama, T., Nakamura, Y., Tamada, K., Shimizu, H., Ninagawa, S., Okada, T., Kamei, Y., Ishikawa-Fujiwara, T., Todo, T., Aoyama, E., Takigawa, M., Harada, A., and Mori, K. (2017) UPR transducer BBF2H7 allows export of type II collagen in a cargo- and developmental stage–specific manner, *J. Cell Biol. 216*, 1761-1774.

[20] Mirigian, L. S., Makareeva, E., Mertz, E. L., Omari, S., Roberts-Pilgrim, A. M., Oestreich, A. K., Phillips, C. L., and Leikin, S. (2016) Osteoblast Malfunction Caused by Cell Stress Response to Procollagen Misfolding in alpha2(I)-G610C Mouse Model of Osteogenesis Imperfecta, *J. Bone Miner. Res. 31*, 1608-1616.

[21] Zang, Y., Wan, M., Liu, M., Ke, H., Ma, S., Liu, L. P., Ni, J. Q., and Pastor-Pareja, J. C. (2015) Plasma membrane overgrowth causes fibrotic collagen accumulation and immune activation in Drosophila adipocytes, *eLife 4*, e07187.

[22] DiChiara, A. S., Taylor, R. J., Wong, M. Y., Doan, N. D., Rosario, A. M., and Shoulders, M. D. (2016) Mapping and Exploring the Collagen-I Proteostasis Network, *ACS Chem. Biol. 11*, 1408-1421.

[23] Ferreira, L. R., Norris, K., Smith, T., Hebert, C., and Sauk, J. J. (1994) Association of Hsp47, Grp78, and Grp94 with procollagen supports the successive or coupled action of molecular chaperones, *J. Cell Biochem. 56*, 518-526.

[24] Lamande, S. R., and Bateman, J. F. (1999) Procollagen folding and assembly: the role of endoplasmic reticulum enzymes and molecular chaperones, *Semin. Cell Dev. Biol. 10*, 455-464.

[25] Wong, M. Y., Doan, N. D., DiChiara, A. S., Papa, L. J., Cheah, J. H., Soule, C. K., Watson, N., Hulleman, J. D., and Shoulders, M. D. (2018) A High-Throughput Assay for Collagen Secretion Suggests an Unanticipated Role for Hsp90 in Collagen Production, *Biochemistry 57*, 2814-2827.

[26] Lee, W. J., Lee, J. H., Ahn, H. M., Song, S. Y., Kim, Y. O., Lew, D. H., and Yun, C. O. (2015) Heat Shock Protein 90 Inhibitor Decreases Collagen Synthesis of Keloid Fibroblasts and Attenuates the Extracellular Matrix on the Keloid Spheroid Model, *Plast. Reconstr. Surg. 136*, 328e-337e.

[27] Noh, H., Kim, H. J., Yu, M. R., Kim, W. Y., Kim, J., Ryu, J. H., Kwon, S. H., Jeon, J. S., Han, D. C., and Ziyadeh, F. (2012) Heat shock protein 90 inhibitor attenuates renal fibrosis through degradation of transforming growth factor-beta type II receptor, *Lab. Invest. 92*, 1583-1596.

[28] Tomcik, M., Zerr, P., Pitkowski, J., Palumbo-Zerr, K., Avouac, J., Distler, O., Becvar, R., Senolt, L., Schett, G., and Distler, J. H. (2014) Heat shock protein 90 (Hsp90) inhibition targets canonical TGF-beta signalling to prevent fibrosis, *Ann. Rheum. Dis. 73*, 1215-1222.

[29] Ambade, A., Catalano, D., Lim, A., and Mandrekar, P. (2012) Inhibition of heat shock protein (molecular weight 90 kDa) attenuates proinflammatory cytokines and prevents lipopolysaccharide-induced liver injury in mice, *Hepatology 55*, 1585-1595.

[30] Ambade, A., Catalano, D., Lim, A., Kopoyan, A., Shaffer, S. A., and Mandrekar, P. (2014) Inhibition of heat shock protein 90 alleviates steatosis and macrophage activation in murine alcoholic liver injury, *J. Hepatol. 61*, 903-911.

[31] Myung, S. J., Yoon, J. H., Kim, B. H., Lee, J. H., Jung, E. U., and Lee, H. S. (2009) Heat shock protein 90 inhibitor induces apoptosis and attenuates activation of hepatic stellate cells, *J. Pharmacol. Exp. Ther. 330*, 276-282.

[32] Zou, J., Guo, Y., Guettouche, T., Smith, D. F., and Voellmy, R. (1998) Repression of Heat Shock Transcription Factor HSF1 Activation by HSP90 (HSP90 Complex) that Forms a Stress-Sensitive Complex with HSF1, *Cell 94*, 471-480.

[33] Moore, C. L., Dewal, M. B., Nekongo, E. E., Santiago, S., Lu, N. B., Levine, S. S., and Shoulders, M. D. (2016) Transportable, Chemical Genetic Methodology for the Small Molecule-Mediated Inhibition of Heat Shock Factor 1, *ACS Chem. Biol. 11*, 200-210.

[34] Young, A., Stoilova-McPhie, S., Rothnie, A., Vallis, Y., Harvey-Smith, P., Ranson, N., Kent, H., Brodsky, F. M., Pearse, B. M., Roseman, A., and Smith, C. J. (2013) Hsc70-induced changes in clathrin-auxilin cage structure suggest a role for clathrin light chains in cage disassembly, *Traffic 14*, 987-996.

[35] Xing, Y., Bocking, T., Wolf, M., Grigorieff, N., Kirchhausen, T., and Harrison, S. C. (2010) Structure of clathrin coat with bound Hsc70 and auxilin: mechanism of Hsc70-facilitated disassembly, *EMBO J. 29*, 655-665.

[36] Crowley, V. M., Khandelwal, A., Mishra, S., Stothert, A. R., Huard, D. J., Zhao, J., Muth, A., Duerfeldt, A. S., Kizziah, J. L., Lieberman, R. L., Dickey, C. A., and Blagg, B. S. (2016) Development of Glucose Regulated Protein 94-Selective Inhibitors Based on the BnIm and Radamide Scaffold, *J. Med. Chem. 59*, 3471-3488.

[37] Patel, P. D., Yan, P., Seidler, P. M., Patel, H. J., Sun, W., Yang, C., Que, N. S., Taldone, T., Finotti, P., Stephani, R. A., Gewirth, D. T., and Chiosis, G. (2013) Paralog-selective Hsp90 inhibitors define tumor-specific regulation of HER2, *Nat. Chem. Biol. 9*, 677-684.

[38] Liu, S., and Street, T. O. (2016) 5'-N-ethylcarboxamidoadenosine is not a paralog-specific Hsp90 inhibitor, *Protein Sci. 25*, 2209-2215.

[39] Jin, L., Pahuja, K. B., Wickliffe, K. E., Gorur, A., Baumgartel, C., Schekman, R., and Rape, M. (2012) Ubiquitin-dependent regulation of COPII coat size and function, *Nature 482*, 495-500.

[40] Lee, J. G., Kim, W., Gygi, S., and Ye, Y. (2014) Characterization of the deubiquitinating activity of USP19 and its role in endoplasmic reticulum-associated degradation, *J. Biol. Chem. 289*, 3510-3517.

[41] Altun, M., Kramer, H. B., Willems, L. I., McDermott, J. L., Leach, C. A., Goldenberg, S. J., Kumar, K. G., Konietzny, R., Fischer, R., Kogan, E., Mackeen, M. M., McGouran, J., Khoronenkova, S. V., Parsons, J. L., Dianov, G. L., Nicholson, B., and Kessler, B. M. (2011) Activity-based chemical proteomics accelerates inhibitor development for deubiquitylating enzymes, *Chem. Biol. 18*, 1401-1412.

[42] Taipale, M., Krykbaeva, I., Koeva, M., Kayatekin, C., Westover, K. D., Karras, G. I., and Lindquist, S. (2012) Quantitative analysis of HSP90-client interactions reveals principles of substrate recognition, *Cell 150*, 987-1001.

[43] Manjarrez, J. R., Sun, L., Prince, T., and Matts, R. L. (2014) Hsp90-dependent assembly of the DBC2/RhoBTB2-Cullin3 E3-ligase complex, *PLoS One 9*, e90054.

[44] McGourty, C. A., Akopian, D., Walsh, C., Gorur, A., Werner, A., Schekman, R., Bautista, D., and Rape, M. (2016) Regulation of the CUL3 Ubiquitin Ligase by a Calcium-Dependent Co-adaptor, *Cell 167*, 525-538.

[45] Taipale, M., Tucker, G., Peng, J., Krykbaeva, I., Lin, Z. Y., Larsen, B., Choi, H., Berger, B., Gingras, A. C., and Lindquist, S. (2014) A quantitative chaperone interaction network reveals the architecture of cellular protein homeostasis pathways, *Cell 158*, 434-448.

[46] Nagata, K., Saga, S., and Yamada, K. M. (1986) A major collagen-binding protein of chick embryo fibroblasts is a novel heat shock protein, *J. Cell Biol. 103*, 223-229.

[47] Ito, S., Ogawa, K., Takeuchi, K., Takagi, M., Yoshida, M., Hirokawa, T., Hirayama, S., Shin-Ya, K., Shimada, I., Doi, T., Goshima, N., Natsume, T., and Nagata, K. (2017) A small-molecule compound inhibits a collagen-specific molecular chaperone and could represent a potential remedy for fibrosis, *J. Biol. Chem. 292*, 20076-20085.

[48] Gorur, A., Yuan, L., Kenny, S. J., Baba, S., Xu, K., and Schekman, R. (2017) COPII-coated membranes function as transport carriers of intracellular procollagen I, *J. Cell Biol. 216*, 1745-1759.

[49] Saito, K., Chen, M., Bard, F., Chen, S., Zhou, H., Woodley, D., Polischuk, R., Schekman, R., and Malhotra, V. (2009) TANGO1 facilitates cargo loading at endoplasmic reticulum exit sites, *Cell 136*, 891-902.

[50] Santos, A. J., Raote, I., Scarpa, M., Brouwers, N., and Malhotra, V. (2015) TANGO1 recruits ERGIC membranes to the endoplasmic reticulum for procollagen export, *eLife 4*, e10982.

[51] Saito, K., Yamashiro, K., Ichikawa, Y., Erlmann, P., Kontani, K., Malhotra, V., and Katada, T. (2011) cTAGE5 mediates collagen secretion through interaction with TANGO1 at endoplasmic reticulum exit sites, *Mol. Biol. Cell 22*, 2301-2308.

[52] Venditti, R., Scanu, T., Santoro, M., Di Tullio, G., Spaar, A., Gaibisso, R., Beznoussenko, G. V., Mironov, A. A., Mironov, A., Jr., Zelante, L., Piemontese, M. R., Notarangelo, A., Malhotra, V., Vertel, B. M., Wilson, C., and De Matteis, M. A. (2012) Sedlin controls the ER export of procollagen by regulating the Sar1 cycle, *Science 337*, 1668-1672.

[53] Nogueira, C., Erlmann, P., Villeneuve, J., Santos, A. J., Martinez-Alonso, E., Martinez-Menarguez, J. A., and Malhotra, V. (2014) SLY1 and Syntaxin 18 specify a distinct pathway for procollagen VII export from the endoplasmic reticulum, *eLife 3*, e02784.

[54] Raote, I., Ortega-Bellido, M., Santos, A. J., Foresti, O., Zhang, C., Garcia-Parajo, M. F., Campelo, F., and Malhotra, V. (2018) TANGO1 builds a machine for collagen export by recruiting and spatially organizing COPII, tethers and membranes, *eLife 7*, e32723.

[55] Ma, W., and Goldberg, J. (2016) TANGO1/cTAGE5 receptor as a polyvalent template for assembly of large COPII coats, *Proc. Natl. Acad. Sci. U.S.A. 113*, 10061-10066.

[56] McCaughey, J., Miller, V. J., Stevenson, N. L., Brown, A. K., Budnik, A., Heesom, K. J., Alibhai, D., and Stephens, D. J. (2016) TFG Promotes Organization of Transitional ER and Efficient Collagen Secretion, *Cell Rep. 15*, 1648-1659.

[57] Wilson, D. G., Phamluong, K., Li, L., Sun, M., Cao, T. C., Liu, P. S., Modrusan, Z., Sandoval, W. N., Rangell, L., Carano, R. A., Peterson, A. S., and Solloway, M. J. (2011) Global defects in collagen secretion in a Mia3/TANGO1 knockout mouse, *J. Cell Biol. 193*, 935-951.

[58] Maki, M., Takahara, T., and Shibata, H. (2016) Multifaceted Roles of ALG-2 in Ca(2+)-Regulated Membrane Trafficking, *Int. J. Mol. Sci. 17*, e1401.

[59] Kanadome, T., Shibata, H., Kuwata, K., Takahara, T., and Maki, M. (2017) The calcium-binding protein ALG-2 promotes endoplasmic reticulum exit site localization and polymerization of Trk-fused gene (TFG) protein, *FEBS J. 284*, 56-76.

[60] Hanna, M. G., Block, S., Frankel, E. B., Hou, F., Johnson, A., Yuan, L., Knight, G., Moresco, J. J., Yates, J. R., Ashton, R., Schekman, R., Tong, Y., and Audhya, A. (2017) TFG facilitates outer coat disassembly on COPII transport carriers to promote tethering and fusion with ER–Golgi intermediate compartments, *Proc. Natl. Acad. Sci. U.S.A. 114*, E7707-E7716.

[61] Zhou, T., Zimmerman, W., Liu, X., and Erikson, R. L. (2006) A mammalian NudC-like protein essential for dynein stability and cell viability, *Proc. Natl. Acad. Sci. U.S.A. 103*, 9039-9044.

[62] Zhu, X. J., Liu, X., Jin, Q., Cai, Y., Yang, Y., and Zhou, T. (2010) The L279P mutation of nuclear distribution gene C (NudC) influences its chaperone activity and lissencephaly protein 1 (LIS1) stability, *J. Biol. Chem. 285*, 29903-29910.

[63] Huang, J., Roberts, A. J., Leschziner, A. E., and Reck-Peterson, S. L. (2012) Lis1 acts as a "clutch" between the ATPase and microtubule-binding domains of the dynein motor, *Cell 150*, 975-986.

[64] Toropova, K., Zou, S., Roberts, A. J., Redwine, W. B., Goodman, B. S., Reck-Peterson, S. L., and Leschziner, A. E. (2014) Lis1 regulates dynein by sterically blocking its mechanochemical cycle, *eLife 3*, e03372.

[65] Takahara, T., Inoue, K., Arai, Y., Kuwata, K., Shibata, H., and Maki, M. (2017) The calcium-binding protein ALG-2 regulates protein secretion and trafficking via interactions with MISSL and MAP1B proteins, *J. Biol. Chem. 292*, 17057-17072.

[66] Zhang, M., Boter, M., Li, K., Kadota, Y., Panaretou, B., Prodromou, C., Shirasu, K., and Pearl, L. H. (2008) Structural and functional coupling of Hsp90- and Sgt1-centred multi-protein complexes, *EMBO J. 27*, 2789-2798.

[67] Glick, B. S. (2017) New insights into protein secretion: TANGO1 runs rings around the COPII coat, *J. Cell Biol. 216*, 859-861.

[68] Prodromou, C. (2016) Mechanisms of Hsp90 regulation, *Biochem. J. 473*, 2439-2452.

[69] Cox, N. J., Unlu, G., Bisnett, B. J., Meister, T. R., Condon, B. M., Luo, P. M., Smith, T. J., Hanna, M., Chhetri, A., Soderblom, E. J., Audhya, A., Knapik, E. W., and Boyce, M. (2018) Dynamic Glycosylation Governs the Vertebrate COPII Protein Trafficking Pathway, *Biochemistry 57*, 91-107.

[70] Zhang, L., Syed, Z. A., van Dijk Hard, I., Lim, J. M., Wells, L., and Ten Hagen, K. G. (2014) O-glycosylation regulates polarized secretion by modulating Tango1 stability, *Proc. Natl. Acad. Sci. U.S.A. 111*, 7296-7301.

[71] Sarmah, S., Barrallo-Gimeno, A., Melville, D. B., Topczewski, J., Solnica-Krezel, L., and Knapik, E. W. (2010) Sec24D-dependent transport of extracellular matrix proteins is required for zebrafish skeletal morphogenesis, *PLoS One 5*, e10367.

CHAPTER IV
Adapting Secretory Proteostasis and Function through the Unfolded Protein Response

Summary

Cells address challenges to protein folding in the secretory pathway by engaging endoplasmic reticulum (ER)-localized protective mechanisms that are collectively termed the unfolded protein response (UPR). By the action of the transmembrane signal transducers IRE1, PERK, and ATF6, the UPR induces networks of genes whose products alleviate the burden of protein misfolding. The UPR also plays instructive roles in cell differentiation and development, aids in the response to pathogens, and coordinates the output of professional secretory cells. These functions add to and move beyond the UPR's classical role in addressing proteotoxic stress. Thus, the UPR is not just a reaction to protein misfolding, but also a fundamental driving force in physiology and pathology. Recent efforts have yielded a suite of chemical genetic methods and small molecule modulators that now provide researchers with both stress-dependent and -independent control of UPR activity. Such tools provide new opportunities to perturb the UPR and thereby study mechanisms for maintaining proteostasis in the secretory pathway. Numerous observations now hint at the therapeutic potential of UPR modulation for diseases related to the misfolding and/or aggregation of ER client proteins. Growing evidence also indicates the promise of targeting ER proteostasis nodes downstream of the UPR. Here, we review selected advances in these areas, providing a resource to inform ongoing studies of secretory proteostasis and function as they relate to the UPR.

Contributions

Introduction

The endoplasmic reticulum (ER) is responsible for secretory proteostasis, involving the coordinated folding, processing, quality control, and trafficking of ~1/3 of the proteome. Protein folding is a highly complex and error-prone process, requiring a delicate balance between function and risk of aggregation in crowded biological microenvironments where total protein concentrations can range from 100–400 mg/mL.[1, 2] ER clients, which include secreted, membrane, and lysosomal proteins, face additional challenges, including unique post-translational modifications (e.g., N-glycosylation) that require specialized cellular machinery,[3] oxidative folding processes associated with selective disulfide bond formation,[4] and both spatial and temporal restraints on the completion of folding, modification, assembly, and transport steps. Cells account for this complexity via a diverse array of folding[2] and quality control mechanisms,[5] some of which are only recently coming to light. The resulting balance of protein synthesis, folding, and recycling is essential for health. Dysregulated proteostasis in the secretory pathway underpins a diverse array of diseases.

Maintaining secretory proteostasis requires the ability to dynamically respond to challenges such as protein misfolding, often by large-scale remodeling of the ER and the ER proteostasis environment.[6] The unfolded protein response (UPR; *Figure 4.1*) is the central stress response pathway involved. The three arms of the metazoan UPR are controlled by the signal transducers IRE1, PERK, and ATF6.[7-10] Activation of these ER transmembrane proteins induces a transcriptional response mediated by three transcription factors, XBP1s,[11, 12] ATF4,[13, 14] and ATF6f (ATF6-fragment).[15] This coordinated transcriptional response alleviates the burden of protein misfolding in the secretory pathway by upregulating ER chaperone, quality control, and secretion mechanisms.[13, 15, 16] UPR activation also inhibits protein translation to lower the net nascent protein load on the ER, a process mediated primarily by PERK activation and subsequent phosphorylation of eIF2α,[8] but also influenced by the selective degradation of ER-directed mRNA transcripts by IRE1.[17, 18] If proteostasis cannot be restored, pro-apoptotic mechanisms within the UPR are engaged leading to programmed cell death primarily through induction of the transcription factor CHOP downstream of PERK.

While extensive research has yielded a relatively well-defined picture of the UPR, the discovery of new regulatory mechanisms continues to shape our understanding of how the UPR relates to ER homeostasis. The ER not only functions as a protein-folding factory, but also participates in calcium storage and lipid biosynthesis.[19] Along these lines, a recent study highlighted the capacity of IRE1's membrane spanning domain to activate the protein in response to lipid perturbation even when the luminal protein misfolding stress-sensing domain is deleted using CRISPR/Cas9.[20] Moreover, the ER is involved in cellular responses to oxidative stress, metabolic imbalance, and pathogen invasion.[21-23] Each of these processes is modulated by the UPR. Thus, despite its moniker, the UPR is not simply a reaction to protein misfolding, but is instead a fundamental driving force for physiology and pathology. The central roles of the UPR in health and disease have catalyzed the development of methods to modulate the UPR, with the goal of better understanding key regulatory axes and identifying opportunities to influence phenotypic outcomes. Below, we review our current picture of the metazoan UPR in the context of secretory proteostasis, from its connections to health and disease, to methods for selectively perturbing the UPR and their potential applications in disease therapy, to efforts to target downstream nodes in ER proteostasis.

The UPR in Health and Disease

Key functional nodes within the ER proteostasis network include chaperones, quality control mechanisms, post-translational modifiers, and trafficking pathways (*Figure 4.2*). Each of these nodes is dynamically regulated by the UPR to match proteostatic capacity to demand, thereby maintaining balanced levels of protein folding and quality control both during normal cellular function and under stressful conditions.
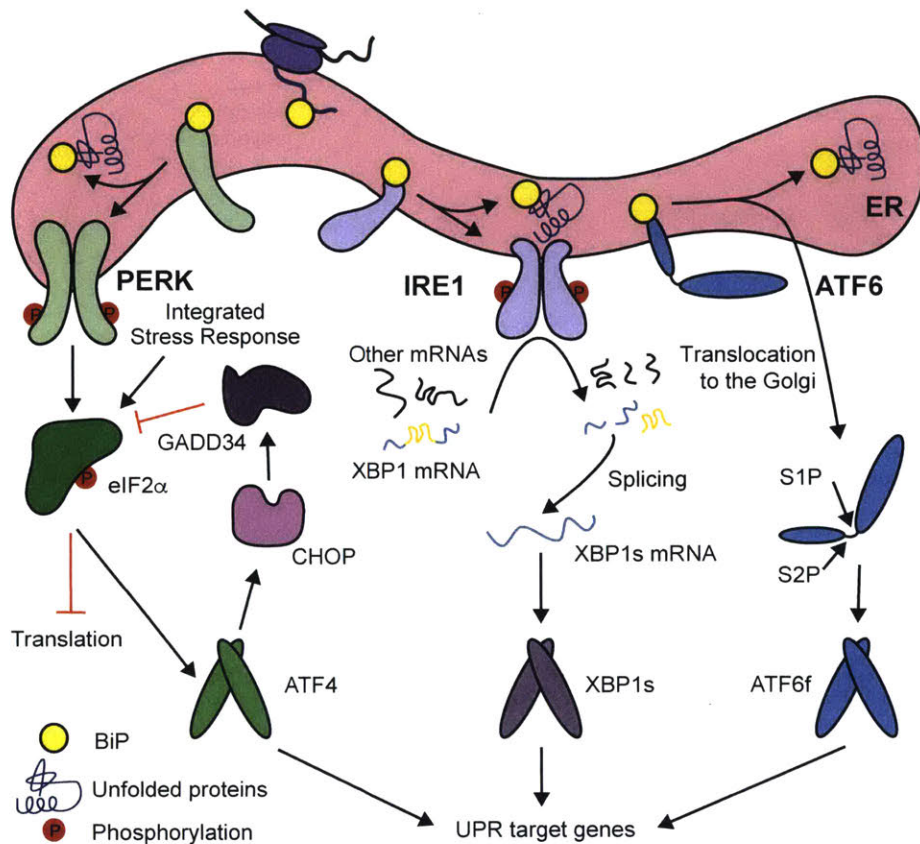
Figure 4.1: *The unfolded protein response (UPR).* Accumulation of misfolding proteins in the endoplasmic reticulum (ER) activates the transmembrane protein UPR signal transducers PERK, IRE1, and ATF6. Dimerization and auto-phosphorylation of PERK and IRE1, or trafficking to the Golgi and subsequent proteolytic processing of ATF6, result in the production of the UPR transcription factors by enhancing translation of ATF4, splicing *XBP1* mRNA to yield *XBP1s*, and releasing ATF6f from the Golgi membrane. These transcription factors proceed to the nucleus and remodel the ER proteostasis environment by upregulating chaperones, quality control components, and other UPR target genes to maintain or recover secretory proteostasis. PERK can globally reduce the nascent protein load on the ER via phosphorylation of eIF2α, a pathway that can be similarly induced by the integrated stress response. The RNase domain of activated IRE1 degrades several ER-targeted transcripts and may play a related role.

Development, Professional Secretory Cells, and Immunity. The UPR has critical roles during development that have been demonstrated in several model systems.[24, 25] In particular, UPR activation appears to upregulate ER-resident chaperones and signaling pathways, which work collectively to relieve stress and regulate development in differentiating cells.[26-28] For example, upon beta-cell differentiation into plasma cells, the ER undergoes extensive XBP1-driven expansion,[28, 29] in part to accommodate high levels of antibody synthesis. Beta-Cell
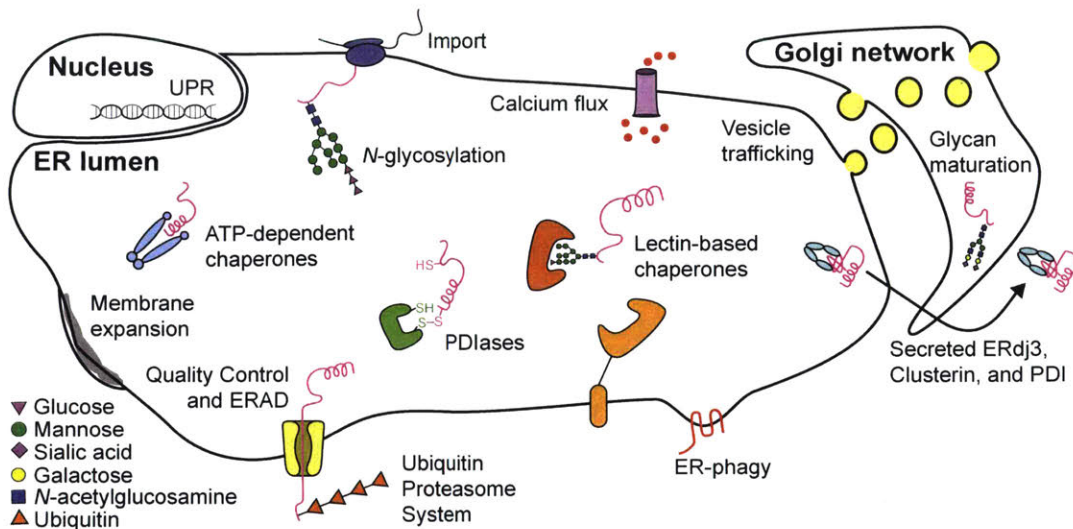
Figure 4.2: *Representative nodes in the secretory proteostasis network.* Diverse proteins and pathways collectively modulate folding, secretion, quality control, and/or degradation of ER clients. ATP-dependent chaperones and PDIases assist in the folding of client proteins, as do lectin-based chaperones such as calnexin and calreticulin. Terminally misfolded proteins are typically cleared by ER-associated degradation (ERAD) via the ubiquitin-proteasome system. ER-phagy can serve as a counterpart to membrane expansion mechanisms, reducing organelle size to regulate ER proteostasis. Meanwhile, calcium flux, vesicle trafficking, and UPR-mediated changes in the chaperone:client balance, import, disulfide bond formation, and *N*-glycosylation of nascent polypeptides help to maintain or create favorable folding conditions and buffer ER protein-folding capacity. A handful of chaperones, including ERdj3, can accompany proteins to the extracellular space.

differentiation *in vitro* also induces the UPR-regulated proteins XBP1s, BiP and Grp94. Notably, the process occurs without expression of CHOP or inhibition of protein translation, suggesting that a physiologic UPR need not involve all three UPR arms, in contrast to the case of attenuating stress-induced protein misfolding. Moreover, induction of XBP1, BiP and Grp94 transcripts apparently occurs prior to any significant protein-folding load on the ER, suggesting further differences between developmental and stress-associated signaling pathways.[30] Other professional secretory cells such as pancreatic beta-cells, hepatic cells, and osteoblasts also must sustain high rates of ER client protein synthesis, folding, and secretion, and thus rely on the UPR and its downstream signaling mechanisms for survival and function.

Other work highlights roles of the UPR in cellular responses to pathogen invasion. Binding of unfolded cholera toxin A subunit induces IRE1α ribonuclease activity but not the canonical UPR involving PERK and ATF6.[31] The fragments of endogenous mRNA produced by IRE1α prompt RIG-I to activate NF-κB and interferon signaling. Other work indicates that Toll-like receptors (TLRs) in macrophages promote splicing of XBP1 to optimize the production of proinflammatory cytokines,[32] although XBP1s can also be essential in protecting against the effects of prolonged inflammation.[33, 34] Intriguingly, viral pathogens are capable of hijacking the UPR to promote proliferation in host cells. For example, IRE1 activity is critical for the replication of at least some strains of the influenza virus.[35] In contrast, HSV1 suppresses both PERK and IRE1 signaling: glycoprotein B interacts with the luminal domain of PERK to block kinase activation, the late viral protein g1 34.5 recruits PP1α to dephosphorylate eIF2α, and the UL41

protein acts as an endoribonuclease to degrade XBP1 mRNA.[36] Similarly, recent studies of *Legionella pneumophila*, the organism responsible for Legionnaires' disease, show that the pathogen forestalls a prototypical ER stress response by repressing translation of a subset of UPR-associated genes to prevent host-cell apoptosis that would otherwise be induced.[37, 38] The relevant bacterial effector proteins may serve as springboards for biomimetic approaches to modulate UPR pathways.

Emerging Functions of the UPR. Beyond established roles in development and immunity, new functions for and consequences of the UPR continue to emerge. ER recycling via ER-phagy is critical for ER homeostasis,[39] and several constituent biochemical pathways were recently mapped.[40, 41] A possible role for the IRE1-XBP1s arm of the UPR in inducing such ER-phagy may exist.[42] By reducing organelle size and/or disposing of dysfunctional ER regions, UPR induction of selective ER-phagy could serve as a counterpart to membrane expansion mechanisms for resolving ER stress.[43]

The discovery that the IRE1-XBP1s axis of the UPR is responsible for cell non-autonomous UPR activation[44] is also intriguing. Such cell-to-cell communication of stress is likely to have important biological consequences that merit further investigation. ATF6 activation was recently shown to increase not just expression but also secretion of ERdj3, an ER-localized HSP40 co-chaperone.[45] The consequent co-secretion of ERdj3 with misfolded client proteins may be protective for the origin cell or ameliorate harmful protein aggregation in the extracellular milieu. Stress-induced ERdj3 secretion thus provides a mechanism by which the UPR can modify not just ER but also extracellular proteostasis.

Emerging functions of the UPR described above focus on direct modulation of protein folding and production. In addition to these mechanisms, a role for the ER in regulating the extent of protein post-translational modifications has emerged. For example, two groups showed that the UPR can modulate hexosamine biosynthesis to promote ER client clearance and prolong life in the face of chronic protein misfolding stress.[46, 47] These studies suggest that UPR activation, and especially the IRE1-XBP1s arm of the UPR, may enhance the extent of client protein *N*- and *O*-glycosylation and/or modify oligosaccharyltransferase efficiency. While the consequences merit further investigation, *N*-glycosylation promotes both ER client folding by providing access to the lectin-based chaperone machinery and the identification of misfolded proteins for ER-associated degradation (ERAD) via the lectin-based quality control machinery, providing a potential rationale for IRE1-XBP1s enhanced *N*-glycosylation.[48]

Surprisingly, the UPR can also remodel the actual molecular architecture of N-glycans added to ER client proteins by modulating their biosynthesis.[49] Stress-independent activation of XBP1s changes transcript levels of *N*-glycan modifying enzymes, leading to altered mature glycan structures on model secreted *N*-glycoproteins. More work is required to establish the biological relevance and consequences of this phenomenon. However, this newly established connection between *N*-glycan signatures and the UPR suggests that the UPR may unexpectedly influence processes such as cell-cell interactions, cell-matrix interactions, and trans-cellular communication by actually modifying the molecular structure of secreted ER clients.[49] In the following chapters, we report the first evidence that XBP1s activation does indeed remodel the distribution of cellular *N*-glycan structures (Chapter V). We also present evidence that the *N*-glycan on the collagen-I C-terminal propeptide domain buffers collagen-I folding against ER stress and destabilizing mutations (Chapter VI). These findings highlight cross talk between UPR signaling and *N*-glycosylation as a new mechanism for regulating secretory proteostasis.

Dysregulated ER Proteostasis and Disease. When proteostasis networks function properly, cells maximize production of properly folded, functional proteins. Meanwhile, quality control mechanisms ensure that only folded proteins are transported to their final locations, while

production of misfolded and aggregated proteins is minimized (*Figure 4.3A*). The UPR regulates this process by sensing the accumulation of misfolded proteins, whether due to genetic mutations or adverse physiological conditions, and remodeling the ER proteostasis network to resolve emerging problems.

Chronically dysregulated ER proteostasis, unresolved by the UPR, leads to diverse protein misfolding and aggregation-related diseases. For many mutations that destabilize or prevent the folding of a protein, the UPR may in principle have the potential to resolve the proteostatic defect – if it is activated. However, just one mutant protein misfolding in a background of thousands of well-behaved proteins may not always be a sufficient signal to trigger a protective UPR. In other cases, the ER may be overwhelmed by high concentrations of an aggregating mutant protein, leading to chronic ER stress and cellular apoptosis. In either scenario, pharmacologic perturbation of the UPR could be therapeutically useful. Moreover, many cancer cells rely on constitutive activation of pro-survival pathways (in particular the IRE-XBP1s arm) within the UPR.[50-52] This observation suggests that UPR inhibition could also prove valuable for diseases that do not stem directly from protein misfolding.

Concept Summary. Even this cursory survey of the roles of the UPR in health and disease reveals a striking functional spectrum. Beyond the traditional UPR, IRE1, PERK, and ATF6 participate in various normal physiological processes, upregulating UPR-associated transcripts to facilitate differentiation and sustain the activity of professional secretory cells. Studies on different models of pathogen infection demonstrate that UPR arms also can be selectively suppressed or activated in order to bypass cytotoxic stress-signaling pathways and promote cell survival. Such mechanisms are appealing starting points for efforts to better understand and manage cases of chronic ER stress associated with disease. Finally, recent findings suggest new roles for the UPR in regulating extracellular proteostasis and signaling, perhaps via post-translational alterations in ER client protein molecular structures.[49]

Targeting the UPR to Modulate ER Proteostasis

The intrinsic functions of the UPR in development and immunity, regulating cell survival and death, and resolving proteostatic stress have motivated extensive method development efforts to uncover small molecule UPR activators and inhibitors. Such tools are enabling detailed dissection of innate UPR function and revealing the therapeutic potential associated with targeting the UPR. This section reviews selected chemical genetic and small molecule approaches to modulate the UPR, and briefly describes insights obtained using such methods in disease model systems.

Stress-Dependent Methods to Modulate the UPR. Traditional approaches to activate the UPR rely on small molecules that cause ER stress. Such compounds include tunicamycin, which inhibits protein *N*-glycosylation; thapsigargin, which disrupts ER calcium homeostasis; and dithiothreitol, which reduces disulfides. All these methods globally and strongly activate the UPR by inducing extensive protein misfolding and aggregation. While such strategies have proven valuable for mapping UPR signaling pathways, they are less suited for mechanistic work as they induce high and physiologically irrelevant levels of stress. Moreover, these compounds are not helpful for therapeutic proof-of-principle studies, as the cytotoxic and pleiotropic side effects of their use obscure any potentially beneficial effects of UPR activation.

One approach to bypass these issues is to administer very low concentrations of ER toxins to induce only moderate stress.[53] Alternatively, a chemical genetic strategy was recently pioneered to transiently activate the UPR without inducing apoptosis.[54] The method involves expressing an ER-targeted HaloTag protein that can be conditionally destabilized by conjugation to a small molecule hydrophobic tag (HyT36). Treatment of ER-targeted HaloTag-expressing cells with HyT36 destabilizes the protein, causing mild but acute accumulation of

misfolded proteins in the ER and consequent UPR activation. While the mechanistic details are not yet fully elucidated, the ERHT/HyT36 system provides a valuable tool to study aspects of global UPR activation distinct from the stress-independent, arm-selective techniques described below.

Stress-Independent Methods to Modulate the UPR. An important element of any strategy to discover and validate small molecule UPR arm activators is verifying that the compounds directly and selectively activate a particular arm of the UPR, without inducing ER protein misfolding or other forms of stress. To address this concern, recent work has delineated comprehensive transcriptional profiling strategies to validate stress-independent, arm-specific UPR activators.[55] While several small molecule leads have been developed that activate endogenous IRE1 and/or ATF6,[55, 56] potency and selectivity, as well as mechanistic characterization, will benefit from further work. Selective small molecule activators of endogenous PERK have thus far proven elusive. On the other hand, salubrinal,[57] guanabenz,[58] ISRIB,[59] and Sephin1[60] can modulate the PERK-induced translational block either up or down by altering eIF2a phosphorylation status (*Figure 4.1*). Other explanations for the biological phenotypes induced by guanabenz and Sephin1 may also prove relevant.[61] Notably, these compounds also modulate overlapping aspects of the integrated stress response, rendering such small molecules neither UPR- nor PERK-specific. For example, salubrinal also targets the constitutively expressed eIF2a phosphatase CReP.[57] Finally, while small molecule inhibitors of endogenous IRE1 have been well characterized,[62-64] compounds that inhibit ATF6 are only now beginning to emerge.[65]

Given the paucity of small molecules to modulate arms of the UPR in a stress-independent manner, over the past fifteen years the field has turned to chemical genetic approaches. One strategy involves fusion of PERK's kinase domain to a tandem-modified FK506-binding domain, Fv2E. Treatment with the small molecule AP20187 induces Fv2E-PERK dimerization, auto-phosphorylation, and downstream PERK signaling.[66, 67] The Fv2E fusion strategy was also applied to recapitulate the regulated RNase activity of IRE1.[68] A complementary strategy based on bump-hole protein engineering employs an I624G mutation to allow binding of a small molecule ligand in the ATP pocket of IRE1, but not other kinase active sites, resulting in selective, stress-independent activation of IRE1's RNase domain. Addition of the ligand 1NM-PP1 to cells expressing IRE1a^{I624G} results in selective IRE1-dependent XBP1 mRNA splicing in the absence of ER stress.[69-71] These PERK and IRE1-targeted chemical genetic strategies allow for uncoupling of UPR arm activation from stress, but require careful engineering of cells to ensure minimal background signaling and robust inducibility. Their application has helped to clarify the consequences of ER stress. For example, chronic PERK activity, but not 1NM-PP1-induced IRE1 activity, is cytotoxic, suggesting a model in which the duration of PERK signaling regulates cell survival versus apoptosis decisions.[66] Intriguingly, the advent of RNase inhibitors targeting wild-type IRE1, such as KIRA6, revealed that IRE1 can also promote apoptosis through formation of oligomers with hyperactive RNase activity.[72] Thus, cell fate may not necessarily be determined by opposing signals from IRE1 and PERK, but rather by the degree and timescale of activation (whether homeostatic or cytotoxic) experienced by both kinases.

An alternative chemical genetic strategy is to confer small molecule-dependent activity upon the UPR transcription factors, independent of the upstream signal transducers IRE1, PERK, and full-length ATF6. This approach has proven most successful for XBP1s and ATF6f (*Figure 4.1*). In the simplest case, tetracycline (Tet)-regulated expression of the UPR transcription factors ATF6f[73] and XBP1s[16, 74] under control of the Tet-repressor domain has been applied in several model systems. A challenge is the requirement for incorporation of the Tet-repressor in target cells and tissues. Moreover, Tet control of protein expression is rarely dose-dependent, resulting in high, non-physiologic levels of the Tet-inducible gene.[16, 75] Such

overexpression of UPR transcription factors can cause off-target mRNA upregulation and often induces apoptosis, limiting its applicability.

More recently, destabilized domains[76, 77] were leveraged to achieve orthogonal activation of XBP1s and/or ATF6f in a single cell. Fusing a destabilized variant of the dihydrofolate reductase from *E. coli* (DHFR) to the N-terminus of ATF6f results in a constitutively expressed DHFR.ATF6f fusion protein that is directed to rapid proteasomal degradation. Addition of the small molecule pharmacologic chaperone trimethoprim stabilizes DHFR, preventing degradation and allowing the DHFR-fused ATF6f transcription factor to function.[16] A related approach was used to control the activity of the XBP1s transcription factor, via fusion to an FKBP12 destabilized domain that can be stabilized by the small molecule Shield-1.[75] The method requires minimal cell engineering, can be used in virtually any cell line of interest (including primary cells), and provides for highly dose-dependent control of transcription factor activity within the physiologically relevant regime. Moreover, fusion of destabilized domains to dominant negative versions of the UPR transcription factors can permit small molecule-dependent inhibition of endogenous transcription factor activity.[75] The advantages of destabilized domains for conferring small molecule control onto UPR and other stress-responsive transcription factors[78] have led to their adoption by a number of research groups in studies of stress-responsive signaling.

Chemical genetic strategies provide valuable temporal and ligand concentration-dependent control of UPR arm activity. However, they must be applied with caution to ensure that engineered domains are compatible with the natural function of the target protein. IRE1^{I624G}, for example, does not share the same mechanism of activation as wild-type IRE1, and may exhibit altered kinase activity; similarly, the fusion of destabilized domains to UPR transcription factors may alter their function. Expression levels of engineered proteins, whether dictated by the leakiness of the system or by copy number per cell, must also be optimized. Thus, while such strategies have enabled substantial progress in the UPR field (especially with regard to potential therapeutic benefits of UPR activation in disease model systems, discussed below), they do not ablate the need for highly selective and potent small molecule UPR modulators that function independently of protein misfolding stress.

Activating the UPR to Address Diseases Linked to Dysregulated ER Proteostasis. Small molecule and chemical genetic methods to modulate the UPR, including those described above, have been extensively applied to test whether UPR-mediated remodeling of the ER proteostasis network is a viable strategy to address ER protein misfolding- and aggregation-related diseases. In particular, various approaches have been used to address dysregulated proteostasis associated with the three types of defects highlighted in *Figure 4.3*.

In the first category of pathologic ER proteostasis defects, protein misfolding/aggregation diseases manifest owing to a combination of underactive quality control and insufficient folding activity (*Figure 4.3B*). The insufficient ER proteostasis environments that characterize such disorders stem from the inability to sense misfolding of an individual mutant protein and/or the permitted secretion of malformed, aggregration-prone proteins into the extracellular milieu.

One example of underactive quality control is the transthyretin (TTR) amyloidoses, wherein misfolded or unstable TTR escapes the ER and later aggregates in peripheral tissues.[79] TTR is a secreted tetrameric protein whose extracellular disassembly provides the necessary template for oligomers to form.[80, 81] Numerous mutations destabilize the tetramer, accelerating disassembly, aggregation, and disease pathology. Two treatments currently exist: (1) gene therapy via liver transplantation, as the liver is the primary source of TTR, and (2) small molecule pharmacologic chaperones that stabilize secreted TTR tetramers and thereby prevent oligomerization.[79]
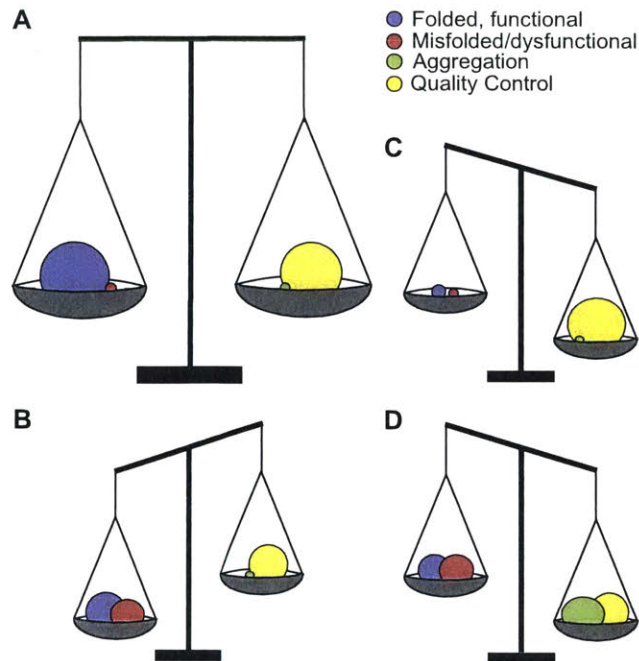
Figure 4.3: *Proteostasis (im)balances in health and disease.* (A) When proteostasis is properly balanced, production of folded, functional client proteins and quality control surveillance are maximized, while misfolding and aggregation are minimized. Maintaining the balance between production of folded, functional proteins and quality control or clearance of misfolded, non-functional proteins is essential for health. In protein misfolding/aggregation disorders linked to defects in ER proteostasis, (B) insufficient quality control, (C) hyperactive quality control and failed folding, and/or (D) failed clearance leading to intracellular protein aggregation and chronic ER stress/UPR activation can all cause a pathologic loss of proteostasis balance. Modulating proteostasis network activities holds potential to resolve such defects.

An alternative potentially synergistic possibility is enhancing the stringency of ER quality control, perhaps via UPR-mediated remodeling of the ER proteostasis network to reduce the secretion and promote the degradation of destabilized TTR variants. By lowering misfolding TTR concentrations in the sera, this strategy would likely reduce pathologic oligomer formation. Indeed, stress-independent activation of ATF6 using the chemical genetic DHFR.ATF6f construct preferentially directs destabilized TTR variants towards ERAD, drastically reducing their secretion.[16, 82] Importantly, the approach does not influence the secretion of stable, wild-type TTR. These results hint at the potential of stress-independent UPR activation to improve cellular capacity to prevent the secretion of misfolded, potentially toxic species. Stress-independent XBP1s and/or ATF6 activation is also effective in preventing the secretion of amyloidogenic light chain.[83] Such findings motivated significant efforts to identify small molecule activators of ATF6 that appear very promising and may provide leads for clinical development.[55]

A second, less intuitive category of ER proteostasis defects arises when cells execute an overly stringent survey of client protein folding status (*Figure 4.3C*). As in cases of healthy proteostasis, aggregation is minimized, but a slow-folding or moderately misfolded ER client is prematurely directed to degradation when, if given sufficient time or assistance, it could have adopted a sufficiently folded, functional state to avert pathology. Such indiscriminate degradation owing to excessively stringent quality control results in loss-of-function phenotypes. Onset of disease occurs when the amount of protein present and able to function in its natural location falls below the required threshold of essential biological activity. Relaxed quality control

may, in these cases, allow such ER clients to adopt sufficiently folded, functional states to traffic to their destinations and thereby ameliorate disease.

One case where hyperactive quality control appears to be associated with pathology is the lysosomal storage diseases. An example is Gaucher's disease, which is caused by mutations that destabilize glucocerebrosidase (GCase). GCase is a hydrolytic enzyme that folds in the ER and is then trafficked to the lysosome, where it degrades encapsulated glycolipids. Misfolding-prone and destabilized GCase variants can be identified by overly stringent ER quality control machinery and targeted for degradation by the proteasome.[84] Even a slight improvement in trafficking to the lysosome would likely provide therapeutic benefit. Mu and coworkers discovered that global UPR activation by induction of mild ER stress significantly enhances trafficking of the L444P GCase variant to the lysosome in an IRE1- and PERK-dependent manner, with a corresponding decrease in degradation.[85] The resulting enhancement in lysosomal GCase activity can be synergistically enhanced by co-administration of GCase-stabilizing pharmacologic chaperones. More recently, both modulation of Ca^{2+} concentrations in the ER and knockdown of specific ER proteostasis network components like FKBP10 and ERdj3, which direct GCase to degradation, were shown to provide similar benefits.[86-88] Thus, reducing quality control stringency via UPR-mediated and/or more targeted perturbations of the ER proteostasis network may hold potential for diseases associated with hyperactive quality control.

A third category of disease-causing ER proteostasis defects emerges when protein misfolding or aggregation in the ER causes chronically unresolved stress and consequent cellular dysfunction. While the IRE1 and ATF6 arms of the UPR can help cells tolerate such chronic stress, as has been observed in fly models of retinal degeneration,[89, 90] chronic activation of the PERK arm leads directly to apoptosis. Thus, long-term UPR activation can severely threaten tissue homeostasis (Figure 4.3D), engendering a vicious cycle in which protein misfolding caused by mutations, aging, and the like results in cell death, increasing the demand on surviving cells in a given tissue to produce folded, functional protein. The increased protein-folding load placed on the surviving cells results in additional chronic ER stress, UPR activation, and further tissue loss. Unresolved chronic ER stress is associated with a wide variety of disorders and diseases, including obesity and neurodegeneration.[91] Strategies to inhibit UPR-mediated apoptosis could be valuable in these situations, as could methods to alleviate chronic stress by enhancing folding or quality control (as in Figures 4.3B and 4.3C).

UPR-mediated ER proteostasis network remodeling has shown promise for several diseases where aggregation is observed in the ER. For example, ATF6 and PERK activation enhance clearance of mutant, aggregating rhodopsin variants that cause retinitis pigmentosa.[92] Similarly, ATF6 activation assists the clearance of the aggregating Z variant of α-1-antitrypsin that normally causes inflammation-induced hepatotoxicity (by a gain-of-function mechanism) and/or proteolytic damage to lung tissue (by a loss-of-function mechanism).[93] Clearance of these protein aggregates from the ER is thus an attractive therapeutic strategy. An alternative approach, instead of activating UPR arms, is to simply lower the net protein load on the ER and allow endogenous mechanisms to resolve the proteostasis defect. This strategy involves application of molecules that function downstream of PERK to translationally attenuate the protein folding challenge faced by the ER. Related strategies have shown promise in both diabetes- and amyotrophic lateral sclerosis-related model systems.[57, 58, 60]

While the division of disease-causing ER proteostasis defects into the three categories in Figures 4.3B–D provides valuable context, few diseases fit seamlessly. For example, lax quality control that permits excessive accumulation of mutant GCase variants may cause chronic, damaging ER stress and/or disrupt general protein trafficking mechanisms, a mechanism that could well be biologically relevant for Gaucher's disease. Intriguingly, GCase deficiency promotes neurodegeneration in Parkinson's disease, possibly through the loss of stabilizing effects on α-synuclein oligomers or by inducing lysosome dysfunction, providing

further evidence that multiple shared mechanisms may underlie and connect otherwise disparate pathologies.[94] Another pertinent example is the collagenopathies such as osteogenesis imperfecta,[95] which appear to encompass all three types of ER proteostasis defects described above. Collagen is a challenging protein to fold,[96] with both globular and triple-helical regions, abundant and critical post-translational modifications, and dimensions too large for standard COP-II transport vesicles. Numerous mutations in collagen strands and in components of the collagen proteostasis network cause diverse collagenopathies, depending on the specific type of collagen involved. Mutations can lead to intracellular collagen accumulation and chronic ER stress, can cause collagen strands to be subjected to excessive quality control and degradation, and/or produce variants that functionally disrupt tissue architecture.[97, 98] Whether ER proteostasis imbalances can be remedied in osteogenesis imperfecta and the other collagenopathies remains to be determined, but such pathologies do highlight the complexity associated with many protein misfolding-related diseases.

Concept Summary. Directly targeting particular arms of the UPR via both stress-dependent and -independent approaches will require minimizing off-target effects and separating overlapping activities, whether this entails uncoupling an apoptotic, late-stage UPR from a stress-sensing, early-stage UPR, uncoupling IRE1 kinase from RNase activity, or uncoupling UPR activation from ER stress. While the development of improved methods continues, the application of existing tools to various model systems confirms the therapeutic potential of adapting secretory proteostasis through the UPR. Although the details of the energy landscapes of misfolding and aggregation-prone proteins are unique, we note that they often share features in terms of their interactions with the ER proteostasis network. Focusing on commonalities may thus yield broadly applicable rather than disease-specific therapeutic approaches.

Beyond the UPR

The preceding discussion emphasizes the substantial promise of UPR modulation to resolve disease-causing ER proteostasis defects. However, global remodeling of the ER proteostasis network by UPR perturbation could have deleterious effects on wild-type, well-behaved proteins or on innate functions of the UPR. Targeted drug delivery to diseased tissues could help sidestep off-target effects, but remains difficult. Alternatively, although the UPR is the master regulator of the ER proteostasis network (*Figure 4.2*), it is still just one of many hubs (*Figure 4.4*). In many cases, it may be possible to design small molecules that target proteostasis nodes downstream of the UPR's transcription/translation response to resolve proteostasis defects. Such approaches could have a larger therapeutic window owing to more precise manipulation of specific mechanisms. Testing these strategies requires selective chemical biology and genetic tools to manipulate the activities of chaperones, quality control pathways, and the like. In addition to the example of translational attenuation discussed above, this section briefly considers two other nodes downstream of the UPR where targeted investigation has become increasingly possible: the ER's ATP-dependent chaperones and ERAD-mediated quality control.

Targeting ATP-Dependent Chaperone Systems in the ER. The ATP-dependent chaperones Grp94 (a heat shock protein 90, or Hsp90, isoform), BiP (a heat shock protein 70, or Hsp70, isoform) and their corresponding Hsp40-like co-chaperones and nucleotide exchange factors are central components of the ER proteostasis network (*Figure 4.2*). They play key roles both in assisting client protein folding and in recognizing and targeting misfolded proteins for quality control. Given the importance of dysregulated proteostasis in cancer and other diseases, Hsp90 inhibitors have been particularly sought after as therapeutics. A number of potent inhibitors have been discovered[99-102] and show promise in cancer model systems. However, the high sequence and structural similarity of Hsp90 isoforms[103] poses difficulties for the design of selective small
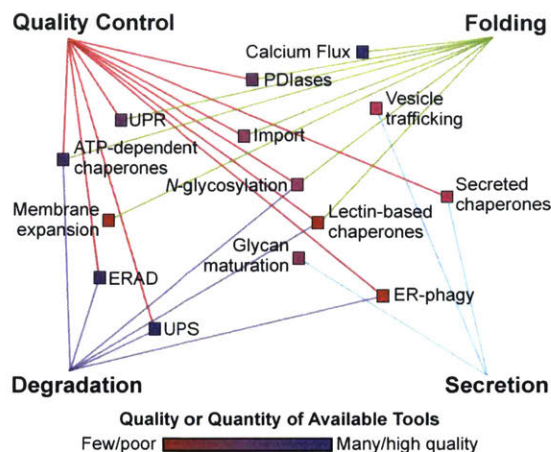
Figure 4.4: *Targeting functional nodes in the ER proteostasis network.* Representative hubs in the ER proteostasis network; colored boxes give a qualitative measure of methods available to modulate each node. Red: few and/or poorly characterized tools; blue: many and/or well-characterized tools. Nodes from *Figure 4.2* are aligned with the broader axes of quality control, folding, degradation, and secretion. For many diseases, coordinated regulation of multiple nodes in one or more of the relevant pathways may produce the most therapeutic benefit.

molecule inhibitors. Indeed, the most widely used Hsp90 inhibitors target multiple Hsp90 isoforms. Such promiscuity remains a significant challenge, even among compounds reported to be selective for the ER-resident isoform Grp94.[104] Particularly with respect to ER proteostasis, development of selective Grp94 inhibitors is essential to reduce undesirable cytotoxicity and ensure that only the intended proteostasis network node is targeted.

Fortunately, the ATP binding pocket of Grp94 is the most distinctive amongst the Hsp90 chaperone isoforms,[105] and recent progress in developing chemical tools for targeting Grp94 is encouraging. Rigorous structure-based analyses revealed that certain members of the purine-scaffold series bind a Grp94-specific hydrophobic pocket, permitting functional characterization of the effects of Grp94 inhibition in cancer cells.[106] Other compound classes also show substantial promise, including radamide derivatives like BnIm, the first Grp94-selective inhibitor discovered.[107] Improvements on BnIm have yielded even more potent Grp94 inhibitors.[108] As such chemical tools have become available, various groups have begun to test the consequences of Grp94 inhibition in protein misfolding-related disease model systems. For example, BnIm-mediated Grp94 inhibition ameliorates excessive quality control of the Ala322Asp variant of GABA$_A$ receptors, whose insufficient trafficking is associated with specific forms of epilepsy.[109] Such results are reminiscent of the ability of UPR activation to reduce excessive quality control of GCase.[85] Here, however, targeting an ER proteostasis node downstream of the UPR offers a more specific biological response.

Selective inhibition of the ER's Hsp70 analogue BiP has proven more challenging. Eukaryotes express as many as thirteen Hsp70 isoforms,[110] including BiP in the ER, mtHsp70 in mitochondria, the constitutive cytosolic Hsc70, and the inducible cytosolic isoform Hsp72. Moreover, BiP association with IRE1, PERK, and ATF6 regulates UPR signaling, making it difficult to target the chaperone without also inducing the UPR.[111, 112] While a few small molecules have been introduced as BiP inhibitors, it is not clear that they act selectively. Nonetheless, because BiP facilitates proper folding of nascent polypeptide chains, targets misfolded proteins for ERAD, and regulates ER stress signaling, selective modulation of BiP activities remains a promising avenue for resolving ER proteostasis defects. The identification of JG-98 as an inhibitor of the Bag3–Hsp72 interaction suggests that targeting protein–protein
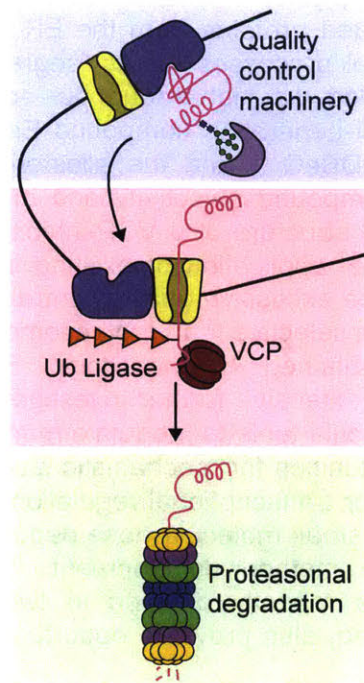
Figure 4.5: *Advances in understanding and adapting ERAD.* Current methods primarily modulate ERAD by influencing transport and/or degradation of misfolded substrates (indicated by red and green boxes, respectively). Other promising targets for therapeutic intervention include the E3 ligase Hrd1 (blue box), which interacts with a broad range of protein substrates via its co-chaperones, and is involved in the selection of terminally misfolded ER clients for ERAD.

interactions of co-chaperones with BiP could be a viable alternative strategy for modulating BiP activity in the absence of small molecules that directly and selectively bind to BiP.[113, 114] The discovery that reversible AMPylation regulates BiP function may likewise yield new opportunities to target BiP activity with selective small molecules.[115]

Targeting ERAD. Quality control mediated by ERAD is another proteostasis node downstream of the UPR where selective modulation offers substantial promise in disease. The process requires coordination of many different proteins in the ER lumen, lipid bilayer, and the cytosol in order to recognize, capture, transport, and degrade aberrant ER client proteins that pose a threat to cell health.[5] ERAD can be viewed as a three-step process, involving (1) selection/recognition of misfolded ER clients, (2) retrotranslocation to the cytosol, and (3) proteasomal degradation (*Figure 4.5*). The detailed mechanism involves intricate networks of chaperone–chaperone and chaperone–client interactions, as well as transport machines that we are only beginning to understand in biochemical detail.[116-118]

The selection/recognition step of ERAD has thus far proven challenging to target selectively, whereas substantial progress has been made in modulating retrotranslocation and degradation. For example, diverse and potent small molecules are available for proteasome inhibition.[119] Targeting the proteasome has proven therapeutically valuable in multiple myeloma, where treatment with bortezomib or other inhibitors appears to enhance proteotoxic load in the ER, triggering stress-induced apoptosis.[120] However, proteasome inhibition can also be severely toxic for normal cells, and it disrupts protein recycling not just in the ER but also in other subcellular compartments.

The need for more selective modulators of ERAD prompted the search for inhibitors of

93

ER retrotranslocation, and specifically for inhibitors of VCP/p97. VCP/p97 is an ATPase that assists in the extraction of misfolded proteins from the ER for proteasomal degradation, and also participates in other biological processes via its "segregase" activity.[121, 122] A number of VCP inhibitors have been identified via high-throughput screening and optimization efforts, ranging from the promiscuous first-generation compound Eeyarestatin I,[123, 124] to more potent second-generation inhibitors like DBeQ[125] and the allosteric inhibitor NMS-873.[126] Rigorous biophysical characterization of compound selectivity and mechanism of action, including the recent determination of a cryo-EM structure of a VCP-inhibitor complex, will continue to guide inhibitor optimization.[127] Meanwhile, application of existing inhibitors has begun to reveal the potential of VCP inhibition to reduce excessive quality control of misfolding ER client proteins in a manner that is considerably more selective than proteasome inhibition.[128, 129]

The promising results obtained by perturbing ER proteostasis network nodes downstream of the UPR motivate extensive further investigation in this area. As highlighted in *Figure 4.4*, we still lack small molecule tools to modulate many of these nodes. The emergence of such tools will open new opportunities for mechanistic studies and disease therapies. In the meantime, Cas9-based methods for transcriptional regulation of individual proteostasis network components in a highly selective, small molecule dose-dependent manner[130] should assist the prioritization of nodes for future method development. The continued advance of high-throughput genetic screening, as recently outlined in two reports that combine CRISPR technology with single cell RNA-seq, also provides opportunities for identifying new targets.[131, 132]

Concept Summary. Advances in targeting Hsp90 and ERAD exemplify how much fruitful ground lies beyond IRE1, ATF6, and PERK. Proteomic and high-throughput genetic screening approaches constitute a critical first step in identifying relevant nodes for specific diseases. Where key downstream nodes are known, the next task is to develop tools and modulators of sufficient potency for mechanistic work and drug development. As is the case for Hsp90 and other abundant protein families, researchers must not only identify what isoform is important for a particular process, but also whether and how the relevant isoform can be targeted. Due to high sequence and structural similarities between isoforms, this second question will likely continue to be a significant hurdle to modulating nodes downstream of the UPR. However, recent progress in Grp94 and VCP inhibition offers hope that combined biophysical and biochemical approaches will yield similar advances for isoform-specific targeting of other ER proteostasis network components.

Conclusions

As the master regulator of proteostasis in the secretory pathway, the UPR plays critical roles in both health and disease. Various groups have demonstrated UPR participation in organ development and innate immunity; more recently, others have revealed alternate cellular mechanisms for resolving stress, whether in the form of newly-discovered extracellular chaperones, membrane expansion, or changes to the oxidative folding environment. The advent of diverse chemical and chemical genetic methods now allows researchers to more precisely control UPR arm activation and inhibition, opening new doors to study not only the fundamental biology of the UPR and the secretory pathway, but also the potential of UPR modulation for therapeutic intervention. When selecting from existing tools to modulate or perturb the UPR, researchers must balance specificity with how much engineering or deviation from the natural system is required. Moreover, new connections between *N*-glycosylation and XBP1s signaling (Chapter V) or protein folding (Chapter VI) serve as a reminder that proteostasis encompasses not only proteins, but also post-translational modifications. The development of improved compounds to target both the UPR and downstream secretory proteostasis network nodes remains an urgent need and, when realized, will surely yield significant new findings. As the field

continues to progress, we anticipate not only the elucidation of new functions of the UPR and improved mechanistic understanding, but also the emergence of additional disease-modifying UPR- and ER-targeted therapeutic strategies.

References

[1] Gershenson, A., Gierasch, L. M., Pastore, A., and Radford, S. E. (2014) Energy landscapes of functional proteins are inherently risky, *Nat. Chem. Biol. 10*, 884-891.

[2] Hartl, F. U., Bracher, A., and Hayer-Hartl, M. (2011) Molecular chaperones in protein folding and proteostasis, *Nature 475*, 324-332.

[3] Aebi, M. (2013) N-linked protein glycosylation in the ER, *Biochem. Biophys. Acta, Mol. Cell Res. 1833*, 2430-2437.

[4] Tu, B. P., and Weissman, J. S. (2004) Oxidative protein folding in eukaryotes: mechanisms and consequences, *J. Cell Biol. 164*, 341-346.

[5] Smith, M. H., Ploegh, H. L., and Weissman, J. S. (2011) Road to ruin: targeting proteins for degradation in the endoplasmic reticulum, *Science 334*, 1086-1090.

[6] Walter, P., and Ron, D. (2011) The Unfolded Protein Response: From Stress Pathway to Homeostatic Regulation, *Science 334*, 1081-1086.

[7] Cox, J. S., Shamu, C. E., and Walter, P. (1993) Transcriptional induction of genes encoding endoplasmic reticulum resident proteins requires a transmembrane protein kinase, *Cell 73*, 1197-1206.

[8] Harding, H. P., Zhang, Y., and Ron, D. (1999) Protein translation and folding are coupled by an endoplasmic-reticulum-resident kinase, *Nature 397*, 271-274.

[9] Haze, K., Yoshida, H., Yanagi, H., Yura, T., and Mori, K. (1999) Mammalian transcription factor ATF6 is synthesized as a transmembrane protein and activated by proteolysis in response to endoplasmic reticulum stress, *Mol. Biol. Cell 10*, 3787-3799.

[10] Tirasophon, W., Welihinda, A. A., and Kaufman, R. J. (1998) A stress response pathway from the endoplasmic reticulum to the nucleus requires a novel bifunctional protein kinase/endoribonuclease (Ire1p) in mammalian cells, *Genes Dev. 12*, 1812-1824.

[11] Yoshida, H., Matsui, T., Yamamoto, A., Okada, T., and Mori, K. (2001) XBP1 mRNA is induced by ATF6 and spliced by IRE1 in response to ER stress to produce a highly active transcription factor, *Cell 107*, 881-891.

[12] Calfon, M., Zeng, H., Urano, F., Till, J. H., Hubbard, S. R., Harding, H. P., Clark, S. G., and Ron, D. (2002) IRE1 couples endoplasmic reticulum load to secretory capacity by processing the XBP-1 mRNA, *Nature 415*, 92-96.

[13] Harding, H. P., Novoa, I., Zhang, Y., Zeng, H., Wek, R., Schapira, M., and Ron, D. (2000) Regulated Translation Initiation Controls Stress-Induced Gene Expression in Mammalian Cells, *Mol. Cell 6*, 1099-1108.

[14] Vattem, K. M., and Wek, R. C. (2004) Reinitiation involving upstream ORFs regulates ATF4 mRNA translation in mammalian cells, *Proc. Natl. Acad. Sci. U.S.A. 101*, 11269-11274.

[15] Adachi, Y., Yamamoto, K., Okada, T., Yoshida, H., Harada, A., and Mori, K. (2008) ATF6 is a transcription factor specializing in the regulation of quality control proteins in the endoplasmic reticulum, *Cell Struct. Funct. 33*, 75-89.

[16] Shoulders, M. D., Ryno, L. M., Genereux, J. C., Moresco, J. J., Tu, P. G., Wu, C., Yates, J. R., 3rd, Su, A. I., Kelly, J. W., and Wiseman, R. L. (2013) Stress-independent activation of XBP1s and/or ATF6 reveals three functionally diverse ER proteostasis environments, *Cell Rep. 3*, 1279-1292.

[17] Hollien, J., Lin, J. H., Li, H., Stevens, N., Walter, P., and Weissman, J. S. (2009) Regulated Ire1-dependent decay of messenger RNAs in mammalian cells, *J. Cell Biol. 186*, 323-331.

[18] Moore, K., and Hollien, J. (2015) Ire1-mediated decay in mammalian cells relies on mRNA sequence, structure, and translational status, *Mol. Biol. Cell 26*, 2873-2884.

[19] Fu, S., Yang, L., Li, P., Hofmann, O., Dicker, L., Hide, W., Lin, X., Watkins, S. M., Ivanov, A. R., and Hotamisligil, G. S. (2011) Aberrant lipid metabolism disrupts calcium homeostasis causing liver endoplasmic reticulum stress in obesity, *Nature 473*, 528-531.

[20] Kono, N., Amin-Wetzel, N., and Ron, D. (2017) Generic membrane-spanning features endow IRE1alpha with responsiveness to membrane aberrancy, *Mol. Biol. Cell 28*, 2318-2332.

[21] Malhotra, J. D., and Kaufman, R. J. (2007) Endoplasmic reticulum stress and oxidative stress: a vicious cycle or a double-edged sword?, *Antioxid. Redox Signaling 9*, 2277-2293.

[22] Roy, C. R., Salcedo, S. P., and Gorvel, J. P. (2006) Pathogen-endoplasmic-reticulum interactions: in through the out door, *Nat. Rev. Immunol. 6*, 136-147.

[23] Volmer, R., and Ron, D. (2015) Lipid-dependent regulation of the unfolded protein response, *Curr. Opin. Cell Biol. 33*, 67-73.

[24] Coelho, Dina S., Cairrão, F., Zeng, X., Pires, E., Coelho, Ana V., Ron, D., Ryoo, Hyung D., and Domingos, Pedro M. (2013) Xbp1-Independent Ire1 Signaling Is Required for Photoreceptor Differentiation and Rhabdomere Morphogenesis in Drosophila, *Cell Rep. 5*, 791-801.

[25] Shen, X., Ellis, R. E., Lee, K., Liu, C. Y., Yang, K., Solomon, A., Yoshida, H., Morimoto, R., Kurnit, D. M., Mori, K., and Kaufman, R. J. (2001) Complementary signaling pathways regulate the unfolded protein response and are required for C. elegans development, *Cell 107*, 893-903.

[26] Dalton, Ryan P., Lyons, David B., and Lomvardas, S. (2013) Co-Opting the Unfolded Protein Response to Elicit Olfactory Receptor Feedback, *Cell 155*, 321-332.

[27] Laguesse, S., Creppe, C., Nedialkova, Danny D., Prévot, P.-P., Borgs, L., Huysseune, S., Franco, B., Duysens, G., Krusy, N., Lee, G., Thelen, N., Thiry, M., Close, P., Chariot, A., Malgrange, B., Leidel, Sebastian A., Godin, Juliette D., and Nguyen, L. (2015) A Dynamic Unfolded Protein Response Contributes to the Control of Cortical Neurogenesis, *Dev. Cell 35*, 553-567.

[28] Reimold, A. M., Iwakoshi, N. N., Manis, J., Vallabhajosyula, P., Szomolanyi-Tsuda, E., Gravallese, E. M., Friend, D., Grusby, M. J., Alt, F., and Glimcher, L. H. (2001) Plasma cell differentiation requires the transcription factor XBP-1, *Nature 412*, 300-307.

[29] Shaffer, A. L., Shapiro-Shelef, M., Iwakoshi, N. N., Lee, A.-H., Qian, S.-B., Zhao, H., Yu, X., Yang, L., Tan, B. K., Rosenwald, A., Hurt, E. M., Petroulakis, E., Sonenberg, N., Yewdell, J. W., Calame, K., Glimcher, L. H., and Staudt, L. M. (2004) XBP1, Downstream of Blimp-1, Expands the Secretory Apparatus and Other Organelles, and Increases Protein Synthesis in Plasma Cell Differentiation, *Immunity 21*, 81-93.

[30] Gass, J. N., Gifford, N. M., and Brewer, J. W. (2002) Activation of an Unfolded Protein Response during Differentiation of Antibody-secreting B Cells, *J. Biol. Chem. 277*, 49047-49054.

[31] Cho, J. A., Lee, A. H., Platzer, B., Cross, B. C., Gardner, B. M., De Luca, H., Luong, P., Harding, H. P., Glimcher, L. H., Walter, P., Fiebiger, E., Ron, D., Kagan, J. C., and Lencer, W. I. (2013) The unfolded protein response element IRE1alpha senses bacterial proteins invading the ER to activate RIG-I and innate immune signaling, *Cell Host Microbe 13*, 558-569.

[32] Martinon, F., Chen, X., Lee, A. H., and Glimcher, L. H. (2010) TLR activation of the transcription factor XBP1 regulates innate immune responses in macrophages, *Nat. Immunol. 11*, 411-418.

[33] Adolph, T. E., Tomczak, M. F., Niederreiter, L., Ko, H. J., Bock, J., Martinez-Naves, E., Glickman, J. N., Tschurtschenthaler, M., Hartwig, J., Hosomi, S., Flak, M. B., Cusick, J. L., Kohno, K., Iwawaki, T., Billmann-Born, S., Raine, T., Bharti, R., Lucius, R., Kweon, M. N., Marciniak, S. J., Choi, A., Hagen, S. J., Schreiber, S., Rosenstiel, P., Kaser, A., and Blumberg, R. S. (2013) Paneth cells as a site of origin for intestinal inflammation, *Nature 503*, 272-276.

[34] Richardson, C. E., Kooistra, T., and Kim, D. H. (2010) An essential role for XBP-1 in host protection against immune activation in C. elegans, *Nature 463*, 1092-1095.

[35] Hassan, I. H., Zhang, M. S., Powers, L. S., Shao, J. Q., Baltrusaitis, J., Rutkowski, D. T., Legge, K., and Monick, M. M. (2012) Influenza A viral replication is blocked by inhibition of the inositol-requiring enzyme 1 (IRE1) stress pathway, *J. Biol. Chem. 287*, 4679-4689.

[36] Zhang, P., Su, C., Jiang, Z., and Zheng, C. (2017) Herpes Simplex Virus 1 UL41 Protein Suppresses the IRE1/XBP1 Signal Pathway of the Unfolded Protein Response via Its RNase Activity, *J. Virol. 91*, e02056-16.

[37] Hempstead, A. D., and Isberg, R. R. (2015) Inhibition of host cell translation elongation by Legionella pneumophila blocks the host cell unfolded protein response, *Proc. Natl. Acad. Sci. U.S.A. 112*, E6790-6797.

[38] Treacy-Abarca, S., and Mukherjee, S. (2015) Legionella suppresses the host unfolded protein response via multiple mechanisms, *Nat. Commun. 6*, 7887.

[39] Bernales, S., McDonald, K. L., and Walter, P. (2006) Autophagy counterbalances endoplasmic reticulum expansion during the unfolded protein response, *PLoS Biol. 4*, e423.

[40] Fumagalli, F., Noack, J., Bergmann, T. J., Cebollero, E., Pisoni, G. B., Fasana, E., Fregno, I., Galli, C., Loi, M., Solda, T., D'Antuono, R., Raimondi, A., Jung, M., Melnyk, A., Schorr, S., Schreiber, A., Simonelli, L., Varani, L., Wilson-Zbinden, C., Zerbe, O., Hofmann, K., Peter, M., Quadroni, M., Zimmermann, R., and Molinari, M. (2016) Translocon component Sec62 acts in endoplasmic reticulum turnover during stress recovery, *Nat. Cell Biol. 18*, 1173-1184.

[41] Khaminets, A., Heinrich, T., Mari, M., Grumati, P., Huebner, A. K., Akutsu, M., Liebmann, L., Stolz, A., Nietzsche, S., Koch, N., Mauthe, M., Katona, I., Qualmann, B., Weis, J., Reggiori, F., Kurth, I., Hubner, C. A., and Dikic, I. (2015) Regulation of endoplasmic reticulum turnover by selective autophagy, *Nature 522*, 354-358.

[42] Margariti, A., Li, H., Chen, T., Martin, D., Vizcay-Barrena, G., Alam, S., Karamariti, E., Xiao, Q., Zampetaki, A., Zhang, Z., Wang, W., Jiang, Z., Gao, C., Ma, B., Chen, Y. G., Cockerill, G., Hu, Y., Xu, Q., and Zeng, L. (2013) XBP1 mRNA splicing triggers an autophagic response in endothelial cells through BECLIN-1 transcriptional activation, *J. Biol. Chem. 288*, 859-872.

[43] Schuck, S., Prinz, W. A., Thorn, K. S., Voss, C., and Walter, P. (2009) Membrane expansion alleviates endoplasmic reticulum stress independently of the unfolded protein response, *J. Cell Biol. 187*, 525-536.

[44] Taylor, R. C., and Dillin, A. (2013) XBP-1 is a cell-nonautonomous regulator of stress resistance and longevity, *Cell 153*, 1435-1447.

[45] Genereux, J. C., Qu, S., Zhou, M. H., Ryno, L. M., Wang, S. Y., Shoulders, M. D., Kaufman, R. J., Lasmezas, C. I., Kelly, J. W., and Wiseman, R. L. (2015) Unfolded protein response-induced ERdj3 secretion links ER stress to extracellular proteostasis, *EMBO J. 34*, 4-19.

[46] Denzel, M. S., Storm, N. J., Gutschmidt, A., Baddi, R., Hinze, Y., Jarosch, E., Sommer, T., Hoppe, T., and Antebi, A. (2014) Hexosamine pathway metabolites enhance protein quality control and prolong life, *Cell 156*, 1167-1178.

[47] Wang, Z. V., Deng, Y., Gao, N., Pedrozo, Z., Li, D. L., Morales, C. R., Criollo, A., Luo, X., Tan, W., Jiang, N., Lehrman, M. A., Rothermel, B. A., Lee, A. H., Lavandero, S., Mammen, P. P., Ferdous, A., Gillette, T. G., Scherer, P. E., and Hill, J. A. (2014) Spliced X-box binding protein 1 couples the unfolded protein response to hexosamine biosynthetic pathway, *Cell 156*, 1179-1192.

[48] Ruiz-Canada, C., Kelleher, D. J., and Gilmore, R. (2009) Cotranslational and posttranslational N-glycosylation of polypeptides by distinct mammalian OST isoforms, *Cell 136*, 272-283.

[49] Dewal, Mahender B., DiChiara, Andrew S., Antonopoulos, A., Taylor, Rebecca J., Harmon, Chyleigh J., Haslam, Stuart M., Dell, A., and Shoulders, Matthew D. (2015) XBP1s Links the Unfolded Protein Response to the Molecular Architecture of Mature N-Glycans, *Chem. Biol. 22*, 1301-1312.

[50] Chen, X., Iliopoulos, D., Zhang, Q., Tang, Q., Greenblatt, M. B., Hatziapostolou, M., Lim, E., Tam, W. L., Ni, M., Chen, Y., Mai, J., Shen, H., Hu, D. Z., Adoro, S., Hu, B., Song, M., Tan, C., Landis, M. D., Ferrari, M., Shin, S. J., Brown, M., Chang, J. C., Liu, X. S., and Glimcher, L. H. (2014) XBP1 promotes triple-negative breast cancer by controlling the HIF1alpha pathway, *Nature 508*, 103-107.

[51] Jamora, C., Dennert, G., and Lee, A. S. (1996) Inhibition of tumor progression by suppression of stress protein GRP78/BiP induction in fibrosarcoma B/C10ME, *Proc. Natl. Acad. Sci. U.S.A. 93*, 7690-7694.

[52] Mimura, N., Fulciniti, M., Gorgun, G., Tai, Y. T., Cirstea, D., Santo, L., Hu, Y., Fabre, C., Minami, J., Ohguchi, H., Kiziltepe, T., Ikeda, H., Kawano, Y., French, M., Blumenthal, M., Tam, V., Kertesz, N. L., Malyankar, U. M., Hokenson, M., Pham, T., Zeng, Q., Patterson, J. B., Richardson, P. G., Munshi, N. C., and Anderson, K. C. (2012) Blockade of XBP1 splicing by inhibition of IRE1alpha is a promising therapeutic option in multiple myeloma, *Blood 119*, 5772-5781.

[53] Rutkowski, D. T. (2006) Adaptation to ER Stress Is Mediated by Differential Stabilities of Pro-Survival and Pro-Apoptotic mRNAs and Proteins, *PLoS Biol. 4*, e374.

[54] Raina, K., Noblin, D. J., Serebrenik, Y. V., Adams, A., Zhao, C., and Crews, C. M. (2014) Targeted protein destabilization reveals an estrogen-mediated ER stress response, *Nat. Chem. Biol. 10*, 957-962.

[55] Plate, L., Cooley, C. B., Chen, J. J., Paxman, R. J., Gallagher, C. M., Madoux, F., Genereux, J. C., Dobbs, W., Garza, D., Spicer, T. P., Scampavia, L., Brown, S. J., Rosen, H., Powers, E. T., Walter, P., Hodder, P., Wiseman, R. L., and Kelly, J. W. (2016) Small molecule proteostasis regulators that reprogram the ER to reduce extracellular protein aggregation, *eLife 5*, e15550.

[56] Mendez, A. S., Alfaro, J., Morales-Soto, M. A., Dar, A. C., McCullagh, E., Gotthardt, K., Li, H., Acosta-Alvear, D., Sidrauski, C., Korennykh, A. V., Bernales, S., Shokat, K. M., and Walter, P. (2015) Endoplasmic reticulum stress-independent activation of unfolded protein response kinases by a small molecule ATP-mimic, *eLife 4*, e05434.

[57] Boyce, M., Bryant, K. F., Jousse, C., Long, K., Harding, H. P., Scheuner, D., Kaufman, R. J., Ma, D., Coen, D. M., Ron, D., and Yuan, J. (2005) A selective inhibitor of eIF2alpha dephosphorylation protects cells from ER stress, *Science 307*, 935-939.

[58] Tsaytler, P., Harding, H. P., Ron, D., and Bertolotti, A. (2011) Selective inhibition of a regulatory subunit of protein phosphatase 1 restores proteostasis, *Science 332*, 91-94.

[59] Sidrauski, C., McGeachy, A. M., Ingolia, N. T., and Walter, P. (2015) The small molecule ISRIB reverses the effects of eIF2alpha phosphorylation on translation and stress granule assembly, *eLife 4*, e05033.

[60] Das, I., Krzyzosiak, A., Schneider, K., Wrabetz, L., D'Antonio, M., Barry, N., Sigurdardottir, A., and Bertolotti, A. (2015) Preventing proteostasis diseases by selective inhibition of a phosphatase regulatory subunit, *Science 348*, 239-242.

[61] Crespillo-Casado, A., Chambers, J. E., Fischer, P. M., Marciniak, S. J., and Ron, D. (2017) PPP1R15A-mediated dephosphorylation of eIF2alpha is unaffected by Sephin1 or Guanabenz, *eLife 6*, e26109.

[62] Cross, B. C., Bond, P. J., Sadowski, P. G., Jha, B. K., Zak, J., Goodman, J. M., Silverman, R. H., Neubert, T. A., Baxendale, I. R., Ron, D., and Harding, H. P. (2012) The molecular basis for selective inhibition of unconventional mRNA splicing by an IRE1-binding small molecule, *Proc. Natl. Acad. Sci. U.S.A. 109*, E869-878.

[63] Volkmann, K., Lucas, J. L., Vuga, D., Wang, X., Brumm, D., Stiles, C., Kriebel, D., Der-Sarkissian, A., Krishnan, K., Schweitzer, C., Liu, Z., Malyankar, U. M., Chiovitti, D., Canny, M., Durocher, D., Sicheri, F., and Patterson, J. B. (2011) Potent and selective inhibitors of the inositol-requiring enzyme 1 endoribonuclease, *J. Biol. Chem. 286*, 12743-12755.

[64] Wang, L., Perera, B. G., Hari, S. B., Bhhatarai, B., Backes, B. J., Seeliger, M. A., Schurer, S. C., Oakes, S. A., Papa, F. R., and Maly, D. J. (2012) Divergent allosteric control of the IRE1alpha endoribonuclease using kinase inhibitors, *Nat. Chem. Biol. 8*, 982-989.

[65] Gallagher, C. M., and Walter, P. (2016) Ceapins inhibit ATF6alpha signaling by selectively preventing transport of ATF6alpha to the Golgi apparatus during ER stress, *eLife 5*, e11880.

[66] Lin, J. H., Li, H., Zhang, Y. H., Ron, D., and Walter, P. (2009) Divergent Effects of PERK and IRE1 Signaling on Cell Viability, *PLoS One 4*, e4170.

[67] Lu, P. D., Jousse, C., Marciniak, S. J., Zhang, Y., Novoa, I., Scheuner, D., Kaufman, R. J., Ron, D., and Harding, H. P. (2004) Cytoprotection by pre-emptive conditional phosphorylation of translation initiation factor 2, *EMBO J. 23*, 169-179.

[68] Back, S. H., Lee, K., Vink, E., and Kaufman, R. J. (2006) Cytoplasmic IRE1 alpha-mediated XBP1 mRNA splicing in the absence of nuclear processing and endoplasmic reticulum stress, *J. Biol. Chem. 281*, 18691-18706.

[69] Dan, H., Upton, J. P., Hagen, A., Callahan, J., Oakes, S. A., and Papa, F. R. (2008) A kinase inhibitor activates the IRE1 alpha RNase to confer cytoprotection against ER stress, *Biochem. Biophys. Res. Commun. 365*, 777-783.

[70] Han, D., Lerner, A. G., Vande Walle, L., Upton, J. P., Xu, W., Hagen, A., Backes, B. J., Oakes, S. A., and Papa, F. R. (2009) IRE1alpha kinase activation modes control alternate endoribonuclease outputs to determine divergent cell fates, *Cell 138*, 562-575.

[71] Papa, F. R., Zhang, C., Shokat, K., and Walter, P. (2003) Bypassing a kinase activity with an ATP-competitive drug, *Science 302*, 1533-1537.

[72] Ghosh, R., Wang, L., Wang, E. S., Perera, B. G., Igbaria, A., Morita, S., Prado, K., Thamsen, M., Caswell, D., Macias, H., Weiberth, K. F., Gliedt, M. J., Alavi, M. V., Hari, S. B., Mitra, A. K., Bhhatarai, B., Schurer, S. C., Snapp, E. L., Gould, D. B., German, M. S., Backes, B. J., Maly, D. J., Oakes, S. A., and Papa, F. R. (2014) Allosteric inhibition of the IRE1alpha RNase preserves cell viability and function during endoplasmic reticulum stress, *Cell 158*, 534-548.

[73] Okada, T., Yoshida, H., Akazawa, R., Negishi, M., and Mori, K. (2002) Distinct roles of activating transcription factor 6 (ATF6) and double-stranded RNA-activated protein kinase-like endoplasmic reticulum kinase (PERK) in transcription during the mammalian unfolded protein response, *Biochem. J. 366*, 585-594.

[74] Lee, A. H., Iwakoshi, N. N., and Glimcher, L. H. (2003) XBP-1 regulates a subset of endoplasmic reticulum resident chaperone genes in the unfolded protein response, *Mol. Cell. Biol. 23*, 7448-7459.

[75] Shoulders, M. D., Ryno, L. M., Cooley, C. B., Kelly, J. W., and Wiseman, R. L. (2013) Broadly applicable methodology for the rapid and dosable small molecule-mediated regulation of transcription factors in human cells, *J. Am. Chem. Soc. 135*, 8129-8132.

[76] Banaszynski, L. A., Chen, L. C., Maynard-Smith, L. A., Ooi, A. G., and Wandless, T. J. (2006) A rapid, reversible, and tunable method to regulate protein function in living cells using synthetic small molecules, *Cell 126*, 995-1004.

[77] Iwamoto, M., Bjorklund, T., Lundberg, C., Kirik, D., and Wandless, T. J. (2010) A general chemical method to regulate protein stability in the mammalian central nervous system, *Chem. Biol. 17*, 981-988.

[78] Moore, C. L., Dewal, M. B., Nekongo, E. E., Santiago, S., Lu, N. B., Levine, S. S., and Shoulders, M. D. (2016) Transportable, Chemical Genetic Methodology for the Small Molecule-Mediated Inhibition of Heat Shock Factor 1, *ACS Chem. Biol. 11*, 200-210.

[79] Johnson, S. M., Connelly, S., Fearns, C., Powers, E. T., and Kelly, J. W. (2012) The transthyretin amyloidoses: from delineating the molecular mechanism of aggregation linked to pathology to a regulatory-agency-approved drug, *J. Mol. Biol. 421*, 185-203.

[80] Hammarstrom, P., Schneider, F., and Kelly, J. W. (2001) Trans-suppression of misfolding in an amyloid disease, *Science 293*, 2459-2462.

[81] Hammarstrom, P., Wiseman, R. L., Powers, E. T., and Kelly, J. W. (2003) Prevention of transthyretin amyloid disease by changing protein misfolding energetics, *Science 299*, 713-716.

[82] Chen, J. J., Genereux, J. C., Qu, S., Hulleman, J. D., Shoulders, M. D., and Wiseman, R. L. (2014) ATF6 activation reduces the secretion and extracellular aggregation of destabilized variants of an amyloidogenic protein, *Chem. Biol. 21*, 1564-1574.

[83] Cooley, C. B., Ryno, L. M., Plate, L., Morgan, G. J., Hulleman, J. D., Kelly, J. W., and Wiseman, R. L. (2014) Unfolded protein response activation reduces secretion and extracellular aggregation of amyloidogenic immunoglobulin light chain, *Proc. Natl. Acad. Sci. U.S.A. 111*, 13046-13051.

[84] Ron, I., and Horowitz, M. (2005) ER retention and degradation as the molecular basis underlying Gaucher disease heterogeneity, *Hum. Mol. Genet. 14*, 2387-2398.

[85] Mu, T. W., Ong, D. S., Wang, Y. J., Balch, W. E., Yates, J. R., 3rd, Segatori, L., and Kelly, J. W. (2008) Chemical and biological approaches synergize to ameliorate protein-folding diseases, *Cell 134*, 769-781.

[86] Ong, D. S., Mu, T. W., Palmer, A. E., and Kelly, J. W. (2010) Endoplasmic reticulum Ca2+ increases enhance mutant glucocerebrosidase proteostasis, *Nat. Chem. Biol. 6*, 424-432.

[87] Ong, D. S., Wang, Y. J., Tan, Y. L., Yates, J. R., 3rd, Mu, T. W., and Kelly, J. W. (2013) FKBP10 depletion enhances glucocerebrosidase proteostasis in Gaucher disease fibroblasts, *Chem. Biol. 20*, 403-415.

[88] Tan, Y. L., Genereux, J. C., Pankow, S., Aerts, J. M., Yates, J. R., 3rd, and Kelly, J. W. (2014) ERdj3 is an endoplasmic reticulum degradation factor for mutant glucocerebrosidase variants linked to Gaucher's disease, *Chem. Biol. 21*, 967-976.

[89] Ryoo, H. D., Domingos, P. M., Kang, M. J., and Steller, H. (2007) Unfolded protein response in a Drosophila model for retinal degeneration, *EMBO J. 26*, 242-252.

[90] Wu, J., Rutkowski, D. T., Dubois, M., Swathirajan, J., Saunders, T., Wang, J., Song, B., Yau, G. D. Y., and Kaufman, R. J. (2007) ATF6α Optimizes Long-Term Endoplasmic Reticulum Function to Protect Cells from Chronic Stress, *Dev. Cell 13*, 351-364.

[91] Lindholm, D., Wootz, H., and Korhonen, L. (2006) ER stress and neurodegenerative diseases, *Cell Death Differ. 13*, 385-392.

[92] Chiang, W. C., Hiramatsu, N., Messah, C., Kroeger, H., and Lin, J. H. (2012) Selective activation of ATF6 and PERK endoplasmic reticulum stress signaling pathways prevent mutant rhodopsin accumulation, *Invest. Ophthalmol. Vis. Sci. 53*, 7159-7166.

[93] Smith, S. E., Granell, S., Salcedo-Sicilia, L., Baldini, G., Egea, G., Teckman, J. H., and Baldini, G. (2011) Activating transcription factor 6 limits intracellular accumulation of mutant alpha(1)-antitrypsin Z and mitochondrial damage in hepatoma cells, *J. Biol. Chem. 286*, 41563-41577.

[94] Abeliovich, A., and Gitler, A. D. (2016) Defects in trafficking bridge Parkinson's disease pathology and genetics, *Nature 539*, 207-216.

[95] Forlino, A., and Marini, J. C. (2016) Osteogenesis imperfecta, *Lancet 387*, 1657-1671.

101

[96] DiChiara, A. S., Taylor, R. J., Wong, M. Y., Doan, N. D., Rosario, A. M., and Shoulders, M. D. (2016) Mapping and Exploring the Collagen-I Proteostasis Network, *ACS Chem. Biol. 11*, 1408-1421.

[97] Mirigian, L. S., Makareeva, E., Mertz, E. L., Omari, S., Roberts-Pilgrim, A. M., Oestreich, A. K., Phillips, C. L., and Leikin, S. (2016) Osteoblast Malfunction Caused by Cell Stress Response to Procollagen Misfolding in alpha2(I)-G610C Mouse Model of Osteogenesis Imperfecta, *J. Bone Miner. Res. 31*, 1608-1616.

[98] Fitzgerald, J., Lamandé, S. R., and Bateman, J. F. (1999) Proteasomal Degradation of Unassembled Mutant Type I Collagen Pro-α1(I) Chains, *J. Biol. Chem. 274*, 27392-27398.

[99] McCleese, J. K., Bear, M. D., Fossey, S. L., Mihalek, R. M., Foley, K. P., Ying, W., Barsoum, J., and London, C. A. (2009) The novel HSP90 inhibitor STA-1474 exhibits biologic activity against osteosarcoma cell lines, *Int. J. Cancer 125*, 2792-2801.

[100] Schulte, T. W., and Neckers, L. M. (1998) The benzoquinone ansamycin 17-allylamino-17-demethoxygeldanamycin binds to HSP90 and shares important biologic activities with geldanamycin, *Cancer Chemother. Pharmacol. 42*, 273-279.

[101] Shi, J., Van de Water, R., Hong, K., Lamer, R. B., Weichert, K. W., Sandoval, C. M., Kasibhatla, S. R., Boehm, M. F., Chao, J., Lundgren, K., Timple, N., Lough, R., Ibanez, G., Boykin, C., Burrows, F. J., Kehry, M. R., Yun, T. J., Harning, E. K., Ambrose, C., Thompson, J., Bixler, S. A., Dunah, A., Snodgrass-Belt, P., Arndt, J., Enyedy, I. J., Li, P., Hong, V. S., McKenzie, A., and Biamonte, M. A. (2012) EC144 is a potent inhibitor of the heat shock protein 90, *J. Med. Chem. 55*, 7786-7795.

[102] Whitesell, L., Mimnaugh, E. G., De Costa, B., Myers, C. E., and Neckers, L. M. (1994) Inhibition of heat shock protein HSP90-pp60v-src heteroprotein complex formation by benzoquinone ansamycins: essential role for stress proteins in oncogenic transformation, *Proc. Natl. Acad. Sci. U.S.A. 91*, 8324-8328.

[103] Chen, B., Piel, W. H., Gui, L., Bruford, E., and Monteiro, A. (2005) The HSP90 family of genes in the human genome: insights into their divergence and evolution, *Genomics 86*, 627-637.

[104] Liu, S., and Street, T. O. (2016) 5'-N-ethylcarboxamidoadenosine is not a paralog-specific Hsp90 inhibitor, *Protein Sci. 25*, 2209-2215.

[105] Soldano, K. L., Jivan, A., Nicchitta, C. V., and Gewirth, D. T. (2003) Structure of the N-terminal domain of GRP94. Basis for ligand specificity and regulation, *J. Biol. Chem. 278*, 48330-48338.

[106] Patel, P. D., Yan, P., Seidler, P. M., Patel, H. J., Sun, W., Yang, C., Que, N. S., Taldone, T., Finotti, P., Stephani, R. a., Gewirth, D. T., and Chiosis, G. (2013) Paralog-selective Hsp90 inhibitors define tumor-specific regulation of HER2, *Nat. Chem. Biol. 9*, 677-684.

[107] Duerfeldt, A. S., Peterson, L. B., Maynard, J. C., Ng, C. L., Eletto, D., Ostrovsky, O., Shinogle, H. E., Moore, D. S., Argon, Y., Nicchitta, C. V., and Blagg, B. S. (2012) Development of a Grp94 inhibitor, *J. Am. Chem. Soc. 134*, 9796-9804.

[108] Crowley, V. M., Khandelwal, A., Mishra, S., Stothert, A. R., Huard, D. J., Zhao, J., Muth, A., Duerfeldt, A. S., Kizziah, J. L., Lieberman, R. L., Dickey, C. A., and Blagg, B. S. (2016) Development of Glucose Regulated Protein 94-Selective Inhibitors Based on the BnIm and Radamide Scaffold, *J. Med. Chem. 59*, 3471-3488.

[109] Di, X. J., Wang, Y. J., Han, D. Y., Fu, Y. L., Duerfeldt, A. S., Blagg, B. S., and Mu, T. W. (2016) Grp94 Protein Delivers gamma-Aminobutyric Acid Type A (GABAA) Receptors to Hrd1 Protein-mediated Endoplasmic Reticulum-associated Degradation, *J. Biol. Chem. 291*, 9526-9539.

[110] Kampinga, H. H., Hageman, J., Vos, M. J., Kubota, H., Tanguay, R. M., Bruford, E. A., Cheetham, M. E., Chen, B., and Hightower, L. E. (2009) Guidelines for the nomenclature of the human heat shock proteins, *Cell Stress Chaperones 14*, 105-111.

102

[111] Carrara, M., Prischi, F., Nowak, P. R., Kopp, M. C., and Ali, M. M. (2015) Noncanonical binding of BiP ATPase domain to Ire1 and Perk is dissociated by unfolded protein CH1 to initiate ER stress signaling, *eLife 4*, e03522.

[112] Pincus, D., Chevalier, M. W., Aragon, T., van Anken, E., Vidal, S. E., El-Samad, H., and Walter, P. (2010) BiP binding to the ER-stress sensor Ire1 tunes the homeostatic behavior of the unfolded protein response, *PLoS Biol. 8*, e1000415.

[113] Li, X., Colvin, T., Rauch, J. N., Acosta-Alvear, D., Kampmann, M., Dunyak, B., Hann, B., Aftab, B. T., Murnane, M., Cho, M., Walter, P., Weissman, J. S., Sherman, M. Y., and Gestwicki, J. E. (2015) Validation of the Hsp70-Bag3 protein-protein interaction as a potential therapeutic target in cancer, *Mol. Cancer Ther. 14*, 642-648.

[114] Li, X., Srinivasan, S. R., Connarn, J., Ahmad, A., Young, Z. T., Kabza, A. M., Zuiderweg, E. R., Sun, D., and Gestwicki, J. E. (2013) Analogs of the Allosteric Heat Shock Protein 70 (Hsp70) Inhibitor, MKT-077, as Anti-Cancer Agents, *ACS Med. Chem. Lett. 4*, 1042-1047.

[115] Preissler, S., Rato, C., Perera, L., Saudek, V., and Ron, D. (2017) FICD acts bifunctionally to AMPylate and de-AMPylate the endoplasmic reticulum chaperone BiP, *Nat. Struct. Mol. Biol. 24*, 23-29.

[116] Christianson, J. C., Olzmann, J. A., Shaler, T. A., Sowa, M. E., Bennett, E. J., Richter, C. M., Tyler, R. E., Greenblatt, E. J., Harper, J. W., and Kopito, R. R. (2011) Defining human ERAD networks through an integrative mapping strategy, *Nat. Cell Biol. 14*, 93-105.

[117] Timms, R. T., Menzies, S. A., Tchasovnikarova, I. A., Christensen, L. C., Williamson, J. C., Antrobus, R., Dougan, G., Ellgaard, L., and Lehner, P. J. (2016) Genetic dissection of mammalian ERAD through comparative haploid and CRISPR forward genetic screens, *Nat. Commun. 7*, 11786.

[118] Stein, A., Ruggiano, A., Carvalho, P., and Rapoport, T. A. (2014) Key steps in ERAD of luminal ER proteins reconstituted with purified components, *Cell 158*, 1375-1388.

[119] Kisselev, A. F., van der Linden, W. A., and Overkleeft, H. S. (2012) Proteasome Inhibitors: An Expanding Army Attacking a Unique Target, *Chem. Biol. 19*, 99-115.

[120] Obeng, E. A., Carlson, L. M., Gutman, D. M., Harrington, W. J., Lee, K. P., and Boise, L. H. (2006) Proteasome inhibitors induce a terminal unfolded protein response in multiple myeloma cells, *Blood 107*, 4907-4916.

[121] Rabinovich, E., Kerem, A., Frohlich, K. U., Diamant, N., and Bar-Nun, S. (2002) AAA-ATPase p97/Cdc48p, a Cytosolic Chaperone Required for Endoplasmic Reticulum-Associated Protein Degradation, *Mol. Cell. Biol. 22*, 626-634.

[122] Meyer, H., Bug, M., and Bremer, S. (2012) Emerging functions of the VCP/p97 AAA-ATPase in the ubiquitin system, *Nat. Cell Biol. 14*, 117-123.

[123] Fiebiger, E., Hirsch, C., Vyas, J. M., Gordon, E., Ploegh, H. L., and Tortorella, D. (2004) Dissection of the dislocation pathway for type I membrane proteins with a new small molecule inhibitor, eeyarestatin, *Mol. Biol. Cell 15*, 1635-1646.

[124] Wang, Q., Li, L., and Ye, Y. (2008) Inhibition of p97-dependent protein degradation by Eeyarestatin I, *J. Biol. Chem. 283*, 7445-7454.

[125] Chou, T. F., Brown, S. J., Minond, D., Nordin, B. E., Li, K., Jones, A. C., Chase, P., Porubsky, P. R., Stoltz, B. M., Schoenen, F. J., Patricelli, M. P., Hodder, P., Rosen, H., and Deshaies, R. J. (2011) Reversible inhibitor of p97, DBeQ, impairs both ubiquitin-dependent and autophagic protein clearance pathways, *Proc. Natl. Acad. Sci. U.S.A. 108*, 4834-4839.

[126] Magnaghi, P., D'Alessio, R., Valsasina, B., Avanzi, N., Rizzi, S., Asa, D., Gasparri, F., Cozzi, L., Cucchi, U., Orrenius, C., Polucci, P., Ballinari, D., Perrera, C., Leone, A., Cervi, G., Casale, E., Xiao, Y., Wong, C., Anderson, D. J., Galvani, A., Donati, D.,

O'Brien, T., Jackson, P. K., and Isacchi, A. (2013) Covalent and allosteric inhibitors of the ATPase VCP/p97 induce cancer cell death, *Nat. Chem. Biol. 9*, 548-556.

[127] Banerjee, S., Bartesaghi, A., Merk, A., Rao, P., Bulfer, S. L., Yan, Y., Green, N., Mroczkowski, B., Neitz, R. J., Wipf, P., Falconieri, V., Deshaies, R. J., Milne, J. L. S., Huryn, D., Arkin, M., and Subramaniam, S. (2016) 2.3 Å resolution cryo-EM structure of human p97 and mechanism of allosteric inhibition, *Science 351*, 871-875.

[128] Han, D. Y., Di, X. J., Fu, Y. L., and Mu, T. W. (2015) Combining valosin-containing protein (VCP) inhibition and suberanilohydroxamic acid (SAHA) treatment additively enhances the folding, trafficking, and function of epilepsy-associated gamma-aminobutyric acid, type A (GABAA) receptors, *J. Biol. Chem. 290*, 325-337.

[129] Wang, F., Song, W., Brancati, G., and Segatori, L. (2011) Inhibition of endoplasmic reticulum-associated degradation rescues native folding in loss of function protein misfolding diseases, *J. Biol. Chem. 286*, 43454-43464.

[130] Maji, B., Moore, C. L., Zetsche, B., Volz, S. E., Zhang, F., Shoulders, M. D., and Choudhary, A. (2017) Multidimensional chemical control of CRISPR-Cas9, *Nat. Chem. Biol. 13*, 9-11.

[131] Adamson, B., Norman, T. M., Jost, M., Cho, M. Y., Nunez, J. K., Chen, Y., Villalta, J. E., Gilbert, L. A., Horlbeck, M. A., Hein, M. Y., Pak, R. A., Gray, A. N., Gross, C. A., Dixit, A., Parnas, O., Regev, A., and Weissman, J. S. (2016) A Multiplexed Single-Cell CRISPR Screening Platform Enables Systematic Dissection of the Unfolded Protein Response, *Cell 167*, 1867-1882.

[132] Dixit, A., Parnas, O., Li, B., Chen, J., Fulco, C. P., Jerby-Arnon, L., Marjanovic, N. D., Dionne, D., Burks, T., Raychowdhury, R., Adamson, B., Norman, T. M., Lander, E. S., Weissman, J. S., Friedman, N., and Regev, A. (2016) Perturb-Seq: Dissecting Molecular Circuits with Scalable Single-Cell RNA Profiling of Pooled Genetic Screens, *Cell 167*, 1853-1866.

CHAPTER V
XBP1s Activation Can Globally Remodel N-Glycan Structure Distribution Patterns

Summary

Classically, the unfolded protein response (UPR) safeguards secretory pathway proteostasis. The most ancient arm of the UPR, the IRE1-activated XBP1s-mediated response, has roles in secretory pathway maturation beyond resolving proteostatic stress. Understanding the consequences of XBP1s activation for cellular processes is critical for elucidating mechanistic connections between XBP1s and development, immunity, and disease. Here, we show that a key functional output of XBP1s activation is a cell type-dependent shift in the distribution of N-glycan structures on endogenous membrane and secreted proteomes. For example, XBP1s activity decreased levels of sialylation and bisecting GlcNAc in the HEK293 membrane proteome and secretome, while substantially increasing the population of oligo-mannose N-glycans only in the secretome. In HeLa cell membranes, stress-independent XBP1s activation increased the population of high-mannose and tetra-antennary N-glycans, and also enhanced core fucosylation. mRNA profiling experiments suggest that XBP1s-mediated remodeling of the N-glycome is, at least in part, a consequence of coordinated transcriptional resculpting of N-glycan maturation pathways by XBP1s. The discovery of XBP1s-induced N-glycan structural remodeling on a glycome-wide scale suggests that XBP1s can act as a master regulator of N-glycan maturation. Moreover, because the sugars on cell surface proteins or on proteins secreted from an XBP1s-activated cell can be molecularly distinct from those of an unactivated cell, these findings reveal a potential new mechanism for translating intracellular stress signaling into altered interactions with the extracellular environment.

Author List

Madeline Y. Wong,* Kenny Chen,* Aristotelis Antonopoulos,* Brian T. Kasper,* Mahender B. Dewal, Rebecca J. Taylor, Charles A. Whittaker, Pyae P. Hein, Anne Dell, Joseph C. Genereux, Stuart M. Haslam,[1] Lara K. Mahal,[1] and Matthew D. Shoulders[1]

*These authors contributed equally
[1]Co-corresponding authors

Contributions

Introduction

The unfolded protein response (UPR) is classically responsible for maintaining proteostasis in the secretory pathway.[1] In the metazoan UPR, three transmembrane proteins (IRE1, PERK, and ATF6) act to detect protein misfolding stress in the endoplasmic reticulum (ER). Once stress is detected, activation of each sensor results in the production of a distinctive transcription factor – XBP1s, ATF4, or ATF6(1–373), respectively.[2-5] The targets of these transcription factors include a suite of ER chaperones and quality control factors that can resolve proteostatic challenges caused by physiologic and environmental changes.[6-8] The central role of the UPR in maintaining secretory proteostasis has catalyzed extensive efforts to delineate relevant mechanisms of induction,[9-11] define cellular effects of UPR activation,[12, 13] and create chemical and genetic tools to modulate the IRE1-XBP1s, PERK-ATF4, and ATF6 branches of the UPR in a selective, stress-independent manner.[14-16]

IRE1-XBP1s is the most ancient UPR arm and is the sole pathway conserved in yeast.[17] It is unsurprising, then, that the IRE1-XBP1s arm has overarching roles in secretory pathway maturation and maintenance beyond just the regulation of protein folding and quality control.[18, 19] Indeed, recent work continues to define essential functions for the IRE1-XBP1s pathway in general, and the XBP1s transcription factor, in particular, that are distinctive from functions of the other UPR arms. For example, XBP1s is critical for development, immune response, memory formation, and cell non-autonomous ER stress signaling.[20-23] Moreover, chronic upregulation of the XBP1s transcription factor is a pathologic hallmark of numerous malignancies.[24, 25] Consistent with its diverse regulatory functions, the IRE1-XBP1s arm can be activated independently of PERK-ATF4 and ATF6 in response to assorted stimuli.[6, 11, 26] This diversity of input signals and output functions emphasizes the broad roles of the IRE1-XBP1s axis beyond sensing and responding to protein misfolding. Attaining mechanistic understanding of these diverse processes requires first comprehensively defining the molecular consequences of XBP1s induction.

The majority of proteins traversing the secretory pathway are co- or post-translationally N-glycosylated.[27] A 14-residue oligosaccharide with the structure $Glc_3Man_9GlcNAc_2$ (Glc = glucose; Man = mannose; GlcNAc = N-acetylglucosamine) is installed on Asn residues within specific amino acid sequons.[28] XBP1s may modulate the extent N-linked glycans are produced by the cell,[29] thereby potentially assisting N-glycoprotein trafficking and directly enhancing proteostasis by providing improved access to the ER's lectin-based chaperone and quality control machineries.[29-31]

Downstream of N-glycan installation, the immature sugar is processed in both the ER and the Golgi to yield the vast array of high- and oligo-mannose, hybrid, and complex N-glycan structures presented on secreted and cell-surface proteins.[32] The molecular architecture of this N-glycome is dynamic, with consequences for both normal and pathologic processes, including cell motility and adhesion, cell–cell interactions, and immune system function.[33-36] However, our understanding of how intracellular signaling pathways coordinately define N-glycan maturation and thereby regulate these processes remains incomplete.

We previously reported that stress-independent activation of XBP1s can significantly modify the structural distribution of N-glycans installed at the single N-glycosylation site of a soluble, secreted glycoprotein domain derived from the CD2 adhesion protein.[37] The shift from a high-mannose to a hybrid N-glycan observed in that work was specifically caused by XBP1s-mediated alterations in N-glycan maturation, not by modified CD2 protein synthesis or trafficking. Those findings hint at the possibility that the IRE1-XBP1s arm of the UPR regulates the molecular architecture of the endogenous N-glycome in a coordinated manner.[38] By this hypothesis, rather than merely regulating occupancy at a given N-glycosylation sequon, XBP1s also influences the ultimate structure of the mature N-glycans displayed on cell-surface or secreted proteins. However, the CD2 protein domain used in our prior studies was an ectopically expressed, soluble, and glycosylation-naïve protein variant,[37, 39] rendering the

biological relevance of the observed XBP1s-dependent remodeling of *N*-glycan maturation uncertain.

Here, we ask (1) whether induction of the transcription factor XBP1s results in altered *N*-glycan structures on endogenous proteins and (2) whether such remodeling is sufficiently significant to be observed in a glycome-wide experiment, not just on individual, purified glycoproteins. We find that stress-independent XBP1s activation substantially alters the composition of the *N*-glycome in a cell type- and proteome-dependent manner. The resulting changes in the distribution of *N*-glycan structures indicate that XBP1s-activated cell membranes, and proteins secreted from XBP1s-activated cells, are often distinguishable at the molecular level from those of cells with basal levels of UPR signaling. These observations suggest a new mechanism to translate intracellular stress signaling to the extracellular milieu, a phenomenon with important implications for our emerging understanding of how the UPR shapes higher order biological activities (e.g., organismal development and immune response) beyond just the maintenance of ER proteostasis.

Results

Experimental Platform and Workflow to Scrutinize Effects of XBP1s on the N-Glycome. In prior work, we used enzymatic digestions, MALDI-TOF mass spectrometry (MS), and TOF/TOF MS/MS analyses to show that the structure of the single *N*-glycan on the secreted CD2 adhesion domain changes upon ER stress-independent XBP1s activation.[37] The biological relevance of those studies was limited by the focus on an ectopically overexpressed, non-native *N*-glycoprotein. Moreover, while a protein-specific *N*-glycan maturation effect is potentially interesting, functionally relevant *N*-glycan remodeling would likely involve larger-scale XBP1s-mediated *N*-glycome remodeling. Our objective here, therefore, was to test the hypotheses that XBP1s-mediated remodeling of *N*-glycan structures (1) occurs on endogenous glycoproteins and (2) is sufficiently significant to be observed in experiments surveying the glycome.

Application of a traditional ER stress inducer such as thapsigargin (Tg) does cause XBP1s activation, and is therefore one approach to test these hypotheses. However, the pleiotropic consequences of ER stressors, including extensive ER protein misfolding and high cytotoxicity, limit the reliability and interpretation of the resulting data. Selective induction of XBP1s using chemical genetics is an alternative strategy that we and others have employed to elucidate the specific consequences of XBP1s activity.[14, 40-43] This strategy results in chronic XBP1s induction in the absence of global UPR activation by ER stress. Such selective and chronic XBP1s activation is observed in numerous relevant settings, including development, the immune response, and cancer.[20, 21, 25]

In this work, we employed two cell lines engineered for ER stress-independent induction of XBP1s transcriptional activity at physiologically relevant levels. In both cell lines, XBP1s expression was placed under control of the tetracycline repressor, providing doxycycline (dox)-regulated expression of XBP1s. The first stable, single-colony cell line we used, termed HEK[XBP1s], was derived from the human embryonic kidney cell line and was reported previously.[14] The second cell line, developed for this study, was a stable, single-colony HeLa cell line (derived from the cervical epithelial line) termed HeLa[XBP1s]. In both HEK[XBP1s] (Figure 5.1A) and HeLa[XBP1s] (Figure 5.1B) cells, treatment with dox upregulated mRNA expression for the established XBP1s target genes *ERDJ4* and *SEC24D* to levels similar to that induced by treatment with the classic ER stressor Tg. In both cases, XBP1s activation occurred independently of global UPR induction, as well-known transcripts induced by the PERK-ATF4 arm of the UPR were not upregulated (e.g., *CHOP* and *GADD34*; Figures 5.1A and 5.1B).

In previous work, we characterized the consequences of sustained XBP1s activation on the transcriptome and the proteome in HEK[XBP1s] cells.[14] As the HeLa[XBP1s] cell line was newly developed, we next assessed the transcript-level changes induced by chronic, selective XBP1s activation in this cell line. We sequenced mRNA from vehicle-, dox-, or Tg-treated HeLa[XBP1s]
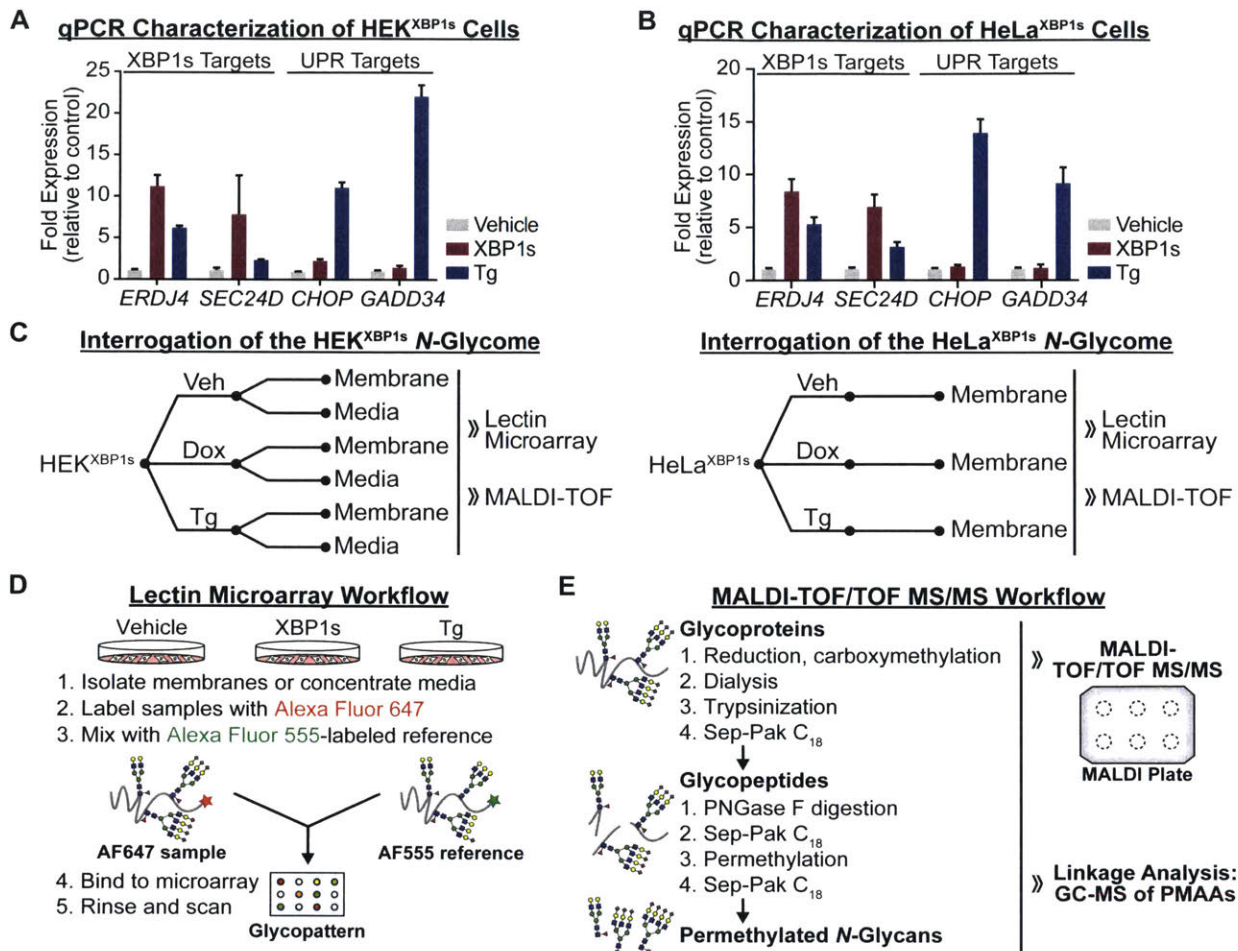
Figure 5.1: *Characterization of HEK^{XBP1s} and HeLa^{XBP1s} cell lines and experimental workflows for glycomic analyses.* Transcript levels of genes regulated by the IRE1-XBP1s arm (*ERDJ4* and *SEC24D*) or the PERK-ATF4 arm (*CHOP* and *GADD34*) of the UPR were measured in HEK^{XBP1s} cells (A) and HeLa^{XBP1s} cells (B) by qRT-PCR. (C) Experimental workflows for isolating membrane proteomes (HEK^{XBP1s} and HeLa^{XBP1s}) or secretomes (HEK^{XBP1s}) in control, XBP1s-activated, or Tg-treated cells. (D) Workflow for glycan analysis by lectin microarrays. (E) Workflow for MS analysis of *N*-glycomes.

cells in biological triplicate. Hierarchical clustering of differentially expressed coding genes showed that XBP1s activation preferentially clustered with vehicle-treated rather than Tg-treated samples, indicating that the transcriptional consequences of chronic XBP1s activation were significantly muted relative to treatment with an ER stressor (Figure S1A). As expected (Figure S1B), and in line with our previous observations for HEK^{XBP1s} cells,[14] gene set enrichment analysis revealed specific gene sets enriched in the XBP1s-activated lines alone (e.g., the unfolded protein binding gene set), others that were enriched in both XBP1s-activated and Tg-treated cells (e.g., the response to ER stress gene set), and a third group of gene sets that were enriched only in the Tg-treated cells (e.g., the intrinsic apoptotic signaling in response to ER stress gene set; see Tables S1A–D for full results). HEK^{XBP1s} and HeLa^{XBP1s} cells thus provide two distinct cell types in which we can induce sustained XBP1s activity over the time period required for an intracellular signal like XBP1s-mediated transcription to propagate to the mature *N*-glycoproteome, mimicking biological cases of extended XBP1s activation.[25, 44]

With appropriate cell lines in hand, we devised an experimental workflow to test for effects of XBP1s activation on *N*-glycan maturation. As illustrated in Figure 5.1C (left panel), we treated HEK[XBP1s] cells in complete media with vehicle (control), dox (to activate XBP1s), or Tg (to induce ER stress and the global UPR), followed by isolation of the membrane fraction by ultracentrifugation. Alternatively, we treated HEK[XBP1s] cells identically in a serum-free media formulation (Freestyle[TM] media) such that we could harvest and concentrate the secreted proteome for downstream glycomic analyses while minimizing contamination from serum glycoproteins. We used Freestyle[TM] media instead of serum-free Dulbecco's Modified Eagle's medium (DMEM) because we have observed that serum-free DMEM causes significant cell stress over longer treatments such as the 24 hours used here. Grp78 was detected in the Freestyle[TM] media only upon XBP1s activation (Figure S2), suggesting that the presence of any ER proteins in the XBP1s-activated secretome is a consequence of UPR activity, as has previously been observed by others.[45] For HeLa[XBP1s] cells, we carried out an identical workflow (Figure 5.1C, right panel) but isolated only the membrane fraction, as in our observation HeLa cells are less secretory than HEK cells.

After isolating membrane and secreted protein fractions, we next carried out two types of glycomic experiments on the samples: 1) Lectin microarray analysis, which enables us to directly compare relative levels of glycan motifs/substructures,[46, 47] and 2) Mass spectrometry, which provides a more detailed analysis of glycan structures. We used dual-color lectin microarrays (Figure 5.1D) on biological triplicates to gain insight into potential glycome remodeling due to XBP1s activation and/or Tg treatment. We measured the relative binding of fluorescently labeled glycoproteome samples and an orthogonally-labeled common reference to immobilized lectins on our microarray. This technology, which leverages the known glycan-binding preferences of lectins to gain insight into changes in the levels of glycan substructures (e.g., α2,6-sialylation), provides a systems-level perspective on differences in the *N*- and *O*-linked glycan epitopes present in a sample.[46, 47] The lectin microarrays we employed were comprised of ~90 unique lectins with diverse sugar-binding specificities (a complete list of the lectins included in the array for HEK[XBP1s] and HeLa[XBP1s] samples is provided in Tables S2A and S2B, respectively). Samples were labeled with Alexa Fluor 647 or 555, mixed with an equal amount of an orthogonally labeled common reference sample (a mix of all replicates for all conditions) and incubated with the microarray. The resulting data (normalized to the reference) provides relative quantitation of the glycan epitopes/substructures recognized by a particular lectin across sample conditions.

We used MALDI-TOF MS and TOF/TOF MS/MS (Figure 5.1E), to obtain a molecular-level view of *N*-glycome composition in all our samples. We performed GC-MS analyses in parallel, which provides further detail by allowing for identification of specific sugar linkages. All MS analyses were performed in a minimum of biological duplicate. The MS data validates and enhances results derived from the lectin microarrays, and adds more detailed structural information. The standard MS techniques used herein pinpoint shifts in the overall composition of the glycome within a sample (i.e., what % of overall glycans in a given sample have sialic acid), but do not provide direct comparison of a specific glycan's absolute abundance between 2 samples (i.e., how much sialic acid is expressed in sample 1 vs. sample 2). Thus, % changes from MS represent shifts in the profile whereas % changes from the lectin microarray represent changes in the overall levels of an epitope. By combining the ability of lectin microarrays to examine directly the relative levels of a glycan motif with the more detailed structural information gathered from MS, we were able to more thoroughly investigate glycosylation changes, with microarray data providing collective expression changes for glycan motifs regardless of parent glycan backbone, and MS supplying more detailed structural analysis of the individual glycans expressed. Together these two techniques enabled us to evaluate, at high resolution, whether and to what extent XBP1s modulates glycan structure. Results from these complementary glycomic analyses are presented in the sections below.

Figure 5.2

A

<u>Lectin Microarray Analysis of HEK^{XBP1s} Membranes</u>
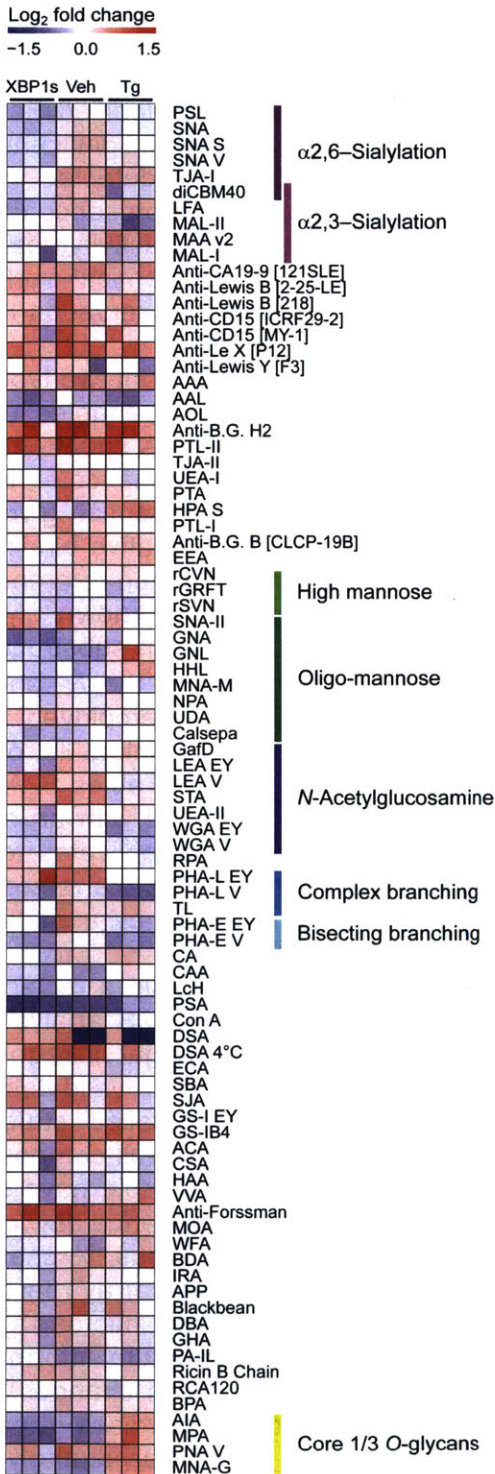
B

<u>MS Analysis of HEK^{XBP1s} Membranes</u>

C

XBP1s-Induced Changes to *N*-Glycans in HEK^{XBP1s} Membranes

Decrease in sialylation (by lectin microarray & MS)			Decrease in bisecting GlcNAc (by lectin microarray & MS)
Bi-	*Tri-*	*Tetra-Antennary*	

110
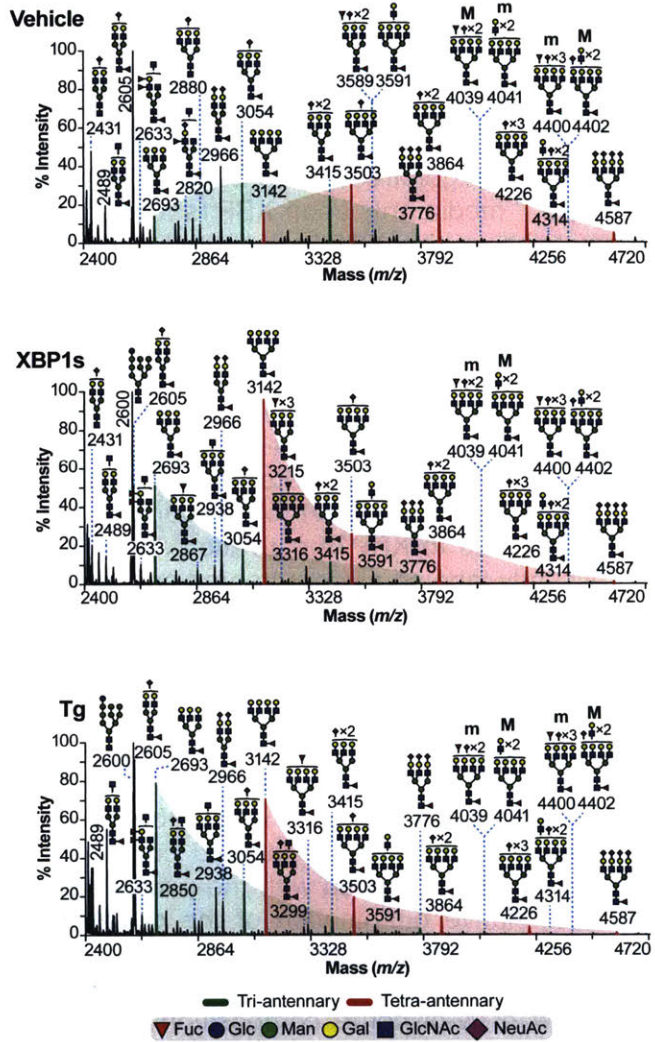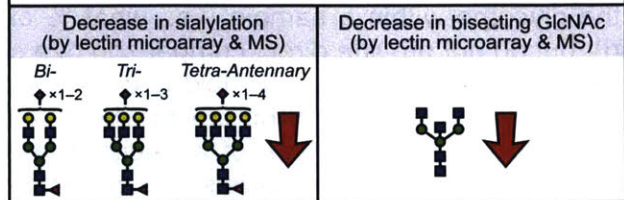
Figure 5.2: *Analysis of the HEK^XBP1s membrane glycoproteome.* (A) Heat map of lectin microarray data generated from isolated HEK^XBP1s membrane glycoproteomes under vehicle-treated, XBP1s-activated, or Tg-treated conditions. Full lectin names, print concentration, and sources are listed in Table S2A. Color intensity represents normalized log2 ratio data relative to a pooled sample reference. Each column represents one biological replicate of the indicated sample. Select lectin groups with shared binding specificities are annotated at right of heatmap. See also Figure S3A and Table S3A in the Appendix. (B) Partial MALDI-TOF mass spectra from the HEK^XBP1s membrane proteome revealed decreased sialylation on bi- (see Figure S3C), tri-, and tetra-antennary *N*-glycans upon XBP1s activation. Green and red peaks correspond to tri- and tetra-antennary *N*-glycans, respectively, with various levels of sialylation. Manually inserted green and red shaded areas highlight the distribution shift of sialylation towards fewer NeuAcs on the corresponding tri- and tetra-antennary *N*-glycans. Structures outside a bracket were not unequivocally defined. "M" and "m" designations indicate major and minor abundances, respectively. Putative structures are based on composition, tandem MS, and knowledge of biosynthetic pathways. All molecular ions are [M+Na]$^+$. High mass spectra are presented in Figure S3B, with GC-MS data in Table S4 (see Appendix). Experiments were carried out on two biological replicates. (C) Table highlighting key XBP1s-induced changes to the HEK^XBP1s membrane *N*-glycome and the analytical methods by which these changes were observed (see also Tables S5 and S6).

XBP1s Remodels the HEK[XBP1s] Cell Membrane N-Glycome. We began by assaying the *N*-glycome of HEK[XBP1s] cell membranes upon XBP1s activation or Tg treatment, following the workflow in Figure 5.1C. Cells were treated with dox to activate XBP1s for 72 h or Tg to induce the UPR for 24 h. Membrane fractions were harvested and processed for glycomic analyses as described above.

Lectin microarray analysis of selective XBP1s activation in HEK[XBP1s] membranes (Table S3A) revealed a significant decrease in $\alpha2,6$-sialylation (Figures 5.2A and S3A left panel). The $\alpha2,6$-sialic acid-specific lectins SNA-I, TJA-I, and PSL-I showed an average 29% decrease in binding relative to the vehicle-treated control. Our engineered recombinant lectin diCBM40, which binds both $\alpha2,3$- and $\alpha2,6$-linked sialic acids with similar affinity,[48] showed a 35% loss in binding, further supporting a decrease in $\alpha2,6$-sialylation. XBP1s activation also decreased bisecting *N*-glycans, as indicated by decreasing PHA-E signal (Figures 5.2A and S3A left panel).

Fully consistent with the lectin microarray analyses, MALDI-TOF MS analysis confirmed that XBP1s activation of HEK[XBP1s] cells resulted in reduced sialylation of cell membrane *N*-glycans and provided additional detail into the specific epitopes affected. Specifically, in XBP1s-activated cells the relative abundance of ions at *m/z* 3054 and 3415 (green peaks in Figure 5.2B, top versus middle panels) – corresponding to tri-antennary *N*-glycans with one or two NeuAc (*N*-acetylneuraminic acid) residues, respectively – decreased with respect to the relative abundance of the ion at *m/z* 2693, an undecorated tri-antennary *N*-glycan. The ratio of ions at *m/z* 3054/2693 indicated an ~84% decrease in sialylation of tri-antennary glycans in XBP1s-activated cells. Similar results were observed for tetra-antennary *N*-glycans in XBP1s-activated cells (red peaks in Figure 5.2B, top versus middle panel). Specifically, the relative abundances of ions at *m/z* 3503, 3864, and 4226, which correspond to singly, doubly and triply sialylated structures, decreased substantially upon XBP1s activation compared with the relative abundance of the undecorated tetra-antennary *N*-glycan with *m/z* 3142. The ratio of ions at *m/z* 3503/3142 indicated an ~85% decrease in sialylation of tetra-antennary glycans in XBP1s-activated cells. Similar but less intense reductions in sialylation were also observed for higher mass *N*-glycans (Figure S3B). For example, in XBP1s-activated cells (Figure S3B, top versus middle panel) the ions at *m/z* 5124 and 5485 – corresponding mainly to tetra-antennary *N*-glycans with two additional LacNAc (*N*-acetyllactosamine) repeats and three or four NeuAc residues – decreased by ~22 and ~44%, respectively, compared to the relative abundance of the ion at *m/z* 4763, which only has two NeuAc residues (Figure S3B, top versus middle panel). Lower mass bi-antennary glycans also displayed an ~80% decrease in sialylation (Figure S3C, left versus middle panel).

We also performed GC-MS linkage analysis (Table S4) to add linkage details to our findings from the MALDI-TOF MS and TOF/TOF MS/MS, and to further verify our findings from the lectin microarrays. The detection of 6-linked galactose showed that a portion of NeuAc residues were $\alpha2,6$-linked. The abundance of these $\alpha2,6$-linked NeuAc residues substantially decreased in XBP1s-activated HEK[XBP1s] cells (23.6% in vehicle samples versus 4.0% in XBP1s-activated samples), in concordance with the MALDI-TOF MS data and with the lectin microarray data that specifically indicated reduced $\alpha2,6$-sialylation (Figure 5.2A). In addition to the results regarding sialylation, we identified 3,4,6-linked mannose by GC-MS, verifying the presence of minor amounts of bisecting *N*-glycans in the sample. The abundance of these epitopes also decreased upon XBP1s activation HEK[XBP1s] cells (Table S4), again in full agreement with our observation of decreased bisecting *N*-glycans by lectin microarrays (Figure 5.2A).

Collectively, these data reveal that XBP1s activation in HEK[XBP1s] cells substantially remodeled the composition of the membrane *N*-glycome, as summarized in Figure 5.2C and Table S5. Moreover, XBP1s-induced changes in the *N*-glycome were consistently detected both

via lectin microarrays and MS. In particular, we observed a strong reduction in sialylation, specifically α2,6-sialylation, as well as a reduction in bisecting N-glycans.

Next, to ask whether global induction of ER protein misfolding, and therefore all UPR pathways, has related effects on N-glycome composition, we performed the same lectin microarray and MALDI-TOF MS analyses on membrane N-glycans from Tg-treated HEK^{XBP1s} cells. We found that the overall glycomic profile closely mirrored that of selective XBP1s activation in that we observed reduced sialylation and a reduction in bisecting GlcNAc (Table S3A). However, the magnitude of the changes was less substantial. For example, statistically significant lectins SNA-I showed a 19% decrease in binding and diCBM40 showed a 40% decrease in binding (Figures 5.2A and S3A right panel). However, while other α2,6-sialic acid lectins (e.g., PSL and TJA-I) also displayed reduced overall binding, they did not meet the statistical cutoff. A reduction in sialylation was also observed in the MALDI-TOF MS data, evident both in the relative abundance of sialylated to undecorated tri-antennary N-glycans (green peaks in Figure 5.2B, top versus bottom panel) and the relative abundance of sialylated to undecorated tetra-antennary N-glycans (red peaks in Figure 5.2B, top versus bottom panel). The ratios of ions at m/z 3054/2693 and m/z 3503/3142 revealed, in both cases, an ~85% decrease in sialylation on tri- and tetra-antennary N-glycans upon Tg treatment. Sialylation of bi-antennary N-glycans also decreased substantially upon Tg treatment (Figure S3C, left versus right panel). GC-MS linkage analysis (Table S4) confirmed a decrease in α2,6-linked NeuAcs, evidenced by a shift in the relative abundance of 6-linked galactose from 23.6% on vehicle-treated HEK^{XBP1s} cells to 4.1% on Tg-treated cells. Like XBP1s, activation of global ER stress by Tg also caused a loss of bisecting GlcNAc, observed by reduced PHA-E binding in lectin microarrays (Figures 5.2A and S3A right panel) and a decrease in 3,4,6-mannose by GC-MS analysis (Table S4). Notably, a gain in Core 1/3 O-glycans was also observed upon Tg treatment, evidenced by increased binding to the lectins AIA, MPA, and MNA-G (Figures 5.2A and S3A right panel). Detailed O-glycan analysis was not performed using MALDI-TOF MS here, but this last result suggests that further investigation may reveal interesting ER stress-mediated changes to O-linked glycoforms, in addition to N-linked glycoforms. The Tg-mediated alterations in the HEK^{XBP1s} membrane glycome are summarized in Table S6.

XBP1s Remodels HeLa^{XBP1s} Membrane N-Glycoproteomes in a Cell Type-Dependent Manner. We next asked whether activation of the XBP1s transcription factor has broader consequences for N-glycome composition across cell types. In particular, because cells derived from different tissues often have very different baseline N-glycome profiles,[49-51] we were interested in the consequences of XBP1s activation in distinctive cellular contexts. To address this question, we employed our HeLa^{XBP1s} cells following the workflow in Figure 5.1C. We first confirmed that the baseline glycome profiles of HEK^{XBP1s} versus HeLa^{XBP1s} cells were different. By lectin microarray analysis, HEK^{XBP1s} and HeLa^{XBP1s} cells displayed strikingly different levels of oligo-mannose N-glycans (Man$_7$–Man$_5$) and sialylation epitopes (Figure S4). Additionally, we found using MALDI-TOF/TOF MS/MS that, whereas high mass poly-LacNAc N-glycans in HEK^{XBP1s} membranes consisted of linear LacNAc repeats (Figure S3D), in HeLa^{XBP1s} cell membranes the high mass N-glycans consisted of I-branched poly-LacNAc repeats (Figures S5A and S5B). GC-MS linkage analysis confirmed the presence of 3,6-linked galactose in HeLa^{XBP1s} cells, in concordance with the presence of I-branched LacNAcs (Table S4).

Coincident with these quite distinctive baseline N-glycome profiles, lectin microarray analysis revealed a unique shift in the membrane glycomic profile upon XBP1s activation in HeLa^{XBP1s} cells (Figures 5.3A and S5C, with complete data presented in Table S3B) relative to that observed in HEK^{XBP1s} cells. XBP1s-activated HeLa^{XBP1s} membranes exhibited increased levels of less processed N-linked glycans relative to vehicle-treated controls, based on increased signal from several high- (Man$_9$–Man$_7$) and oligo-mannose binding lectins (rGRFT, HHL, NPA, GNL, and Calsepa).[52] Additionally, XBP1s activation resulted in increased core

Figure 5.3

A

Lectin Microarray Analysis of HeLa^XBP1s Membranes
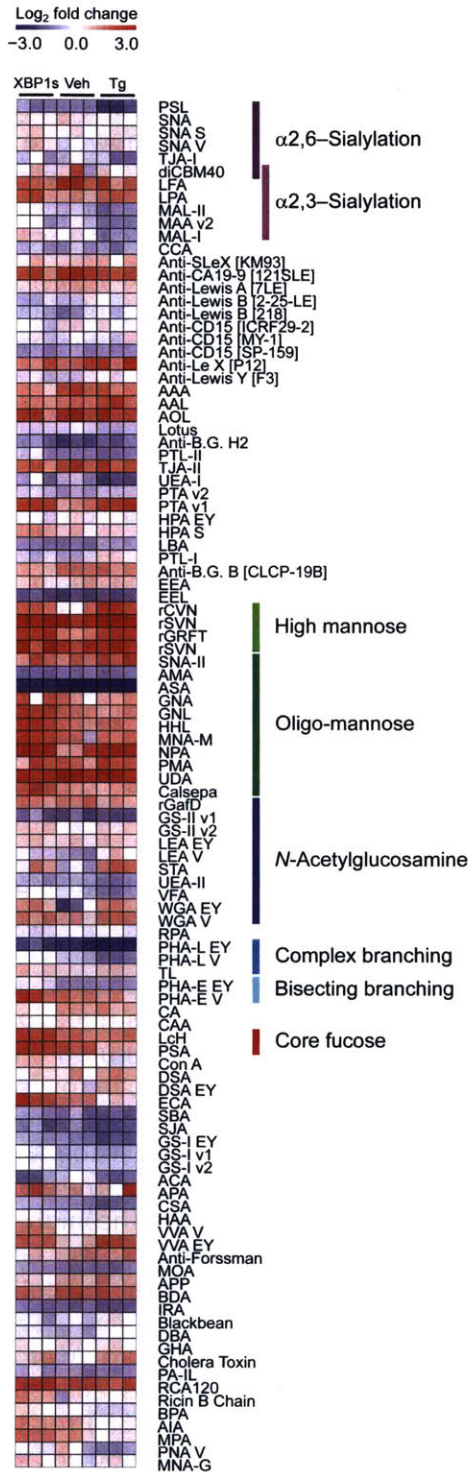
B

MS Analysis of HeLa^XBP1s Membranes

C
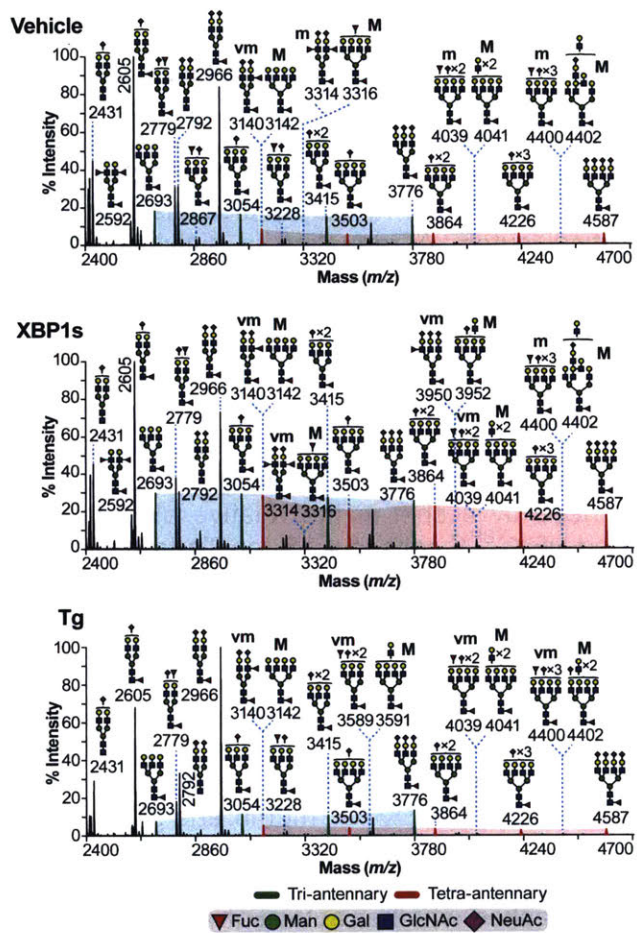
XBP1s-Induced Changes to *N*-Glycans in HeLa^XBP1s Membranes

Increase in high- & oligo-mannose (by lectin microarray and flow)

Increase in core fucose (by lectin microarray & MS)

Increase in tetra-antennary *N*-glycans (by lectin microarray & MS)
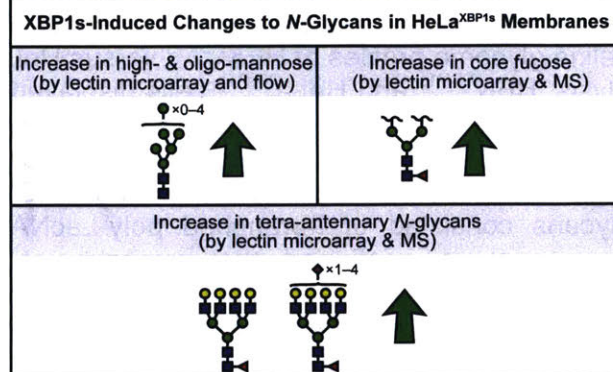
Figure 5.3: *Analysis of the HeLa^(XBP1s) membrane glycoproteome.* (A) Heat map of lectin microarray data generated from isolated HeLa^(XBP1s) membrane glycoproteomes under vehicle-treated, XBP1s-activated, or Tg-treated conditions. Full lectin names, print concentration, and sources are listed in Table S2B. Color intensity represents normalized log2 ratio data relative to a pooled sample reference. Each column represents one biological replicate of the indicated sample. Select lectin groups with shared binding specificities are annotated at right of heatmap. See also Figure S5C and Table S3B in the Appendix. (B) Partial MALDI-TOF mass spectra from the HeLa^(XBP1s) membrane proteome revealed increased abundance of tetra-antennary *N*-glycans upon XBP1s activation. Green and red peaks correspond to tri- and tetra-antennary *N*-glycans, respectively, with various levels of sialylation. Manually inserted green and red shaded areas highlight the increased abundance of tetra-antennary *N*-glycans. Structures outside a bracket were not unequivocally defined, and "M," "m," and "vm," designations indicate major, minor, and very minor abundances, respectively. Putative structures are based on composition, tandem MS, and knowledge of biosynthetic pathways. All molecular ions are [M+Na]⁺. High mass spectra and selected MS/MS fragmentations are presented in Figures S5A and S5B, with GC-MS data in Table S4 (see Appendix). Experiments were carried out on two biological replicates. (C) Table highlighting key XBP1s-induced changes to the HeLa^(XBP1s) membrane *N*-glycome and the analytical methods by which these changes were observed (see also Tables S5 and S6).

fucosylation (PSA and LcH) and β1,6-GlcNAc branching (PHA-L, which binds tri- and tetra-antennary glycans). In contrast to HEK^{XBP1s} cells, no shift in sialylation was observed.

MALDI-TOF MS analysis was fully consistent with these findings from the lectin microarray analyses. Similar to the microarray data, no significant changes in sialylation were observed. For example, the ratio of ions at *m/z* 3054, 3415 and 3776 (corresponding to tri-antennary *N*-glycans with one to three NeuAc residues, respectively) over the ion at *m/z* 2693 (corresponding to non-sialylated tri-antennary *N*-glycan) did not shift between vehicle-treated controls and XBP1s-activated HeLa^{XBP1s} cells (comparing the intensity of the green peaks within the green shaded areas in Figure 5.3B, top versus middle panels). Similarly, no change was observed in sialylation of the tetra-antennary *N*-glycans (comparing the intensity of the red peaks within the red shaded areas in Figure 5.3B, top versus middle panels). Instead, consistent with the lectin microarrays, the most striking consequence of XBP1s activation for the HeLa^{XBP1s} membrane glycome was a 97% increase in the relative abundance of tetra- versus tri-antennary *N*-glycans (intensity of all the red versus all the green peaks in Figure 5.3B, top versus middle panels). GC-MS linkage analysis further confirmed the MALDI-TOF MS and microarray data, demonstrating (i) an increase in tetra-antennary *N*-glycans revealed by increased 2,4- and 2,6-linked mannose in the partially permethylated alditol acetates (Table S4) and (ii) an increase in core fucosylation revealed by increased levels of 4,6-linked GlcNAc (Table S4).

As performed in our experiments here, MALDI-TOF MS does not permit quantitation of relative high- and oligo-mannose levels between two samples, and so we were unable to further validate that particular finding from the lectin microarray using the MS data. Moreover, the experimental strategy of cell membrane isolation by ultracentrifugation can be problematic when assessing changes in levels of less processed *N*-glycans on membrane proteins, as all cell membranes are isolated in the ultracentrifugation step. That is, the method cannot distinguish *N*-glycans displayed on intracellular, immature membrane glycoproteins from *N*-glycans on membrane glycoproteins that have trafficked through the Golgi to the cell surface. To further evaluate the increase in high-mannose levels observed via lectin microarrays upon XBP1s activation (Figure 5.3A) we turned to a lectin-based flow cytometry approach.[53] We treated HeLa^{XBP1s} cells with dox to induce XBP1s or vehicle for 72 h, as above, and then incubated live cells with biotinylated HHL, a lectin that preferentially binds high-mannose *N*-glycans.[54] We then stained cells with Cy5-streptavidin to enable quantitative fluorescent detection of high-mannose *N*-glycans. XBP1s-activated HeLa^{XBP1s} cells displayed a modest but significant increase in HHL signal when compared to vehicle-treated control cells (Figure S6A, left and middle panels). Control experiments inhibiting lectin binding with methyl α-D-mannopyranoside confirmed that HHL was binding high-mannose *N*-glycans, as expected (Figure S6A, right panel). Treatment of cells with the mannosidase I inhibitor 1-deoxymannojirimycin[55, 56] also led to an increase in HHL binding without causing a loss of cell viability (Figures S6B–D), further confirming that HHL detected high-mannose structures.

Collectively, these data reveal that XBP1s activation in HeLa^{XBP1s} cells substantially remodels the composition of the HeLa membrane *N*-glycome in well-defined ways observed via multiple analytical methods, as summarized in Figure 5.3C and Table S5. In particular, we observed strong increases in both core fucosylation and the abundance of tetra-antennary *N*-glycans via both lectin microarray and MALDI-TOF MS analyses. In addition, the combination of lectin microarrays and lectin flow cytometry experiments revealed that XBP1s activation increased levels of high- and oligo-mannose *N*-glycans on the HeLa^{XBP1s} cell surface.

Stress-induced, global UPR activation via Tg treatment of HeLa^{XBP1s} cells also caused increases in high-mannose levels (Figures 5.3A and S5C right panel; see also Table S3B), demonstrated by significantly increased binding of the α1,2-mannose-specific antiviral lectins rSVN and rCVN, in addition to rGRFT. However, Tg treatment also induced a loss of β1,6

branched N-glycans based on decreased signal from PHA-L, revealing a distinct difference from the XBP1s-induced signal (Figures 5.3A and S5C right panel). Again, no changes in sialylation were observed. These observations were fully confirmed by our MALDI-TOF MS analysis, in which Tg-treated cells did not display increased tetra-antennary N-glycans relative to vehicle-treated cells (red versus green shading in Figure 5.3B, top versus bottom panels). Instead, as anticipated from the lectin microarray work (Figure 5.3A), a decrease in 2,4- and 2,6-linked mannose was observed by GC-MS after Tg treatment, the latter being consistent with a loss of β1,6 branched N-glycans (Table S4). The Tg-mediated alterations in the HeLa^{XBP1s} membrane glycome are summarized in Table S6.

In summary, the ability of XBP1s (or even stress-induced UPR activation) to globally remodel the N-glycome extends beyond just a single cell type. The molecular nature of the resulting changes in N-glycan distribution patterns is cell type-dependent, likely owing at least in part to differences in baseline N-glycome composition, and the size of the effects can be quite substantial.

XBP1s Remodels the N-Glycan Composition of the HEK^{XBP1s} Secretome. The observation that XBP1s activation alters the structure distribution pattern of N-glycans on cell surfaces in a cell type-dependent manner suggests that cell–cell interactions, cell adhesion, and receptor signaling could all be modified by XBP1s. Beyond cell membranes, a large portion of the N-glycoproteome is comprised of soluble, secreted proteins. Alterations in the N-glycans on secreted proteins could provide a mechanism for long-distance transmission of cell stress signals, or engender other important phenotypes. Therefore, we next asked whether XBP1s-mediated remodeling extends to the secreted N-glycome.

The secretome from XBP1s-activated HEK^{XBP1s} cells displayed substantial N-glycome remodeling when analyzed by lectin microarrays (Figures 5.4A and S7 left panel, with complete data presented in Table S3C). A decrease in both β1,6 N-glycan branching (PHA-L) and bisecting GlcNAc (PHA-E) was observed, as was also observed in the HEK^{XBP1s} membrane N-glycome. A modest (27%) but not statistically significant decrease in binding to α2,6-sialylation epitopes based on the lectins SNA-I and diCMB40 was also observed. These results were confirmed by our MALDI-TOF MS and GC/MS analyses (see below). We also observed a significant increase in binding to oligo-mannose-targeting lectins (HHL, UDA, NPA, Calsepa, and GNA), a consequence of XBP1s activation that was notably absent from the HEK^{XBP1s} membrane samples (Figure 5.2A).

MALDI-TOF MS analysis (Figure 5.4B) revealed that N-glycans isolated from the secretome of HEK^{XBP1s} cells featured abundant core fucosylated bi- (purple peaks; m/z 2244, 2605, and 2966), tetra- (red peaks; m/z 3142, 3503, 3864, 4226 and 4587), and to a lesser extent tri-antennary (green peaks; m/z 2693, 3054, 3415, and 3776) N-glycans. Upon XBP1s activation, the tri- and tetra-antennary, but not the bi-antennary, N-glycans exhibited reduced sialylation (Figure 5.4B, upper versus middle panel). For example, the relative abundance of molecular ions at m/z 3503, 3864, 4226, and 4587 compared to m/z 3142 was reduced from vehicle-treated levels in the XBP1s-activated secretome (red peaks in Figure 5.4B, upper versus middle panel). GC-MS analysis confirmed a small reduction in α2,6-sialylation overall (Table S4: 6-linked Gal: Veh - 98.9% versus XBP1s - 63.8%), consistent with our lectin microarray analysis. The decreases in bisected N-glycans and β1,6-branching observed by microarray were also confirmed by our GC-MS data (Table S4: 3,4,6-linked Man: Veh - 4.7% versus XBP1s - 2.3%; 2,6-linked Man: Veh - 105.8% versus XBP1s - 71.1%).

Collectively, these data reveal that XBP1s activation in HEK^{XBP1s} cells substantially remodels the composition of the secreted N-glycome, as summarized in Figure 5.4C and Table S5. Lectin microarray analyses revealed an increase in oligo-mannose levels. Both lectin microarray and MALDI-TOF MS analyses demonstrated a decrease in bisecting N-glycans.

117

Figure 5.4

A — Lectin Microarray Analysis of the HEK^XBP1s Secretome

B — MS Analysis of the HEK^XBP1s Secretome

C — XBP1s-Induced Changes to N-Glycans in the HEK^XBP1s Secretome

118
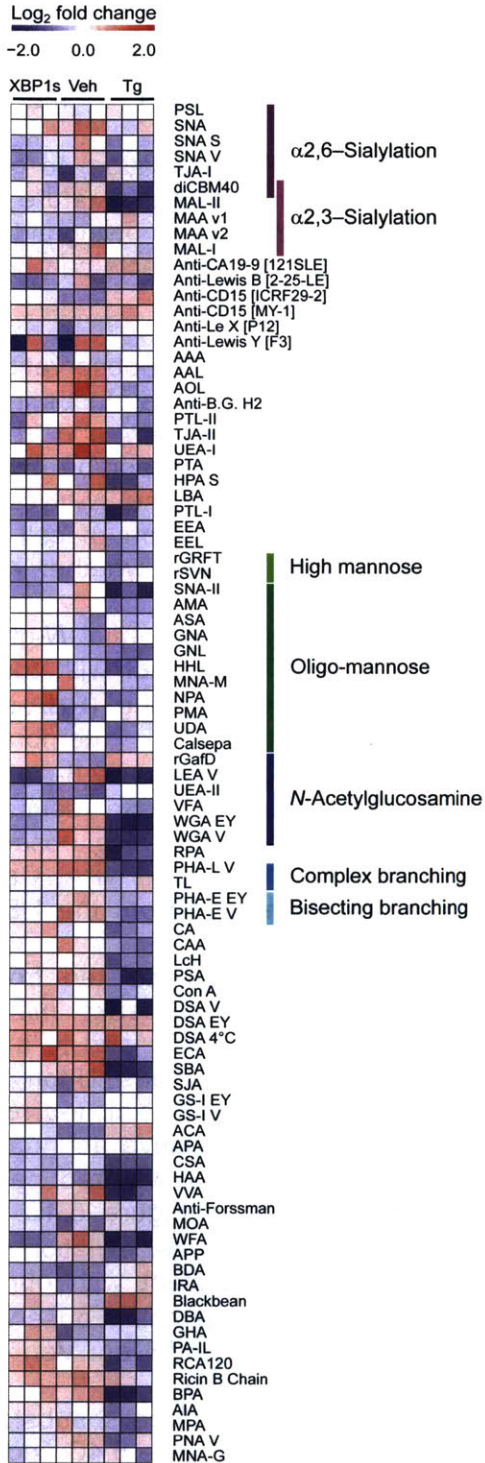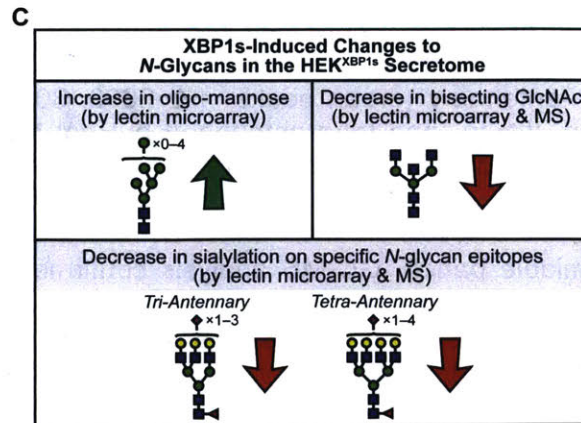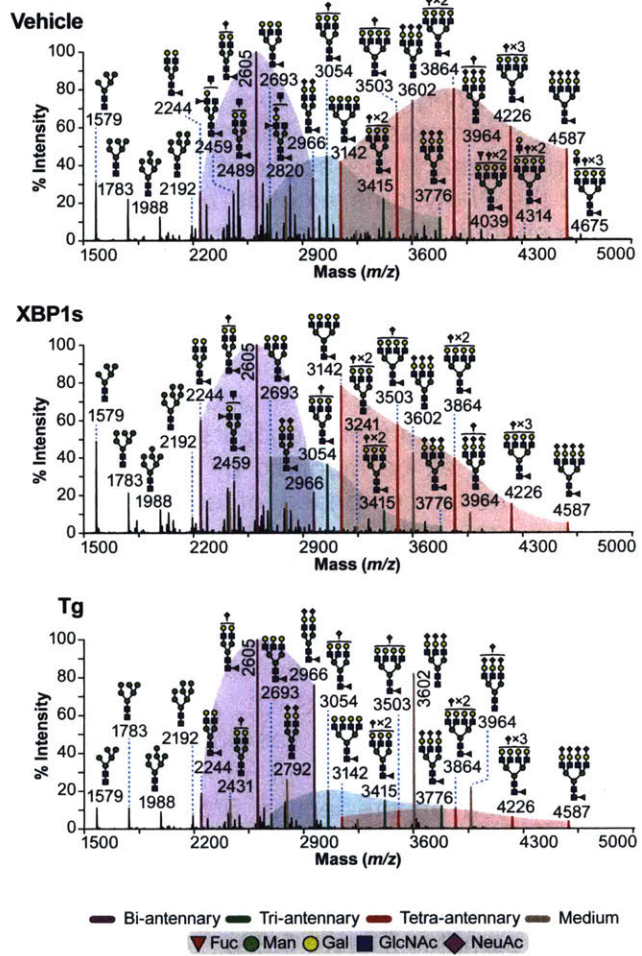
Figure 5.4: *Glycomic analysis of the HEK^{XBP1s} secretome.* (A) Heat map of lectin microarray data generated from isolated HEK^{XBP1s} secreted glycoproteomes under vehicle-treated, XBP1s-activated, or Tg-treated conditions. Full lectin names, print concentration, and sources are listed in Table S2A. Color intensity represents normalized log2 ratio data relative to a pooled sample reference. Each column represents one biological replicate of the indicated sample. Select lectin groups with shared binding specificities are annotated at right of heatmap. See also Figure S7 and Table S3C. (B) MALDI-TOF mass spectra of the HEK^{XBP1s} secretome revealed modestly decreased sialylation of tri-antennary and tetra-antennary N-glycans upon XBP1s activation. Purple, green and red peaks correspond to bi-, tri- and tetra-antennary N-glycans with various levels of sialylation, respectively. Manually inserted purple, green and red shaded areas highlight the distribution shift of sialylation on the corresponding bi-, tri- and tetra-antennary N-glycans. Non-core fucosylated structures corresponded to residual fetal bovine serum-derived N-glycans that we also characterized previously (brown peaks; *m/z* 2792, 3602, and 3964) [57]. These fetal bovine serum-derived N-glycans were observed consistently in all samples and were excluded from the analyses as previously described [58, 59]. Structures outside a bracket have not been unequivocally defined. Putative structures are based on composition, tandem MS, and knowledge of biosynthetic pathways. All molecular ions are [M+Na]$^+$. Experiments were carried out on two biological replicates. GC-MS data are presented in Table S4 (see Appendix). (C) Table highlighting key XBP1s-induced changes to the HEK^{XBP1s} secreted N-glycome and the analytical methods by which these changes were observed (see also Tables S5 and S6).

119

MALDI-TOF MS showed a loss of sialylation on specific N-glycan epitopes (for tri- and tetra-antennary but not bi-antennary N-glycans), suggesting possible protein-specific effects influencing loss of sialylation. While the reduction in bisecting N-glycans and a loss of sialylation are both consistent with our observations for the HEK^{XBP1s} membrane N-glycomes (Figure 5.2C), the increase in oligo-mannose levels is limited to the secretome. This observation highlights the importance of separately evaluating membrane and secreted N-glycans in glycomic experiments, as we have also noted in prior work.[60-63]

We also analyzed the consequences of stress-induced, global UPR activation via Tg treatment on the composition of the secreted N-glycome in HEK^{XBP1s} cells. The secreted N-glycome of Tg-treated HEK^{XBP1s} cells was similar to that of XBP1s-activated cells in that reduced levels of β1,6 N-glycan branching (decreased binding to PHA-L) and bisecting GlcNAc (decreased binding to PHA-E) were observed (Figures 5.4A and S7 right panel; see also Table S3C). MALDI-TOF MS and GC-MS analyses (Figure 5.4B, top versus bottom panel, and Table S4) confirmed the decreases in levels of tri- and tetra-antennary N-glycans (β1,6 N-glycan branching shown by changes in 2,4-linked mannose and 2,6-linked mannose levels). There was no impact on overall α2,6-sialylation, although a decrease in α2,3-sialylation was observed via reduced binding to MAL-I, MAL-II, and diCBM40 (Figure 5.4A). In contrast to the secretome of XBP1s-activated cells, Tg treatment did not induce an increase in oligo-mannose levels. The Tg-mediated alterations in the HEK^{XBP1s} secreted glycoproteome are summarized in Table S6.

XBP1s Does Not Significantly Alter the Proteomic Composition of the HEK^{XBP1s} Secretome. We questioned whether XBP1s-driven changes in the N-glycan architecture of the HEK^{XBP1s} secretome could be attributed to changes in the composition of the proteome caused by XBP1s activation. To assess this hypothesis, we induced XBP1s by dox treatment for 24 hours and then collected conditioned, serum-free DMEM after a 6 hour incubation on vehicle-treated or dox-treated cells in biological triplicate. We used serum-free DMEM, not Freestyle™ media optimized for HEK cell growth, because Freestyle™ media contains various proprietary peptides that interfere with MS analysis. The use of serum-free DMEM limited us to a 6 hour incubation instead of the longer incubation used in our glycomic experiments, to avoid extensive autophagy induction and cell toxicity. We precipitated total protein from the collected media, followed by denaturing, reducing, alkylating, and trypsinizing to peptides. The digested samples were then labeled with isobaric mass tags to permit quantitative assessment of proteome composition and identify any XBP1s-mediated changes. Using this approach, we were able to quantify 700 secreted proteins across all samples (Table S7). Using significance thresholds of unadjusted p-value ≤ 0.05 and fold-change ≥ 1.5, we found that levels of only 16 of these 700 proteins were changed by XBP1s activation (we further note that no proteins qualified using a false discovery rate (FDR) threshold of 10% by Benjamini-Hochberg analysis). Of those 16 proteins, only 9 are annotated in UniProtKB as being N-glycosylated. These data lead us to conclude that the observed changes in the secreted glycome are unlikely to be driven by changes in the composition of the secreted proteome. Rather, we hypothesize that altered biosynthesis of N-glycans is the controlling factor.

Selective XBP1s Induction Remodels the Glycogene Transcriptome. Control mechanisms of N-glycan biosynthesis are not well understood. One possible way in which induction of XBP1s could lead to an altered glycome is through enhanced flux through the secretory pathway that overwhelms the N-glycan maturation machinery and causes uncoordinated changes in structure distribution patterns. However, global changes in protein secretion that might suggest the Golgi N-glycan maturation pathway is overwhelmed upon XBP1s activation have not been detected (see Table S7 and work by Wiseman and colleagues[14]). Moreover, while the changes in oligo-mannose levels in HeLa^{XBP1s} membranes and HEK^{XBP1s} secretomes could possibly be linked to increased flux overwhelming the N-glycan maturation pathway, it seems unlikely that many of

the other changes observed (e.g., increases in core fucosylation, alterations in *N*-glycan branching, and changes in specific types of sialylation) could be explained by this hypothesis. Furthermore, changes in oligo-mannose levels were not observed in the HEK^{XBP1s} membrane glycome. Thus, although we cannot completely exclude the possibility, we believe the 'overwhelming flux' hypothesis is unlikely to be correct.

A second possible explanation is that XBP1s activation either directly (as it is a transcription factor), or indirectly (as it can alter the expression of other transcription factors[12, 14, 64]) modifies the expression of enzymes involved in *N*-glycan maturation and biosynthesis (known as glycogenes), leading to transcriptionally encoded changes in *N*-glycome architecture. This explanation would be consistent with prior work showing that alterations in the expression of glycogene transcripts can induce changes in the *N*-glycome.[51, 65] To evaluate this possibility, we used both previously published data for HEK^{XBP1s} cells[14] and our new RNA-Seq data for HeLa^{XBP1s} cells to characterize the glycogene transcript-level changes induced by chronic, selective XBP1s activation in vehicle- versus dox-treated cells. In both HEK^{XBP1s} cells and HeLa^{XBP1s} cells, gene set enrichment analyses highlighted *N*-linked glycosylation as a highly enriched pathway under conditions of chronic, selective XBP1s activation (see Table S1C and Ref. 14). Indeed, GO groups related to protein glycosylation are the most enriched gene sets in XBP1s-activated cells, after XBP1s- and IRE1-regulated aspects of the classical unfolded protein response. Furthermore, standard GSEA indicated strong enrichment of genes involved in biosynthesis of the lipid-linked oligosaccharide in both HEK^{XBP1s} and HeLa^{XBP1s} cells (Figure S8 and Table S1E). However, our data suggest that XBP1s activation does not globally alter transcripts involved in *N*-glycan maturation, as examination of the comprehensive gene set involved in *N*-glycan trimming and elaboration showed no global enrichment in XBP1s-activated cells (Figure S8 and Table S1E). This observation suggests that XBP1s regulates specific subsets of these glycogenes, consistent with the changes in levels of only specific epitopes observed in our glycomic analyses.

We next manually curated a list of 1310 genes related to protein *N*-glycosylation from gene ontology data sets (the gene sets used and full list of *N*-glycosylation-related genes analyzed are shown in Tables S8A and S8B, respectively). These *N*-glycosylation-related genes included not just enzymes involved in *N*-glycan biosynthesis, trimming, and modification, but also sugar transporters, monosaccharide synthases, and catabolic enzymes. We plotted the changes in expression levels of these individual glycogenes upon XBP1s activation for HEK^{XBP1s} and HeLa^{XBP1s} cells in the volcano plots in Figures 5.5A and 5.5C, respectively. As some glycogenes are expressed at low levels,[66] we also employed a commercially available Qiagen RT^2 Profiler Human Glycosylation PCR Array to analyze transcript levels of selected glycogenes. For transcripts that were detected by multiple methods, we used the qPCR array results, rationalizing that the latter is more sensitive.[67]

For HEK^{XBP1s} cells, 39 of 950 expressed glycogene transcripts were altered (shown in teal in Figure 5.5A) by chronic XBP1s activation using a significance cutoff of false discovery rate (FDR; for microarray data) or *p*-value (for qPCR array data) $\leq$ 0.05, in addition to a fold-change threshold $\geq$ 1.5 (see Table S8C for a complete list of results). This analysis (1) confirms the remodeling of the glycogene transcriptome induced by XBP1s activation in HEK^{XBP1s} cells, especially for a number of glycogenes directly involved in *N*-glycan maturation, and (2) is in concordance with key features of the *N*-glycome remodeling caused by XBP1s activation in both the membrane and secretome of HEK^{XBP1s} cells (Figures 5.2C and 5.4C, respectively). Specifically (Figure 5.5B), reduced levels of *MGAT3* (shown in pink in Figure 5.5A), which is the enzyme responsible for introduction of bisecting GlcNAc residues,[68] correlate with the reduction in bisected *N*-glycans observed most significantly in the HEK^{XBP1s} membrane *N*-glycome (Figure 5.2C), but also in the secreted *N*-glycome (Figure 5.4C). Although below the threshold for significance, there was also a decrease in the transcripts of the α2,6-sialyltransferase *ST6GAL1* (0.69-fold for XBP1s-activated HEK^{XBP1s}/control, *p* = 0.23) and an increase in transcript levels

A Glycogene Transcripts in HEK^{XBP1s} Cells

B Correlation Between XBP1s-Induced Changes to the HEK^{XBP1s} Transcriptome and *N*-Glycome

C Glycogene Transcripts in HeLa^{XBP1s} Cells

D Correlation Between XBP1s-Induced Changes to the HeLa^{XBP1s} Transcriptome and *N*-Glycome
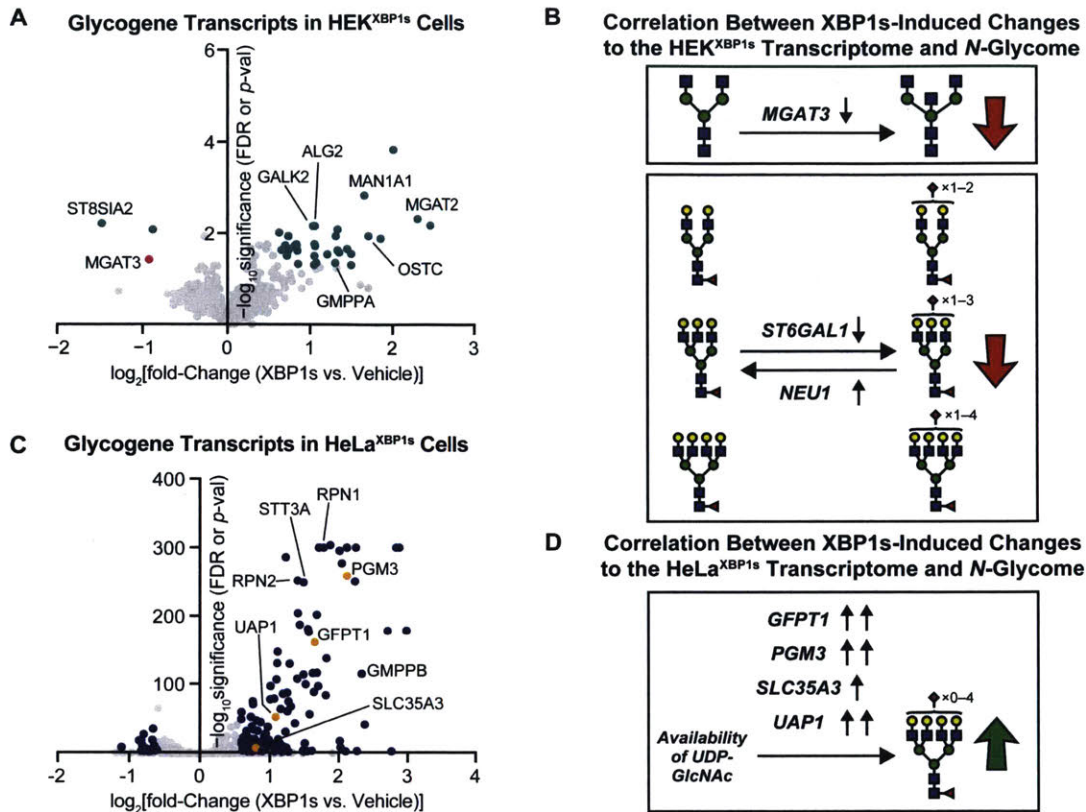
Figure 5.5: Glycogene analysis in HEK^{XBP1s} and HeLa^{XBP1s} cells. Transcript levels for numerous enzymes involved in key steps of *N*-glycan biosynthesis and maturation pathways were modulated upon XBP1s activation. (A) Volcano plot showing XBP1s activation-induced changes in glycosylation-related transcripts in HEK^{XBP1s} cells. Glycogenes analyzed are listed in Table S8B. Data shown were obtained either from the Qiagen Human Glycosylation qPCR Array or, if a transcript of interest was not included in the qPCR array, extracted from previously published HEK^{XBP1s} microarray data.[14] Transcripts shown in teal meet significance and fold-change thresholds of FDR or *p*-value ≤ 0.05 and fold-change ≥ 1.5 upon XBP1s activation. Gene symbols are shown for outliers and transcripts of particular interest. (B) XBP1s-mediated changes in the HEK^{XBP1s} transcriptome directly correlate with some of the observed changes to the HEK^{XBP1s} *N*-glycome. Decreased expression of *MGAT3* (shown in pink in (A)), a GlcNAc transferase,[68] could account for the loss in bisecting GlcNAc observed on HEK^{XBP1s} secreted and membrane glycoproteins upon XBP1s activation. Decreased expression of *ST6GAL1* and increased expression of *NEU1* could account for the reduced sialylation observed for HEK^{XBP1s} secreted and membrane glycoproteins upon XBP1s activation. (C) Volcano plot showing XBP1s activation-induced changes in glycosylation-related transcripts in HeLa^{XBP1s} cells, generated from Qiagen Human Glycosylation qPCR Array or, if a transcript of interest was not included in the qPCR array, RNA-Seq data as described for (A). Transcripts shown in blue meet significance and fold-change thresholds of FDR or *p*-value ≤ 0.05 and fold-change ≥ 1.5 upon XBP1s activation. (D) XBP1s-mediated changes in the HeLa^{XBP1s} transcriptome directly correlate with at least some of the observed changes to the HeLa^{XBP1s} *N*-glycome. Increased expression of *GFPT1*, *PGM3*, *SLC35A3*, or *UAP1*, which are enzymes involved in regulating UDP-GlcNAc availability (shown in orange in (C)), could contribute to increased tetra-antennary *N*-glycans observed on HeLa^{XBP1s} membranes upon XBP1s activation (see also Figure S9).

122

for the sialidase *NEU1* (1.4-fold for XBP1s-activated HEK^{XBP1s}/control, $p = 0.25$) that together may explain the loss of α2,6 sialylation, especially on more branched structures, that we observed (Figure 5.5B). Interestingly, the transcript levels for *ST6GAL1* go up in HeLa^{XBP1s} cells upon XBP1s activation (Table S8D, 2.3-fold for XBP1s-activated HeLa^{XBP1s}/control, $p = 0.060$), which may balance the increase in *NEU1* (3.0-fold for XBP1s-activated HeLa^{XBP1s}/control, $p = 0.003$) and account for the lack of change in α2,6 sialic acid in that cell line (see Figure 5.3A).

For HeLa^{XBP1s} cells, 155 of 1075 expressed glycogene transcripts were altered (shown in blue in Figure 5.5C) by chronic XBP1s activation using a significance cutoff of FDR (for RNA-Seq data) or *p*-value (for qPCR array data) ≤ 0.05, in addition to a fold-change threshold ≥ 1.5 (see Table S8D for a complete list of results). As was also the case in the HEK^{XBP1s} transcriptome data, this analysis again (1) confirms the remodeling of the glycogene transcriptome induced by XBP1s activation in HeLa^{XBP1s} cells, especially for genes directly involved in *N*-glycan maturation, and (2) is again in concordance with at least one key feature of the *N*-glycome remodeling caused by XBP1s activation in HeLa^{XBP1s} cell membranes (Figure 5.3C). Specifically (Figure 5.5D), XBP1s-activated HeLa^{XBP1s} cells exhibited increased expression of genes affecting nucleotide sugar availability (see Figure S9 and Table S8D). Among the genes significantly overexpressed, we observed the plasma membrane glucose transporters, such as GLUT1 and GLUT3 (encoded by *SLC2A1* and *SLC2A3*, respectively), which could potentially increase glucose availability in the cytosol. We also identified enzymes in the hexosamine biosynthetic pathway required for the biosynthesis of the UDP-GlcNAc sugar donor,[69, 70] including *PGM3* (4.35-fold increase), *UAP1* (2.12-fold increase), and the rate-limiting *GFPT1* (3.15-fold increase).[71] The increased expression of these enzymes, which collectively regulate cytosolic levels of UDP-GlcNAc (Figure S9), was accompanied by an increase in the Golgi UDP-GlcNAc transporter *SLC35A3*, thus increasing potential UDP-GlcNAc availability for Golgi glycosyltransferases.[72] Levels of UDP-GlcNAc have been shown to affect the levels of tri- and tetra-antennary *N*-glycans.[70] It is therefore likely that, even in the absence of altered *MGAT5* expression, the abundance of tetra-antennary *N*-glycans observed in HeLa^{XBP1s} cells might increase with XBP1s activation owing to increased UDP-GlcNAc concentrations in the Golgi.[73] The connection between changes in UDP-GlcNAc levels and MGAT5-dependent β1,6 branching, in the absence of protein expression changes, is well-established.[74] In further support of this hypothesis, changes in the level of functional *PGM3* have been correlated with altered *N*-glycomic profiles in human neutrophils[75] and increased expression of *UAP1* in human prostate cancer cells has been correlated with increased abundance of UDP-GlcNAc.[76]

Collectively, these transcript-level results provide insights into the detailed mechanism of some of the XBP1s-induced changes in cellular *N*-glycome architectures. XBP1s activation results in a coordinated remodeling of the glycogene transcriptome at multiple levels, particularly for specific subsets of genes involved in *N*-glycan maturation and the availability of nucleotide sugar donors. In several cases, the alterations in glycogene transcript levels are in concordance with *N*-glycome-level changes induced by XBP1s activation. It should be noted, however, that *N*-glycosylation is regulated at many levels – not just transcriptionally, but also at the levels of translation (via miRNA), enzyme localization, and sugar metabolism/transport.[29, 30, 51, 66, 77-79] Thus, it is expected that XBP1s-induced changes in the glycogene transcriptome would be unlikely to explain all the *N*-glycome changes we observed. Altogether, these data are consistent with XBP1s-mediated changes in the *N*-glycome reflecting a coordinated response driven in part by glycogene transcriptome remodeling.

Discussion

Previous work has revealed the paradigmatic role of XBP1s in resolving protein misfolding stress and in expanding the secretory pathway.[80] The impact of XBP1s activation on *N*-glycan maturation for specific ectopically overexpressed secreted proteins and possibly on *N*-glycan site occupancy has also been demonstrated.[29, 37] However, prior to this work no evidence

that selective XBP1s activation could alter *N*-glycan maturation on endogenous proteins existed. Strikingly, our data show that chronic XBP1s activation can generate changes of sufficient magnitude to be detected in glycome-wide assays that assess global carbohydrate composition in cell membranes and in the secretome.

XBP1s-mediated changes in *N*-glycome composition occur across multiple cell types. The nature of these changes displays some cell type specificity, likely owing to the very different baseline *N*-glycome compositions of different cells. As summarized in Table S5, upon stress-independent XBP1s activation in three different sets of samples, we observed increases in oligo-mannose levels in two sample sets (HEK^{XBP1s} secretomes and HeLa^{XBP1s} membranes), reductions in sialylation in two sample sets (HEK^{XBP1s} membranes and HEK^{XBP1s} secretomes), decreases in bisecting GlcNAc in two sample sets (HEK^{XBP1s} membranes and HEK^{XBP1s} secretomes), and increases in core fucosylation and tetra-antennary *N*-glycans in one sample set (HeLa^{XBP1s} membranes). These results were consistently observed by multiple analytical methods, including lectin microarrays, MALDI-TOF MS, GC-MS, and lectin flow cytometry. We note that, in many cases, related changes were also induced by global, stress-mediated UPR activation using Tg treatment (Table S6). However, the capacity of XBP1s activation alone to remodel the *N*-glycome indicates the key role of this transcription factor in driving *N*-glycome remodeling and highlights that ER stress is not required.

XBP1s-induced alterations in *N*-glycome architectures appear to reflect, at least in part, a coordinated response owing to remodeling of the glycogene transcriptome. In particular, changes in the expression of specific transcripts encoding enzymes directly involved in *N*-glycan maturation were induced by XBP1s activation. *N*-Glycome remodeling driven by changes in glycogene transcript levels is consistent with observations that differential expression of glycosylation-related transcripts,[51, 67] or even individual glycogenes,[81] can give rise to cell type- and disease-specific glycosylation profiles. However, full prediction of how glycogene transcriptome changes will be reflected in secreted and cell membrane *N*-glycome architectures remains very challenging, likely owing to the fact that multiple variables, and not just glycogene transcript levels, play a role in *N*-glycan biosynthesis.[29, 30, 51, 66, 77-79] Thus, unsurprisingly, although some features of the XBP1s-remodeled *N*-glycome correlate with transcript-level changes in glycogene expression (Figures 5.5B and 5.5D), a direct correlation with all features does not exist. Further studies to fully detail the intermediate steps that propagate XBP1s activation to an altered *N*-glycome are an important subject for future work.

Another interesting aspect of our data is the differential impact of XBP1s activation on membrane versus secreted glycoproteins, particularly with respect to an increase in oligo-mannose levels in the secretome. Secreted glycoproteins on exosomes have been shown to have a glycosylation phenotype distinct from that of the originating cell surfaces.[60] Moreover high- and oligo-mannose is one of the enriched epitopes observed in exosomes, and there are data indicating that glycosylation can act as a trafficking marker for these secreted vesicles.[82] In the body, secreted glycoproteins can impact biology at sites distant to the cell, and there are many glycan-binding proteins involved in innate immunity. Non-cell autonomous UPR signaling has been observed in metabolic regulation, immune system activation, and tumorigenesis,[83] all processes that are also responsive to target- and/or epitope-specific glycosylation.[35, 84, 85] Cumulatively, these observations raise the intriguing possibility that cells may use XBP1s-enacted changes in glycosylation in the secretome to signal for immune system functions or for cell non-autonomous stress signaling.[23, 86-88]

Selective activation of the IRE1-XBP1s arm of the UPR, even in the absence of ER protein misfolding stress, is a widely observed biological phenomenon. For example, selective XBP1s activation is involved in both memory formation and circadian clock regulation.[22, 89] Similarly, selective induction of the XBP1s arm of the UPR plays a critical role in aspects of the immune response and in development.[20, 21] The molecular mechanisms by which XBP1s activity influences these processes continue to be investigated, but our results indicate that a functional

role for XBP1s-altered glycosylation must be considered. Beyond these phenomena, selective, chronic XBP1s activation is also commonly observed in malignancies.[24, 25] One common view suggests that the IRE1-XBP1s pathway is constitutively activated in these cancer cells to resolve ER protein misfolding stress. However, it is noteworthy that another common feature of cancer cells is modified *N*-glycosylation.[85, 90-93] Our findings open the possibility of a mechanistic connection between chronic XBP1s activity in cancer and metastasis-promoting, neoplastic *N*-glycosylation patterns.

In conclusion, our results suggest a new function of the XBP1s arm of the UPR that merits more detailed mechanistic and functional investigation to fully elucidate the roles of the UPR in cellular homeostasis. In ongoing work focused on this objective, we are now using glycoproteomic strategies to identify the specific endogenous proteins whose *N*-glycans are altered by XBP1s, and characterizing the functional consequences of such *N*-glycome remodeling. This phenomenon provides a potential new pathway for intracellular stress signaling to be propagated to the extracellular milieu. XBP1s-induced changes to the *N*-glycome also may hold implications for pathologic processes, such as development of neoplastic glycosylation patterns that support cancer metastasis.

Materials and Methods

Cells and Reagents. HEK[XBP1s] cells were cultured as previously described.[37] HeLa-TREx cells were obtained from Invitrogen and cultured in complete DMEM supplemented with 10% fetal bovine serum (FBS), as well as geneticin sulfate (G418, 500 µg/mL) to maintain the tetracycline repressor. A pLenti4.XBP1s construct, along with a linear marker for puromycin resistance, was transfected into HeLa.TREx cells using Xfect (Clontech). Stable HeLa[XBP1s] cell lines were selected by culturing in puromycin (0.5 µg/mL). An optimized single colony was selected and characterized by Western blot and qPCR before use in these experiments.

RNA Extraction, Real-Time qPCR, RNA-Seq and Membrane/Secretome Preparation for Glycomic Analyses. Detailed protocols are provided in the Appendix.

Lectin Microarray Glycomic Analyses. Lectin microarrays were generated as previously described.[94] Briefly, they were manufactured in-house with a Nano-plotter v2.0 piezoelectric non-contact array printer (GeSiM) using a nano A-J tip. Arrays were printed on Nexterion Slide H (Schott Nexterion) under 50% relative humidity at a surface temperature of 12 °C. Commercial lectins and antibodies were purchased from Vector Labs, R&D Systems, Santa Cruz, TCI, AbCam, E.Y. Labs, or Sigma-Aldrich. The recombinant lectins rGRFT, rCVN, and rSVN were generous gifts from Dr. B. O'Keefe (NCI Frederick). For a list of all printed lectins see Table S2. We note that while the diversity of printed lectins allows for a wide range in the detection of glycan epitopes, we are unable to observe some epitopes (e.g. α2,8-linked sialic acids) on our current array.

Prior to sample hybridization, lectin microarray slides were blocked for 1 h with 50 mM ethanolamine in 50 mM sodium borate buffer (pH 8.8) followed by three washes with 0.005% PBS-T (pH 7.4). Sample protein concentration and the degree of fluorescent label incorporation was determined by measuring absorbances at 280, 555, and 650 nm per the manufacturer's instructions on a NanoDrop ND-2000c spectrophotometer (Thermo Scientific). Equal protein amounts (5 µg) of sample and contrasting labeled reference were mixed in 0.005% PBS-T (pH 7.4) for a final concentration of 67 ng/µL of protein. Slides were then loaded into a hybridization cassette (Arrayit) to isolate individual arrays (24 per slide). Samples were loaded onto individual arrays along with one array for the reference vs reference sample per slide. Samples were hybridized for 2 h at 25 °C with gentle agitation. After hybridization, samples were removed and arrays were washed 4× with 0.005% PBS-T (pH 7.4) for 10 min each. Slides were removed,

submerged in ddH$_2$O, and spun dry. Arrays were scanned using a GenePix 4300A array scanner (PMT 550 laser power 100% for both fluorescent channels).

Background-subtracted median fluorescence intensities were extracted using GenePix Pro v7.2. Non-active lectins were defined as having an average of both channel SNRs <3 in >90% of the data and removed prior to further analysis. Data were median-normalized in each fluorescent channel and the log2 of the sample/reference ratio was calculated for each technical replicate for each lectin. Technical replicates were then averaged for each lectin within each array. The ratios across individual biological triplicates per lectin were compared across treatments using a two-tailed Student's *t*-test.

MALDI-TOF MS and TOF/TOF MS/MS Glycomic Analyses. All samples were treated as described previously.[95] Briefly, each sample was subjected to sonication in the presence of detergent (CHAPS), reduction in 4 M guanidine-HCl (Pierce, Cramlington, Northumberland, UK), carboxymethylation, and trypsin digestion. The digested glycoproteins were then purified by HLB plus C$_{18}$-Sep-Pak (Waters Corp, Hertfordshire, UK; 186000132). *N*-Glycans were released by peptide *N*-glycosidase F (E.C. 3.5.1.52; Roche Applied Science, Burgess Hill, UK) digestion. Released *N*-glycans were permethylated using the sodium hydroxide procedure and purified by classic C$_{18}$-Sep-Pak (Waters, WAT051910). Permethylated *N*-glycans were eluted at the 50% acetonitrile fraction. We note that polysialylated structures deriving from the activity of α2,8-sialyltransferase (ST8SIA2) enzyme were not analyzed from the above released *N*-glycans.[96]

MS and MS/MS data were acquired using a 4800 MALDI-TOF/TOF (Applied Biosystems, Darmstadt, Germany) mass spectrometer. Permethylated samples were dissolved in 10 μL of methanol and 1 μL of dissolved sample was premixed with 1 μL of matrix (10 mg/mL 3,4-diaminobenzophenone in 75% (v/v) aqueous acetonitrile), spotted onto a target plate and dried under vacuum. For the MS/MS studies the collision energy was set to 1 kV, and Ar was used as the collision gas. The 4700 Calibration standard kit, calmix (Applied Biosystems), was used as the external calibrant for the MS mode and [Glu1] fibrinopeptide B human (Sigma-Aldrich) was used as an external calibrant for the MS/MS mode.

MS and MS/MS data were processed using Data Explorer 4.9 Software (Applied Biosystems). The processed spectra were subjected to manual assignment and annotation with the aid of a glycobioinformatics tool, GlycoWorkBench.[97] The proposed assignments for the selected peaks were based on ^{12}C isotopic composition together with knowledge of the biosynthetic pathways. The proposed structures were then confirmed by data obtained from MS/MS and linkage analysis experiments.

For MALDI-TOF analysis of secretome samples, we note that even in commercial "serum-free" media, MS detects serum-derived glycoproteins.[57] In prior work we have shown that the FBS-derived glycans can be identified and removed from the analysis without impacting analysis of other glycan structures, just as we have done here.[58, 59] Thus, the presence of these peaks does not significantly impact detection of other glycan structures in the secretome.

GC-MS Glycan Linkage Analyses. Partially methylated alditol acetates (PMAAs) were prepared as previously described.[95] Linkage analyses of PMAAs were performed on a Scion 456-GC SQ instrument (Bruker) fitted with a BR-5ms fused silica capillary column (15 m × 0.25 mm i.d.; Bruker). The sample was dissolved in ~20 μL of hexanes and injected manually (4–5 μL) at a split ratio of 1/10. Injector temperature was set at 250 °C. Helium was used a carrier gas at constant flow of 1 mL/min. PMAAs were eluted with the following linear gradient oven: initially the oven temperature was set at 60 °C for 1 min, heated to 300 °C at a rate of 8 °C per min, then held at 300 °C for 1 min.

Lectin Flow Cytometry and Metabolic Assays. Detailed protocols are provided in the Appendix.

126

Proteomics Analysis. Detailed protocols are provided in the Appendix.

Data Deposition

The RNA-Seq data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) Database, http://www.ncbi.nlm.nih.gov/geo, with accession no. GSE112589.

Funding

References

[1] Wong, M. Y., DiChiara, A. S., Suen, P. H., Chen, K., Doan, N. D., and Shoulders, M. D. (2017) Adapting Secretory Proteostasis and Function Through the Unfolded Protein Response, *Curr. Top. Microbiol. Immunol. 414*, 1-25.

[2] Calfon, M., Zeng, H., Urano, F., Till, J. H., Hubbard, S. R., Harding, H. P., Clark, S. G., and Ron, D. (2002) IRE1 couples endoplasmic reticulum load to secretory capacity by processing the XBP-1 mRNA, *Nature 415*, 92-96.

[3] Harding, H. P., Novoa, I., Zhang, Y., Zeng, H., Wek, R., Schapira, M., and Ron, D. (2000) Regulated translation initiation controls stress-induced gene expression in mammalian cells, *Mol. Cell 6*, 1099-1108.

[4] Vattem, K. M., and Wek, R. C. (2004) Reinitiation involving upstream ORFs regulates ATF4 mRNA translation in mammalian cells, *Proc. Natl. Acad. Sci. U.S.A. 101*, 11269-11274.

[5] Haze, K., Yoshida, H., Yanagi, H., Yura, T., and Mori, K. (1999) Mammalian transcription factor ATF6 is synthesized as a transmembrane protein and activated by proteolysis in response to endoplasmic reticulum stress, *Mol. Biol. Cell 10*, 3787-3799.

[6] Volmer, R., and Ron, D. (2015) Lipid-dependent regulation of the unfolded protein response, *Curr. Opin. Cell Biol. 33*, 67-73.

[7] Malhotra, J. D., and Kaufman, R. J. (2007) Endoplasmic reticulum stress and oxidative stress: a vicious cycle or a double-edged sword?, *Antioxid. Redox Signaling 9*, 2277-2293.

[8] Roy, C. R., Salcedo, S. P., and Gorvel, J. P. (2006) Pathogen-endoplasmic-reticulum interactions: in through the out door, *Nat. Rev. Immunol. 6*, 136-147.

[9] Carrara, M., Prischi, F., Nowak, P. R., Kopp, M. C., and Ali, M. M. (2015) Noncanonical binding of BiP ATPase domain to Ire1 and Perk is dissociated by unfolded protein CH1 to initiate ER stress signaling, *eLife 4*, e03522.

[10] Pincus, D., Chevalier, M. W., Aragón, T., van Anken, E., Vidal, S. E., El-Samad, H., and Walter, P. (2010) BiP Binding to the ER-Stress Sensor Ire1 Tunes the Homeostatic Behavior of the Unfolded Protein Response, *PLoS Biol. 8*, e1000415.

[11] Amin-Wetzel, N., Saunders, R. A., Kamphuis, M. J., Rato, C., Preissler, S., Harding, H. P., and Ron, D. (2017) A J-Protein Co-chaperone Recruits BiP to Monomerize IRE1 and Repress the Unfolded Protein Response, *Cell 171*, 1625-1637.

[12] Acosta-Alvear, D., Zhou, Y., Blais, A., Tsikitis, M., Lents, N. H., Arias, C., Lennon, C. J., Kluger, Y., and Dynlacht, B. D. (2007) XBP1 controls diverse cell type- and condition-specific transcriptional regulatory networks, *Mol. Cell 27*, 53-66.

[13] Kroeger, H., Grimsey, N., Paxman, R., Chiang, W. C., Plate, L., Jones, Y., Shaw, P. X., Trejo, J., Tsang, S. H., Powers, E., Kelly, J. W., Wiseman, R. L., and Lin, J. H. (2018) The unfolded protein response regulator ATF6 promotes mesodermal differentiation, *Sci. Signal. 11*, eaan5785.

[14] Shoulders, M. D., Ryno, L. M., Genereux, J. C., Moresco, J. J., Tu, P. G., Wu, C., Yates, J. R., 3rd, Su, A. I., Kelly, J. W., and Wiseman, R. L. (2013) Stress-independent activation of XBP1s and/or ATF6 reveals three functionally diverse ER proteostasis environments, *Cell Rep. 3*, 1279-1292.

[15] Mendez, A. S., Alfaro, J., Morales-Soto, M. A., Dar, A. C., McCullagh, E., Gotthardt, K., Li, H., Acosta-Alvear, D., Sidrauski, C., Korennykh, A. V., Bernales, S., Shokat, K. M., and Walter, P. (2015) Endoplasmic reticulum stress-independent activation of unfolded protein response kinases by a small molecule ATP-mimic, *eLife 4*, e05434.

[16] Plate, L., Cooley, C. B., Chen, J. J., Paxman, R. J., Gallagher, C. M., Madoux, F., Genereux, J. C., Dobbs, W., Garza, D., Spicer, T. P., Scampavia, L., Brown, S. J., Rosen, H., Powers, E. T., Walter, P., Hodder, P., Wiseman, R. L., and Kelly, J. W.

(2016) Small molecule proteostasis regulators that reprogram the ER to reduce extracellular protein aggregation, *eLife 5*, e15550.

[17] Walter, P., and Ron, D. (2011) The unfolded protein response: from stress pathway to homeostatic regulation, *Science 334*, 1081-1086.

[18] Margariti, A., Li, H., Chen, T., Martin, D., Vizcay-Barrena, G., Alam, S., Karamariti, E., Xiao, Q., Zampetaki, A., Zhang, Z., Wang, W., Jiang, Z., Gao, C., Ma, B., Chen, Y. G., Cockerill, G., Hu, Y., Xu, Q., and Zeng, L. (2013) XBP1 mRNA splicing triggers an autophagic response in endothelial cells through BECLIN-1 transcriptional activation, *J. Biol. Chem. 288*, 859-872.

[19] Thibault, G., Shui, G. H., Kim, W., McAlister, G. C., Ismail, N., Gygi, S. P., Wenk, M. R., and Ng, D. T. W. (2012) The Membrane Stress Response Buffers Lethal Effects of Lipid Disequilibrium by Reprogramming the Protein Homeostasis Network, *Mol. Cell 48*, 16-27.

[20] Shen, X., Ellis, R. E., Lee, K., Liu, C. Y., Yang, K., Solomon, A., Yoshida, H., Morimoto, R., Kurnit, D. M., Mori, K., and Kaufman, R. J. (2001) Complementary signaling pathways regulate the unfolded protein response and are required for C. elegans development, *Cell 107*, 893-903.

[21] Martinon, F., Chen, X., Lee, A. H., and Glimcher, L. H. (2010) TLR activation of the transcription factor XBP1 regulates innate immune responses in macrophages, *Nat. Immunol. 11*, 411-418.

[22] Martínez, G., Vidal, René L., Mardones, P., Serrano, Felipe G., Ardiles, Alvaro O., Wirth, C., Valdés, P., Thielen, P., Schneider, Bernard L., Kerr, B., Valdés, Jose L., Palacios, Adrian G., Inestrosa, Nibaldo C., Glimcher, Laurie H., and Hetz, C. (2016) Regulation of Memory Formation by the Transcription Factor XBP1, *Cell Rep. 14*, 1382-1394.

[23] Taylor, R. C., and Dillin, A. (2013) XBP-1 is a cell-nonautonomous regulator of stress resistance and longevity, *Cell 153*, 1435-1447.

[24] Cubillos-Ruiz, J. R., Silberman, P. C., Rutkowski, M. R., Chopra, S., Perales-Puchalt, A., Song, M., Zhang, S., Bettigole, S. E., Gupta, D., Holcomb, K., Ellenson, L. H., Caputo, T., Lee, A. H., Conejo-Garcia, J. R., and Glimcher, L. H. (2015) ER Stress Sensor XBP1 Controls Anti-tumor Immunity by Disrupting Dendritic Cell Homeostasis, *Cell 161*, 1527-1538.

[25] Chen, X., Iliopoulos, D., Zhang, Q., Tang, Q., Greenblatt, M. B., Hatziapostolou, M., Lim, E., Tam, W. L., Ni, M., Chen, Y., Mai, J., Shen, H., Hu, D. Z., Adoro, S., Hu, B., Song, M., Tan, C., Landis, M. D., Ferrari, M., Shin, S. J., Brown, M., Chang, J. C., Liu, X. S., and Glimcher, L. H. (2014) XBP1 promotes triple-negative breast cancer by controlling the HIF1alpha pathway, *Nature 508*, 103-107.

[26] Karagöz, G. E., Acosta-Alvear, D., Nguyen, H. T., Lee, C. P., Chu, F., and Walter, P. (2017) An unfolded protein-induced conformational switch activates mammalian IRE1, *eLife 6*, e30700.

[27] Ruiz-Canada, C., Kelleher, D. J., and Gilmore, R. (2009) Cotranslational and posttranslational N-glycosylation of polypeptides by distinct mammalian OST isoforms, *Cell 136*, 272-283.

[28] Kornfeld, R., and Kornfeld, S. (1985) Assembly of Asparagine-Linked Oligosaccharides, *Annu. Rev. Biochem. 54*, 631-664.

[29] Wang, Z. V., Deng, Y., Gao, N., Pedrozo, Z., Li, D. L., Morales, C. R., Criollo, A., Luo, X., Tan, W., Jiang, N., Lehrman, M. A., Rothermel, B. A., Lee, A. H., Lavandero, S., Mammen, P. P., Ferdous, A., Gillette, T. G., Scherer, P. E., and Hill, J. A. (2014) Spliced X-box binding protein 1 couples the unfolded protein response to hexosamine biosynthetic pathway, *Cell 156*, 1179-1192.

[30] Denzel, M. S., Storm, N. J., Gutschmidt, A., Baddi, R., Hinze, Y., Jarosch, E., Sommer, T., Hoppe, T., and Antebi, A. (2014) Hexosamine pathway metabolites enhance protein quality control and prolong life, *Cell 156*, 1167-1178.

[31] Thibault, G., Ismail, N., and Ng, D. T. W. (2011) The unfolded protein response supports cellular robustness as a broad-spectrum compensatory pathway, *Proc. Natl. Acad. Sci. U.S.A. 108*, 20597-20602.

[32] Moremen, K. W., Tiemeyer, M., and Nairn, A. V. (2012) Vertebrate protein glycosylation: diversity, synthesis and function, *Nat. Rev. Mol. Cell. Biol. 13*, 448-462.

[33] Kannagi, R., Izawa, M., Koike, T., Miyazaki, K., and Kimura, N. (2004) Carbohydrate-mediated cell adhesion in cancer metastasis and angiogenesis, *Cancer Sci. 95*, 377-384.

[34] Zandberg, W. F., Kumarasamy, J., Pinto, B. M., and Vocadlo, D. J. (2012) Metabolic inhibition of sialyl-Lewis X biosynthesis by 5-thiofucose remodels the cell surface and impairs selectin-mediated cell adhesion, *J. Biol. Chem. 287*, 40021-40030.

[35] Pang, P. C., Tissot, B., Drobnis, E. Z., Morris, H. R., Dell, A., and Clark, G. F. (2009) Analysis of the human seminal plasma glycome reveals the presence of immunomodulatory carbohydrate functional groups, *J. Proteome Res. 8*, 4906-4915.

[36] Varki, A. (2017) Biological roles of glycans, *Glycobiology 27*, 3-49.

[37] Dewal, M. B., DiChiara, A. S., Antonopoulos, A., Taylor, R. J., Harmon, C. J., Haslam, S. M., Dell, A., and Shoulders, M. D. (2015) XBP1s Links the Unfolded Protein Response to the Molecular Architecture of Mature N-Glycans, *Chem. Biol. 22*, 1301-1312.

[38] Ramanathan, A. A., Mehta, A., and Artlett, C. M. (2017) The Impact of the Unfolded Protein Response and the Hexosamine Biosynthetic Pathway on Glycosylation, *Glycobiology Insights 6*.

[39] Hanson, S. R., Culyba, E. K., Hsu, T.-L., Wong, C.-H., Kelly, J. W., and Powers, E. T. (2009) The core trisaccharide of an N-linked glycoprotein intrinsically accelerates folding and enhances stability, *Proc. Natl. Acad. Sci. U.S.A. 106*, 3131-3136.

[40] Lee, A.-H., Iwakoshi, N. N., and Glimcher, L. H. (2003) XBP-1 regulates a subset of endoplasmic reticulum resident chaperone genes in the unfolded protein response, *Mol. Cell. Biol. 23*, 7448-7459.

[41] Han, D., Lerner, A. G., Vande Walle, L., Upton, J. P., Xu, W., Hagen, A., Backes, B. J., Oakes, S. A., and Papa, F. R. (2009) IRE1alpha kinase activation modes control alternate endoribonuclease outputs to determine divergent cell fates, *Cell 138*, 562-575.

[42] Cooley, C. B., Ryno, L. M., Plate, L., Morgan, G. J., Hulleman, J. D., Kelly, J. W., and Wiseman, R. L. (2014) Unfolded protein response activation reduces secretion and extracellular aggregation of amyloidogenic immunoglobulin light chain, *Proc. Natl. Acad. Sci. U.S.A. 111*, 13046-13051.

[43] Genereux, J. C., Qu, S., Zhou, M., Ryno, L. M., Wang, S., Shoulders, M. D., Kaufman, R. J., Lasmezas, C. I., Kelly, J. W., and Wiseman, R. L. (2015) Unfolded protein response-induced ERdj3 secretion links ER stress to extracellular proteostasis, *EMBO J. 34*, 4-19.

[44] Sekiya, M., Maruko-Otake, A., Hearn, S., Sakakibara, Y., Fujisaki, N., Suzuki, E., Ando, K., and Iijima, K. M. (2017) EDEM Function in ERAD Protects against Chronic ER Proteinopathy and Age-Related Physiological Decline in Drosophila, *Dev. Cell 41*, 652-664.

[45] Lee, A. S. (2014) Glucose-regulated proteins in cancer: molecular mechanisms and therapeutic potential, *Nat. Rev. Cancer 14*, 263-276.

[46] Propheter, D. C., Hsu, K. L., and Mahal, L. K. (2011) Recombinant lectin microarrays for glycomic analysis, *Methods Mol. Biol. 723*, 67-77.

[47] Pilobello, K. T., Slawek, D. E., and Mahal, L. K. (2007) A ratiometric lectin microarray approach to analysis of the dynamic mammalian glycome, *Proc. Natl. Acad. Sci. U.S.A. 104*, 11534-11539.

130

[48] Ribeiro, J. P., Pau, W., Pifferi, C., Renaudet, O., Varrot, A., Mahal, L. K., and Imberty, A. (2016) Characterization of a high-affinity sialic acid-specific CBM40 from Clostridium perfringens and engineering of a divalent form, *Biochem. J. 473*, 2109-2118.

[49] Fujitani, N., Furukawa, J., Araki, K., Fujioka, T., Takegawa, Y., Piao, J., Nishioka, T., Tamura, T., Nikaido, T., Ito, M., Nakamura, Y., and Shinohara, Y. (2013) Total cellular glycomics allows characterizing cells and streamlining the discovery process for cellular biomarkers, *Proc. Natl. Acad. Sci. U.S.A. 110*, 2105-2110.

[50] North, S. J., Huang, H. H., Sundaram, S., Jang-Lee, J., Etienne, A. T., Trollope, A., Chalabi, S., Dell, A., Stanley, P., and Haslam, S. M. (2010) Glycomics profiling of Chinese hamster ovary cell glycosylation mutants reveals N-glycans of a novel size and complexity, *J. Biol. Chem. 285*, 5759-5775.

[51] Agrawal, P., Kurcon, T., Pilobello, K. T., Rakus, J. F., Koppolu, S., Liu, Z. Y., Batista, B. S., Eng, W. S., Hsu, K. L., Liang, Y. X., and Mahal, L. K. (2014) Mapping posttranscriptional regulation of the human glycome uncovers microRNA defining the glycocode, *Proc. Natl. Acad. Sci. U.S.A. 111*, 4338-4343.

[52] Mori, T., O'Keefe, B. R., Sowder, R. C., 2nd, Bringans, S., Gardella, R., Berg, S., Cochran, P., Turpin, J. A., Buckheit, R. W., Jr., McMahon, J. B., and Boyd, M. R. (2005) Isolation and characterization of griffithsin, a novel HIV-inactivating protein, from the red alga Griffithsia sp, *J. Biol. Chem. 280*, 9345-9353.

[53] Moriwaki, K., and Miyoshi, E. (2014) Basic Procedures for Lectin Flow Cytometry, In *Methods in Molecular Biology* (Hirabayashi, J., Ed.), pp 147-152, Humana Press, New York, NY.

[54] Kaku, H., Van Damme, E. J. M., Peumans, W. J., and Goldstein, I. J. (1990) Carbohydrate-binding specificity of the daffodil (Narcissus pseudonarcissus) and amaryllis (Hippeastrum hybr.) bulb lectins, *Arch. Biochem. Biophys. 279*, 298-304.

[55] Fuhrmann, U., Bause, E., Legler, G., and Ploegh, H. (1984) Novel mannosidase inhibitor blocking conversion of high mannose to complex oligosaccharides, *Nature 307*, 755-758.

[56] Mendoza, L., Olaso, E., Anasagasti, M. J., Fuentes, A. M., and Vidal-Vanaclocha, F. (1998) Mannose receptor-mediated endothelial cell activation contributes to B16 melanoma cell adhesion and metastasis in liver, *J. Cell. Physiol. 174*, 322-330.

[57] Monk, C. R., Sutton-Smith, M., Dell, A., and Garden, O. A. (2006) Letter to the Glyco-Forum: Preparation of CD25+ and CD25− CD4+ T cells for glycomic analysis—a cautionary tale of serum glycoprotein sequestration, *Glycobiology 16*, 11G-13G.

[58] Chen, Q., Muller, J. S., Pang, P. C., Laval, S. H., Haslam, S. M., Lochmuller, H., and Dell, A. (2015) Global N-linked Glycosylation is Not Significantly Impaired in Myoblasts in Congenital Myasthenic Syndromes Caused by Defective Glutamine-Fructose-6-Phosphate Transaminase 1 (GFPT1), *Biomolecules 5*, 2758-2781.

[59] Rillahan, C. D., Antonopoulos, A., Lefort, C. T., Sonon, R., Azadi, P., Ley, K., Dell, A., Haslam, S. M., and Paulson, J. C. (2012) Global metabolic inhibitors of sialyl- and fucosyltransferases remodel the glycome, *Nat. Chem. Biol. 8*, 661-668.

[60] Batista, B. S., Eng, W. S., Pilobello, K. T., Hendricks-Munoz, K. D., and Mahal, L. K. (2011) Identification of a conserved glycan signature for microvesicles, *J. Proteome Res. 10*, 4624-4633.

[61] Krishnamoorthy, L., Bess, J. W., Jr., Preston, A. B., Nagashima, K., and Mahal, L. K. (2009) HIV-1 and microvesicles from T cells share a common glycome, arguing for a common origin, *Nat. Chem. Biol. 5*, 244-250.

[62] Escrevente, C., Grammel, N., Kandzia, S., Zeiser, J., Tranfield, E. M., Conradt, H. S., and Costa, J. (2013) Sialoglycoproteins and N-glycans from secreted exosomes of ovarian carcinoma cells, *PLoS One 8*, e78631.

[63] Morrison, C. J., Easton, R. L., Morris, H. R., McMaster, W. R., Piret, J. M., and Dell, A. (2000) Modification of a recombinant GPI-anchored metalloproteinase for secretion alters the protein glycosylation, *Biotechnol. Bioeng. 68*, 407-421.

[64] Dong, L., Jiang, C. C., Thorne, R. F., Croft, A., Yang, F., Liu, H., de Bock, C. E., Hersey, P., and Zhang, X. D. (2011) Ets-1 mediates upregulation of Mcl-1 downstream of XBP-1 in human melanoma cells upon ER stress, *Oncogene 30*, 3716-3726.

[65] Chen, H. L., Li, C. F., Grigorian, A., Tian, W., and Demetriou, M. (2009) T cell receptor signaling co-regulates multiple Golgi genes to enhance N-glycan branching, *J. Biol. Chem. 284*, 32454-32461.

[66] Nairn, A. V., York, W. S., Harris, K., Hall, E. M., Pierce, J. M., and Moremen, K. W. (2008) Regulation of glycan structures in animal tissues: transcript profiling of glycan-related genes, *J. Biol. Chem. 283*, 17298-17313.

[67] Comelli, E. M., Head, S. R., Gilmartin, T., Whisenant, T., Haslam, S. M., North, S. J., Wong, N. K., Kudo, T., Narimatsu, H., Esko, J. D., Drickamer, K., Dell, A., and Paulson, J. C. (2006) A focused microarray approach to functional glycomics: transcriptional regulation of the glycome, *Glycobiology 16*, 117-131.

[68] Narasimhan, S. (1982) Control of glycoprotein synthesis. UDP-GlcNAc:glycopeptide beta 4-N-acetylglucosaminyltransferase III, an enzyme in hen oviduct which adds GlcNAc in beta 1-4 linkage to the beta-linked mannose of the trimannosyl core of N-glycosyl oligosaccharides, *J. Biol. Chem. 257*, 10235-10242.

[69] Slawson, C., Copeland, R. J., and Hart, G. W. (2010) O-GlcNAc signaling: a metabolic link between diabetes and cancer?, *Trends Biochem. Sci. 35*, 547-555.

[70] Dennis, J. W., Nabi, I. R., and Demetriou, M. (2009) Metabolism, cell surface organization, and disease, *Cell 139*, 1229-1241.

[71] Broschat, K. O., Gorka, C., Page, J. D., Martin-Berger, C. L., Davies, M. S., Huang Hc, H. C., Gulve, E. A., Salsgiver, W. J., and Kasten, T. P. (2002) Kinetic characterization of human glutamine-fructose-6-phosphate amidotransferase I: potent feedback inhibition by glucosamine 6-phosphate, *J. Biol. Chem. 277*, 14764-14770.

[72] Maszczak-Seneczko, D., Sosicka, P., Olczak, T., Jakimowicz, P., Majkowski, M., and Olczak, M. (2013) UDP-N-acetylglucosamine transporter (SLC35A3) regulates biosynthesis of highly branched N-glycans and keratan sulfate, *J. Biol. Chem. 288*, 21850-21860.

[73] Sasai, K., Ikeda, Y., Fujii, T., Tsuda, T., and Taniguchi, N. (2002) UDP-GlcNAc concentration is an important factor in the biosynthesis of beta1,6-branched oligosaccharides: regulation based on the kinetic properties of N-acetylglucosaminyltransferase V, *Glycobiology 12*, 119-127.

[74] Lau, K. S., Partridge, E. A., Grigorian, A., Silvescu, C. I., Reinhold, V. N., Demetriou, M., and Dennis, J. W. (2007) Complex N-Glycan Number and Degree of Branching Cooperate to Regulate Cell Proliferation and Differentiation, *Cell 129*, 123-134.

[75] Sassi, A., Lazaroski, S., Wu, G., Haslam, S. M., Fliegauf, M., Mellouli, F., Patiroglu, T., Unal, E., Ozdemir, M. A., Jouhadi, Z., Khadir, K., Ben-Khemis, L., Ben-Ali, M., Ben-Mustapha, I., Borchani, L., Pfeifer, D., Jakob, T., Khemiri, M., Asplund, A. C., Gustafsson, M. O., Lundin, K. E., Falk-Sorqvist, E., Moens, L. N., Gungor, H. E., Engelhardt, K. R., Dziadzio, M., Stauss, H., Fleckenstein, B., Meier, R., Prayitno, K., Maul-Pavicic, A., Schaffer, S., Rakhmanov, M., Henneke, P., Kraus, H., Eibel, H., Kolsch, U., Nadifi, S., Nilsson, M., Bejaoui, M., Schaffer, A. A., Smith, C. I., Dell, A., Barbouche, M. R., and Grimbacher, B. (2014) Hypomorphic homozygous mutations in phosphoglucomutase 3 (PGM3) impair immunity and increase serum IgE levels, *J. Allergy Clin. Immunol. 133*, 1410-1419.

[76] Itkonen, H. M., Engedal, N., Babaie, E., Luhr, M., Guldvik, I. J., Minner, S., Hohloch, J., Tsourlakis, M. C., Schlomm, T., and Mills, I. G. (2015) UAP1 is overexpressed in

prostate cancer and is protective against inhibitors of N-linked glycosylation, *Oncogene 34*, 3744-3750.

[77] Deng, Y., Wang, Z. V., Tao, C., Gao, N., Holland, W. L., Ferdous, A., Repa, J. J., Liang, G., Ye, J., Lehrman, M. A., Hill, J. A., Horton, J. D., and Scherer, P. E. (2013) The Xbp1s/GalE axis links ER stress to postprandial hepatic metabolism, *J. Clin. Invest. 123*, 455-468.

[78] Neelamegham, S., and Mahal, L. K. (2016) Multi-level regulation of cellular glycosylation: from genes to transcript to enzyme to structure, *Curr. Opin. Struct. Biol. 40*, 145-152.

[79] Nairn, A. V., Aoki, K., dela Rosa, M., Porterfield, M., Lim, J. M., Kulik, M., Pierce, J. M., Wells, L., Dalton, S., Tiemeyer, M., and Moremen, K. W. (2012) Regulation of glycan structures in murine embryonic stem cells: combined transcript profiling of glycan-related genes and glycan structural analysis, *J. Biol. Chem. 287*, 37835-37856.

[80] Shaffer, A. L., Shapiro-Shelef, M., Iwakoshi, N. N., Lee, A. H., Qian, S. B., Zhao, H., Yu, X., Yang, L., Tan, B. K., Rosenwald, A., Hurt, E. M., Petroulakis, E., Sonenberg, N., Yewdell, J. W., Calame, K., Glimcher, L. H., and Staudt, L. M. (2004) XBP1, downstream of Blimp-1, expands the secretory apparatus and other organelles, and increases protein synthesis in plasma cell differentiation, *Immunity 21*, 81-93.

[81] Ohtsubo, K., Chen, M. Z., Olefsky, J. M., and Marth, J. D. (2011) Pathway to diabetes through attenuation of pancreatic beta cell glycosylation and glucose transport, *Nat. Med. 17*, 1067-1075.

[82] Liang, Y., Eng, W. S., Colquhoun, D. R., Dinglasan, R. R., Graham, D. R., and Mahal, L. K. (2014) Complex N-linked glycans serve as a determinant for exosome/microvesicle cargo recruitment, *J. Biol. Chem. 289*, 32526-32537.

[83] Imanikia, S., Sheng, M., and Taylor, R. C. (2017) Cell Non-autonomous UPR(ER) Signaling, *Curr. Top. Microbiol. Immunol. 414*, 27-43.

[84] Wellen, K. E., Lu, C., Mancuso, A., Lemons, J. M., Ryczko, M., Dennis, J. W., Rabinowitz, J. D., Coller, H. A., and Thompson, C. B. (2010) The hexosamine biosynthetic pathway couples growth factor-induced glutamine uptake to glucose metabolism, *Genes Dev. 24*, 2784-2799.

[85] Gaziel-Sovran, A., Segura, M. F., Di Micco, R., Collins, M. K., Hanniford, D., Vega-Saenz de Miera, E., Rakus, J. F., Dankert, J. F., Shang, S., Kerbel, R. S., Bhardwaj, N., Shao, Y., Darvishian, F., Zavadil, J., Erlebacher, A., Mahal, L. K., Osman, I., and Hernando, E. (2011) miR-30b/30d regulation of GalNAc transferases enhances invasion and immunosuppression during metastasis, *Cancer Cell 20*, 104-118.

[86] Williams, K. W., Liu, T., Kong, X., Fukuda, M., Deng, Y., Berglund, E. D., Deng, Z., Gao, Y., Liu, T., Sohn, J. W., Jia, L., Fujikawa, T., Kohno, D., Scott, M. M., Lee, S., Lee, C. E., Sun, K., Chang, Y., Scherer, P. E., and Elmquist, J. K. (2014) Xbp1s in Pomc neurons connects ER stress with energy balance and glucose homeostasis, *Cell Metab. 20*, 471-482.

[87] Sun, J., Liu, Y., and Aballay, A. (2012) Organismal regulation of XBP-1-mediated unfolded protein response during development and immune activation, *EMBO Rep. 13*, 855-860.

[88] Mahadevan, N. R., Rodvold, J., Sepulveda, H., Rossi, S., Drew, A. F., and Zanetti, M. (2011) Transmission of endoplasmic reticulum stress and pro-inflammation from tumor cells to myeloid cells, *Proc. Natl. Acad. Sci. U.S.A. 108*, 6561-6566.

[89] Zhu, B., Zhang, Q., Pan, Y., Mace, E. M., York, B., Antoulas, A. C., Dacso, C. C., and O'Malley, B. W. (2017) A Cell-Autonomous Mammalian 12 hr Clock Coordinates Metabolic and Stress Rhythms, *Cell Metab. 25*, 1305-1319.

[90] Stowell, S. R., Ju, T., and Cummings, R. D. (2015) Protein glycosylation in cancer, *Annu. Rev. Pathol. 10*, 473-510.

[91] Hakomori, S. (1996) Tumor malignancy defined by aberrant glycosylation and sphingo(glyco)lipid metabolism, *Cancer Res. 56*, 5309-5318.

[92] de Leoz, M. L. A., Young, L. J. T., An, H. J., Kronewitter, S. R., Kim, J. H., Miyamoto, S., Borowsky, A. D., Chew, H. K., and Lebrilla, C. B. (2011) High-mannose glycans are elevated during breast cancer progression, *Mol. Cell. Proteomics 10*.

[93] Hauselmann, I., and Borsig, L. (2014) Altered tumor-cell glycosylation promotes metastasis, *Front. Oncol. 4*, 28.

[94] Pilobello, K. T., Agrawal, P., Rouse, R., and Mahal, L. K. (2013) Advances in lectin microarray technology: optimized protocols for piezoelectric print conditions, *Curr. Protoc. Chem. Biol. 5*, 1-23.

[95] Jang-Lee, J., North, S. J., Sutton-Smith, M., Goldberg, D., Panico, M., Morris, H., Haslam, S., and Dell, A. (2006) Glycomic profiling of cells and tissues by mass spectrometry: fingerprinting and sequencing methodologies, *Methods Enzymol. 415*, 59-86.

[96] Chen, C., Constantinou, A., Chester, K. A., Vyas, B., Canis, K., Haslam, S. M., Dell, A., Epenetos, A. A., and Deonarain, M. P. (2012) Glycoengineering approach to half-life extension of recombinant biotherapeutics, *Bioconjug. Chem. 23*, 1524-1533.

[97] Ceroni, A., Maass, K., Geyer, H., Geyer, R., Dell, A., and Haslam, S. M. (2008) GlycoWorkbench: a tool for the computer-assisted annotation of mass spectra of glycans, *J. Proteome Res. 7*, 1650-1659.

CHAPTER VI
Identification of a Functional Role for the Conserved *N*-Glycan
on the Collagen-I C-Terminal Propeptide

Summary

In the human body, collagen-I constitutes the primary component of bone, skin, and other organs. Collagen-I is synthesized in the endoplasmic reticulum (ER) with globular N- and C-terminal propeptides flanking the triple-helical domain; folding begins with the C-terminal propeptide (C-Pro), which initiates assembly and determines the stoichiometry of the trimeric procollagen molecule. A key feature of the C-Pro domain is a single conserved *N*-glycosylation sequon. What is the purpose of the collagen-I *N*-glycan? C-Pro domains are cleaved before collagen is incorporated into the extracellular matrix (ECM), making an extracellular role unlikely. *N*-glycans facilitate protein folding in the ER by providing access to specialized chaperone and quality control systems, but attempts to establish an intracellular role have so far been unsuccessful. However, we hypothesized that the collagen-I *N*-glycan is critical for instances in which collagen-I folding is challenged by ER stress or by destabilizing mutations. One such example is C-Pro mutations that delay collagen-I folding, resulting in retention of collagen molecules and diseases such as osteogenesis imperfecta (OI). We tested our hypothesis using the OI variant Colα1(I) C1299W and found that the *N*-glycan is essential for folding and secretion. Furthermore, we show that even wild type Colα1(I) requires the *N*-glycan under conditions of ER stress. Our results identify, for the first time, a functional role for the highly conserved *N*-glycan on the collagen-I C-Pro domain and demonstrate that a single *N*-glycan can effectively buffer protein folding against stress incurred during development, disease, or tissue repair.

Contributions

Introduction

The collagen protein family plays critical structural and dynamic roles in numerous organisms. In the human body, collagens provide mechanical strength to tissues and mediate processes such as development, wound healing, and cell-cell interactions.[1] Among the fibrillar collagens, collagen-I constitutes the primary component of bone, skin, and other organs. While the unique structural properties of the collagen-I triple helix facilitate its biological functions, they also present constraints for its folding and secretion. Production of collagen-I molecules involves a dedicated cohort of chaperones and enzymes that shepherd the protein through folding, post-translational modification, and triple helix assembly.[2] A crucial step in collagen-I biosynthesis is the folding of the globular C-terminal propeptide (C-Pro) domain, which is responsible for initiating collagen-I folding, determining chain register within the triple helix, and maintaining the appropriate 2:1 $Col\alpha1(I):Col\alpha2(I)$ stoichiometry.[3-5]

The C-Pro domain contains a highly conserved N-glycan on each chain of the collagen-I triple helix. Previous work has shown that other conserved features of the C-Pro domain, such as the network of cysteine residues that forms stabilizing inter- and intra-chain disulfides, are critical for C-Pro function and collagen-I stability.[6] These findings suggest that the single N-glycan in the C-Pro domain is also required for folding and/or function. Yet the *raison d'être* of the N-glycan has remained elusive: the extracellular protease bone morphogenic protein-1 (BMP-1) cleaves the C-terminal propeptide prior to matrix incorporation of collagen-I fibers, suggesting that the N-glycan cannot have a biological function in the ECM.[7] And while $Col\alpha1(I)$ chains with destabilizing C-Pro mutations can be targeted for ER-associated degradation,[8] an intracellular role for the N-glycan in this process, or in collagen-I folding in general,[9] has not been found.

Glycobiology, however, argues that the C-Pro N-glycan must exist for a reason. Besides functioning as extracellular signals, N-glycans influence folding and quality control in the endoplasmic reticulum (ER). Nascent N-glycoproteins with unfolded or misfolded regions bind to chaperones calnexin and calreticulin upon exposure of the $Glc_1Man_9GlcNAc_2$ sugar (where Glc = Glucose, Man = Mannose, and GlcNAc = N-Acetylglucosamine).[10] Removal and re-addition of the glucose residue by glucosidase II and UDP-Glc:glycoprotein glucosyltransferase, respectively, facilitate additional cycles of engagement with calnexin and calreticulin, whereas folding-competent glycoproteins are released for transport to the Golgi. Terminally misfolded proteins have not only the glucose residue, but also additional mannose units, removed; the action of ER α-mannosidase I triggers extraction of glycoproteins to the cytosol and subsequent degradation by the proteasome.

We hypothesized that the C-Pro N-glycan might be required for collagen-I folding in the context of ER stress or destabilizing mutations. The function of the N-glycan would thus only be detectable when collagen folding was challenged. To test this hypothesis, we generated a non-glycosylated C-Pro variant by genetic substitution of the native acceptor asparagine with glutamine. We then expressed glycosylated or non-glycosylated versions of a disease-associated $Col\alpha1(I)$ C-Pro domain in mammalian cells and assayed the secretion of these constructs. We find that the $Col\alpha1(I)$ C1299W mutant is highly dependent on the N-glycan, while the N-glycan is dispensable for wild type $Col\alpha1(I)$. ER stress, however, reveals an essential function for the N-glycan on wild type $Col\alpha1(I)$ as well. Our results provide the first evidence of a function for the N-glycan in collagen-I proteostasis, and motivate continued efforts to examine the consequences of this highly conserved post-translational modification.

Results

The Collagen-I N-Glycan is Highly Conserved Across Species. The *N*-glycosylation motif on Colα1(I) is located in the C-terminal propeptide (C-Pro) domain at residue 1365 (Figure 6.1A). Studies of the human collagen-III homotrimer revealed that the C-Pro domain adopts a flower-like structure, with α-helical coiled coils forming the stalk, inter- and intra-chain disulfide-bonds that form the Ca^{2+}-binding base, and antiparallel β-sheets that form the inner and outer faces of the petals.[11] Mapping the location of this asparagine onto the crystal structure of a Colα1(I) C-Pro homotrimer[5] shows that the *N*-glycan is positioned on the outer face of each petal, directed into the solvent (Figure 6.1C). Sequence alignments of the C-Pro domain show that the *N*-

A. Colα1(I) Structure

1–22   162   1219

- ■ Signal sequence
- ■ N-propeptide
- ■ Triple-helical domain
- ■ C-propeptide

B. Species Conservation of the Colα1(I) *N*-Glycan

Human	TFLRLMSTEASQ**NIT**YHCKNSVAYMDQQTGNLKKAL
Beaver	TFLRLMSTEASQ**NIT**YHCKNSVAYMDQQTGNLKKSL
Frog	TFLRLMATEASQ**NIT**YHCKNSVAYMDQATGNLKKAL
Coelecanth	TFLRLMSTEASQ**NIT**YHCKNSIAYMDQQAGNLKKSL
Shark	TFLRLMSSEATQ**NIT**YHCKNSIAYMDEEAGNLKKAI
Lamprey	TFLRLLSTDASQ**KIT**YHCKNSVAYLDSRAGNLKKAL
Tunicate	KILRYFSKFAVQ**KFK**FKCMGVVAAYDTQYNSYTSSL
Lancelet	QFLRLSSTGARQ**NIT**YQCKNSVAYYDAYNSNYKKAL
Sea Urchin	TFLRLLSTSARQ**TVT**YFCKNVQG-----------NP
Mollusk	KVLQFDSQAARQ**GVI**YHCLNSHTYGTRFTGSTEQAL
Red Fire Ant	--LQLLSKHANQ**NIT**YHCKNSVAYFDY------GNP
Hydra	RFLRLNSQFVRQ**NLT**YHCLNSHAH**NDS**-FGNNRPYV
Sponge	QFIRAASRHAVQ**SFT**YKCRNSAAAVIFRTQDN----

C.

N X S/T
$_{1365}$N I T Y$_{1368}$

D. Conservation of the *N*-Glycan across Human Fibrillar Collagens

Colα1(I)	TFLRLMSTEASQ**NIT**YHCKNSVAYMDQQTGNLKKAL	*1388*
Colα2(I)	AFMRLLANYASQ**NIT**YHCKNSIAYMDEETGNLKKAV	*1290*
Colα1(II)	TFLRLLSTEGSQ**NIT**YHCKNSIAYLDEAAGNLKKAL	*1411*
Colα1(III)	AFLRLLSSRASQ**NIT**YHCKNSIAYMDQASGNVKKAL	*1390*
Colα1(V)	TFLRLLSASAHQ**NVT**YHCYQSVAWQDAATGSYDKAL	*1764*
Colα2(V)	TFLRLLSKEASQ**NIT**YICKNSVGYMDDQAKNLKKAV	*1423*
Colα3(V)	NFLKLLSATARQ**NFT**YSCQNAAAWLDEATGDYSHSA	*1669*
Colα1(XI)	TFLKLLTASARQ**NFT**YHCHQSAAWYDVSSGSYDKAL	*1732*
Colα2(XI)	TFLRLLSVSAHQ**DVS**YPCSGAA---------RDGPL	*1664*
Colα1(XXIV)	NFLHLLSSEATH**IIT**IHCLNTPRWTSTQTS-GPGLP	*1630*
Colα1(XXVII)	NFLHLLSSEVTQ**HIT**IHCLNMTVWQEGTGQTPAKQA	*1786*
Colα1(I)	QGCNLDAIKVFC**NME**T-GETCVYPTQPSVAQKNWYI	*1314*
Colα2(XI)	QGCARDAFRVFC**NFT**AGGETCVTPRDDVT-------	*1620*
Colα1(XXIV)	LGCPSDAIEVFC**NFS**AGGQTCLPPVSVTKLEF----	*1597*
Colα1(XXVII)	LGCSSDTIEVSC**NFT**HGGQTCLKPITASKVEF----	*1742*

Figure 6.1: *The collagen-I N-glycan is highly conserved across species.* (A) Schematic of the Colα1(I) sequence. The *N*-glycosylation motif is located in the C-terminal propeptide (C-Pro) domain. A representative glycan structure is shown, using standard symbols from the Consortium for Functional Glycomics (CFG) nomenclature. (B) The Colα1(I) *N*-glycosylation motif, highlighted in yellow, is highly conserved across species. Sequences of Colα1(I) or relevant analogues were aligned using Constraint-based Multiple Alignment Tool (COBALT; https://www.ncbi.nlm.nih.gov/tools/cobalt/). (C) Crystal structure of the collagen-I C-Pro domain (PDB: 5K31), showing the location of the *N*-glycan in magenta. The asparagine was changed to a glutamine for crystallization. (D) The Colα1(I) *N*-glycosylation motif, highlighted in yellow, is highly conserved across human fibrillar collagens. Collagen chains Colα2(XI), Colα1(XXIV), and Colα1(XXVII) have an *N*-glycosylation sequon in a different location from that of Colα1(I).

N-Glycosylation of Non-Fibrillar Human Collagens

Gene	N-Glycosylated?	Number of Sites
COL4A1	✓	1
COL4A2	✓	1
COL6A1	✓	5
COL6A2	✓	6
COL6A3	✓	10
COL6A5	✓	4
COL6A6	✓	9
COL7A1	✓	3
COL8A1		0
COL8A2		0
COL9A1	✓	1
COL9A2		0
COL9A3	✓	1
COL10A1	✓	(1)
COL12A1	✓	5
COL13A1		0
COL14A1	✓	5
COL15A1	✓	6
COL16A1	✓	2
COL17A1	✓	1
COL18A1	✓	4
COL19A1		0
COL20A1	✓	1
COL21A1	✓	1
COL22A1		1
COL23A1		0
COL25A1		0
COL26A1	✓	(2)
COL28A1	✓	(4)

Table 6.1: *N-Glycosylation patterns are not conserved among non-fibrillar, human collagens.* Table of non-fibrillar human collagen genes, showing number of annotated *N*-glycosylation sites (UniProtKB). Sequences that contained *N*-glycosylation sequons but were not annotated as being *N*-glycosylated have the number of sites shown in parentheses.

glycosylation motif of Colα1(I) is highly conserved across species (Figure 6.1B). Remarkably, the *N*-glycan is also present in all human fibrillar collagens, although the *N*-glycosylation sequon is shifted earlier in Colα2(XI), Colα1(XXIV), and Colα1(XXVII) (Figure 6.1D). Non-fibrillar human collagens have varying numbers of *N*-glycosylation sites distributed across the protein, with many chains containing multiple sequons (Table 6.1).

The N-Glycan is Required for Secretion of Misfolding Collagen-I Variants. The preservation of the *N*-glycosylation motif across species and collagens suggests that the *N*-glycan is important for collagen folding and/or function. *N*-glycans have been shown to stabilize protein structure and promote folding on isolated domains,[12] and the collagen-I C-Pro domain must fold first before triple helix assembly can occur.[2] One approach to determine biological roles of *N*-glycans involves mutating asparagine residues in *N*-glycosylation sequons to glutamines. Previous studies on Colα1(I) using this method surprisingly found that removal of the *N*-glycan caused no change in collagen-I secretion or matrix deposition by mouse fibroblasts.[9] However, those studies only examined wild type Colα1(I), and so we wondered if the function of the *N*-glycan might be revealed by performing similar experiments with misfolding Colα1(I) variants.

We selected the disease-associated Colα1(I) C1299W variant for our studies. The mutation removes a cysteine residue from the disulfide bonding network,[11] and might be expected to impact chain selection.[6] In addition, Colα1(I) C1299W results in a relatively mild osteogenesis imperfecta phenotype, and so we reasoned that protective functions of the *N*-glycan would be revealed by exacerbated misfolding of Colα1(I) C1299W upon genetic glycan removal.[13] We elected to perform initial studies on the isolated C-Pro domain, which we have shown can be expressed in various cell types in native, trimeric form.[14] We therefore generated *N*-terminal HA-tagged constructs of Colα1(I) wild type, Colα1(I) N1365Q, Colα1(I) C1299W, and Colα1(I) C1299W/N1365Q C-Pro domains, and expressed the constructs in HEK293 cells.

Immunoblotting analysis of conditioned media showed that removal of the *N*-glycan does not prevent secretion of Colα1(I) wild type C-Pro, consistent with previous reports;[9] trimerization of Colα1(I) N1365Q also did not appear to be affected (Figure 6.2A). In contrast, Colα1(I)
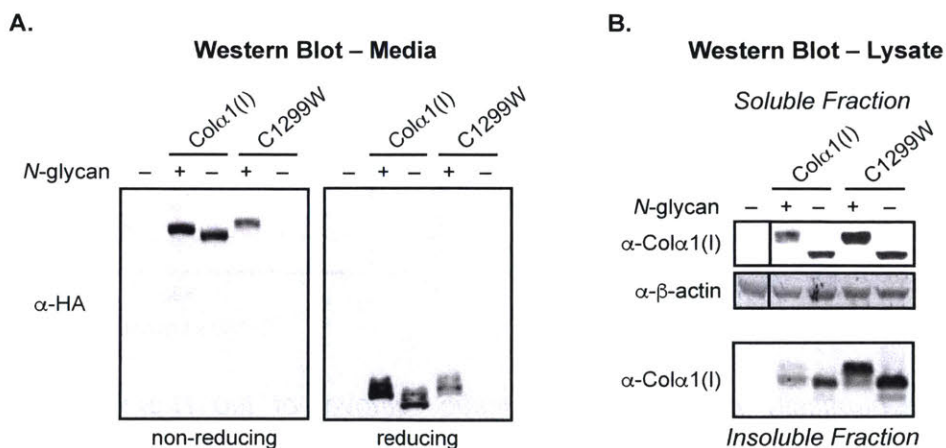
Figure 6.2: The N-glycan is required for secretion of misfolding collagen-I variants. (A) Western blot of Colα1(I) C-Pro domains expressed in and secreted from HEK293 cells. Cells were plated and allowed to adhere overnight before transfection with the indicated Colα1(I) C-Pro constructs and Lipofectamine3000. Secretion of Colα1(I) C1299W C-Pro is prevented by removal of the N-glycan. (B) Western blot of soluble and insoluble intracellular fractions following expression of Colα1(I) C-Pro domains in HEK293 cells.

C1299W C-Pro was no longer secreted upon removal of the N-glycan. Immunoblotting analysis of the intracellular fraction revealed that the loss of Colα1(I) C1299W/N1365Q secretion was not due to retention of the double mutant (Figure 6.2B). While Colα1(I) N1365Q appeared more prone to aggregate than the Colα1(I) wild type construct, Colα1(I) C1299W/N1365Q, surprisingly, did not show an increased tendency to aggregate relative to Colα1(I) C1299W (Figure 6.2B).

Chemically Phenocopying Genetic Removal of the N-Glycan Sequon. In the absence of increased aggregation or intracellular retention, our results suggested that removal of the N-glycan might direct the Colα1(I) C1299W C-Pro towards proteasomal or autophagasomal degradation. Cellular quality control of N-glycoproteins in the ER predominantly involves two chaperones, calnexin and calreticulin.[10] These glycoprotein-binding lectins capture and release partially folded or misfolded N-glycoproteins displaying a $Glc_1Man_9GlcNAc_2$ sugar; substrates that fail to fold are subsequently targeted for extraction from the ER and degradation by the proteasome.

We examined the effects of preventing Colα1(I) C1299W entry into the calnexin/calreticulin cycle by treating cells with the α-glucosidase inhibitor castanospermine.[15-17] Treating cells with castanospermine did not alter secretion of Colα1(I) wild type or Colα1(I) N1365Q (Figure 6.3A). Similar to the wild type, the secretion of Colα1(I) C1299W/N1365Q was also not affected by treatment with castanospermine; however, all tested concentrations of inhibitor markedly reduced the secretion of Colα1(I) C1299W (Figure 6.3B), showing that chemically preventing access to the calnexin/calreticulin cycle mimics genetic removal of the N-glycan on a misfolding Colα1(I) variant.
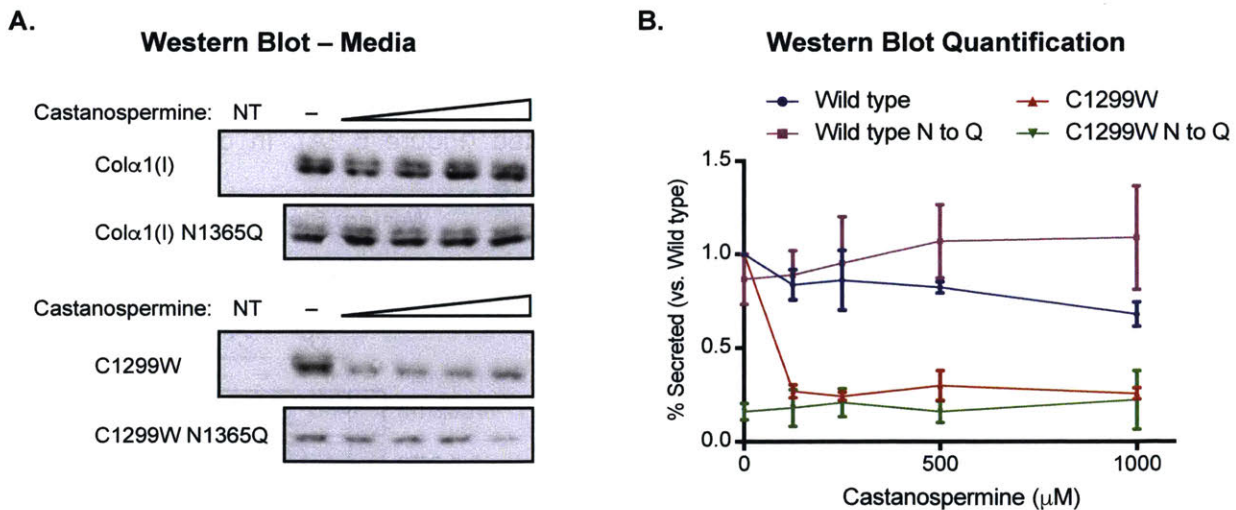
Figure 6.3: *Chemically phenocopying genetic removal of the N-glycan sequon.* (A) Representative Western blot of Colα1(I) C-Pro domains expressed in and secreted from HEK293 cells treated with increasing concentrations of castanospermine. HEK293 cells were plated and allowed to adhere overnight. The next day, cells were transfected with the indicated Colα1(I) C-Pro constructs using Lipofectamine3000. The following day, cells received fresh media and either DMSO or increasing concentrations of castanospermine: 125, 250, 500, or 1000 μM. Media was collected for 6 h before harvesting for immunoblotting. NT = no transfection. (B) Quantification of Western blots shown in (A), normalized to the amount of Colα1(I) C-Pro secreted from DMSO-treated cells.

The N-Glycan Protects Colα1(I) Against Stress-Induced Aggregation. These data showed that the *N*-glycan is essential when destabilizing amino acid substitutions induce collagen-I misfolding. Wild type collagen, however, folds well. Given that the *N*-glycan is highly conserved, we wondered if there might be cases where the *N*-glycan is required for wild type Colα1(I) to fold. Collagen-producing tissues are known to experience physiological ER stress induced by the high secretory output of ECM proteins that accompanies development and remodeling processes.[18-20] We therefore hypothesized that the *N*-glycan might protect wild type Colα1(I) from misfolding under challenging conditions of physiological ER stress.

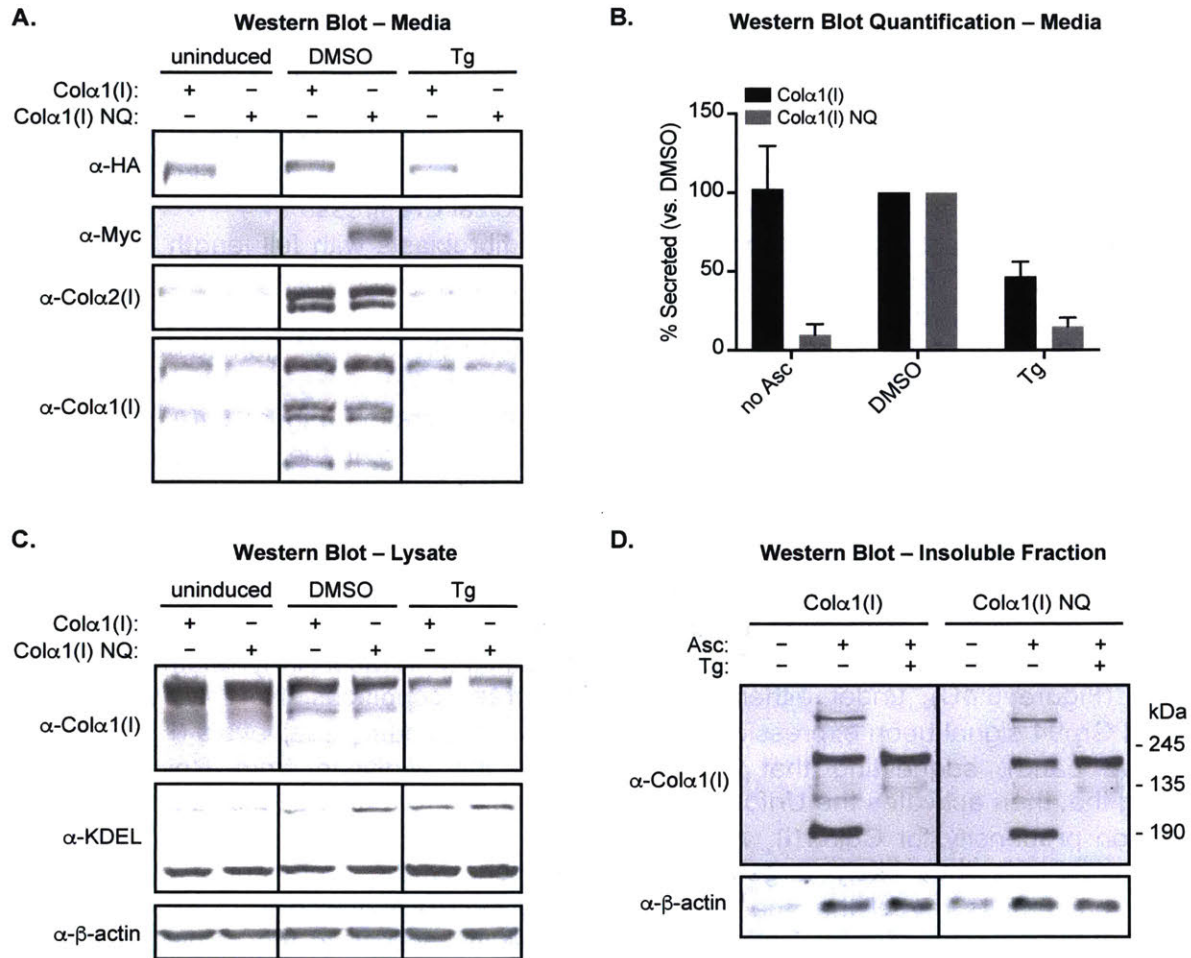To test our hypothesis, we transduced primary fibroblasts with full length Colα1(I) wild type or Colα1(I) N1365Q adenoviruses, induced endogenous collagen-I secretion with ascorbate, and assayed Colα1(I) wild type or Colα1(I) N1365Q behavior in the absence (DMSO) or presence of ER stress (induced by treatment with the Ca^{2+} homeostasis disruptor Thapsigargin; Tg). While full length Colα1(I) N1365Q was secreted robustly by vehicle-treated cells, inducing ER stress with Tg strongly reduced mutant secretion (Figure 6.4A); secretion of wild type Colα1(I), by contrast, was not significantly affected by Tg. In the absence of ascorbate, which is required to facilitate proper hydroxylation of the collagen-I triple helix by prolyl-4-hydroxylase,[21, 22] Colα1(I) N1365Q was also secreted much less efficiently than under ascorbate-supplemented conditions (Figures 6.4A and 6.4B).

We next examined whether the reduced secretion of Colα1(I) N1365Q might be due to intracellular retention of misfolding chains. Consistent with the results obtained using the isolated C-Pro domain, Colα1(I) N1365Q did not accumulate intracellularly relative to Colα1(I) wild type (Figure 6.4C), under either basal or ER stress conditions. Notably, we observed increased Grp94 signal upon expression of Colα1(I) N1365Q under basal levels (Figure 6.4C α-KDEL, top band), suggesting that simply removing the *N*-glycan from Colα1(I) triggers misfolding that then activates the Unfolded Protein Response. We did not observe differences in aggregation propensity for Colα1(I) wild type or Colα1(I) N1365Q (Figure 6.4D); the Colα1(I) N1365Q protein is thus likely targeted for degradation by the proteasome or autophagy. Together, our results reveal that the function of the Colα1(I) *N*-glycan is to protect collagen chains from stress-induced aggregation, likely by providing access to the calnexin/calreticulin chaperone cycle.

Discussion

A specific *N*-glycan site on the transmembrane glycoprotein CD147 was found to facilitate access to calnexin-mediated folding; genetic removal of the acceptor asparagine caused BiP and Hrd1-dependent routing of CD147 N152Q to ER-associated degradation (ERAD).[23] We have shown here that the single *N*-glycan on the Colα1(I) C-Pro domain is similarly needed to allow access to the calnexin/calreticulin chaperone cycle. Remarkably, the importance of the *N*-glycan for wild type Colα1(I) is not apparent under basal conditions, but becomes detectable when folding is challenged, whether by hydroxylation deficient conditions, chemically-induced ER stress, or by the presence of a disease-associated mutation. We do not observe significant increases in Colα1(I) aggregation, suggesting that when access to calnexin/calreticulin is prevented, the protein is directed to cellular quality control pathways. Patient cell lines expressing mutants in the Colα1(I) C-Pro domain that prevent trimer formation and interchain disulfide formation show elevated BiP and Grp94 levels; BiP was also found to stably associate with Colα1(I) in that same study.[24] Whether BiP, Grp94, or other chaperones are involved in quality control of misfolding, non-glycosylated Colα1(I) is the subject of ongoing work.

A. Western Blot – Media

B. Western Blot Quantification – Media

C. Western Blot – Lysate

D. Western Blot – Insoluble Fraction

Figure 6.4: *The N-glycan protects Colα1(I) against stress-induced aggregation.* (A) Western blot of conditioned media from primary fibroblasts expressing full length Colα1(I) with or without the *N*-glycan. Cells were plated and allowed to adhere overnight before being transduced with HA.Colα1(I) wild type or cMyc.Colα1(I) N1365Q adenoviruses. The next day, cells were split 1:4 into 12-well plates. The following day, cells received fresh media, endogenous collagen-I expression was induced with ascorbate, and cells received DMSO or 750 nM Tg. Conditioned media was collected for 24 h before harvesting for immunoblotting. (B) Quantification of secreted adenoviral Colα1(I) constructs, normalized to levels of Colα1(I) collected from DMSO and ascorbate-treated cells. Note that removal of the *N*-glycan reduces secretion of wild type Colα1(I) under both hydroxylation-deficient (no ascorbate) and ER stress (Tg treatment) conditions. (C) Western blot of soluble lysate from cells treated as described in (A). (D) Western blot of the insoluble intracellular fraction from cells treated as described in (A). Insoluble pellets were washed with PBS and then incubated in SDS solubilization buffer to dissolve proteins prior to immunoblotting analysis.

We previously found that stress-independent activation of the XBP1s branch of the UPR shifts the distribution of mature *N*-glycoforms on the Colα1(I) C-Pro domain in osteosarcoma cells.[14] Furthermore, we have shown that XBP1s activation also alters the structures of mature *N*-glycans on endogenous membrane-associated and secreted glycoproteins (e.g., Chapter V). As XBP1s activation has been linked to cartilage development, it will be interesting to determine if the structures of the Colα1(I) and Colα2(I) *N*-glycans similarly shift in a differentiation and/or tissue-dependent manner. These findings also raise the possibility that the *N*-glycan might have functions beyond promoting folding of the C-Pro domain. A number of collagen types (IV, VIII, XV, XVIII, and XIX) contain domains that can be cleaved by extracellular proteases to release bioactive signaling molecules, or matricryptins.[7] While an analogous process has not been identified for collagen-I, the C-Pro domain is known to persist in serum,[25] where it could ostensibly interact with glycan-binding components of the innate immune system such as galectins.[26] Moreover, *O*-linked glycosylation of collagen-IV has been shown to mediate endocytic uptake of collagen by the uPARAP/Endo180 mannose receptor,[27] and so perhaps the *N*-glycan on the C-Pro domain mediates similar interactions with other cell surface proteins.

In summary, we demonstrate that the Colα1(I) *N*-glycan is necessary to rescue misfolding chains by allowing the protein to access the calnexin/calreticulin chaperone cycle. This work expands the mechanisms that govern collagen-I secretion (e.g., Chapters II and III), and points to possible connections between stress-associated changes in *N*-glycan maturation and Colα1(I) C-Pro signaling (e.g., Chapter V). The agreement between our results with full length Colα1(I) and the isolated Colα1(I) C-Pro domain validates our transient expression system as a means to examine *N*-glycan function in the context of diverse collagen-I variants. In addition to testing other disease-associated Colα1(I) C-Pro substitutions, engineered changes to the *N*-glycan itself may also prove informative for identifying glycan- versus sequence (or conformation)-specific determinants of chaperone-client binding—through installation of additional *N*-glycan sequons, relocation of the *N*-glycan to another solvent-exposed region of the C-Pro domain, or addition of chemically-induced constraints on *N*-glycan maturation.[28-30] We anticipate that future studies will provide further insight into collagen-I proteostasis, as well as general principles of *N*-glycan dependent quality control in the ER that may well extend to other proteins in the secretory pathway.[31, 32]

Materials and Methods

Cell Lines and Reagents. HEK293 cells were grown in complete DMEM (Corning) supplemented with 10% fetal bovine serum (FBS), 100 IU penicillin, 100 μg/mL streptomycin, and 2 mM L-glutamine (Corning). Healthy dermal fibroblasts (GM05294; Coriell Cell Repositories) were cultured in complete MEM supplemented with 15% FBS, 100 IU penicillin/100 μg/mL streptomycin, and 2 mM L-glutamine (Corning). Secretion of endogenous collagen-I was induced in primary fibroblasts using 200 μM sodium ascorbate (Amresco). Immunoblots were probed with the following primary antibodies, as indicated: HA (ThermoFisher 26183), Colα2(I) (Abcam ab96723), Colα1(I) (NIH LF-68), Colα1(I) (Sigma HPA008405), β-actin (Sigma A1978), FLAG (Agilent 131 200473), c-Myc (Sigma M4439).

Vector Construction. Site-directed mutagenesis was performed on a previously described PPT.HA.Colα1(I).pENTR1A plasmid, using the QuikChange II XL Site-Directed Mutagenesis Kit (Agilent).[33] Collagen pENTR1A vectors were then recombined into pAd.CMV.V5.DEST Gateway destination vectors[34] using LR clonase (Life Technologies). The following primers were used to generate the C1299W [Forward primer: 5'-CTGGTGAGACCTGGGTGTACCCCACTC-3'; Reverse primer: 5'- GAGTGGGGTACACCCAGGTCTCACCAG-3'] or the N1365Q [Forward Primer: 5'-CCGAGGCCTCCCAGCAGATCACCTACCA-3'; Reverse Primer: 5'-CAGTGGTAGGTGATCTGCTGGGAGGCCT-3'] variants.

143

Adenoviral Production. Adenoviral vectors encoding HA.Colα1(I), HA.Colα1(I)C1299W, Myc.Colα1(I)N1365Q, or Myc.Colα1(I)C1299W/N1365Q were used to produce adenovirus according to the manufacturer's instructions (LifeTechnologies ViraPower Adenoviral Expression System). Viruses were subsequently amplified in 293A cells and titrated with GM05294 cells to select an appropriate volume to use for experiments.

Adenoviral Transduction. GM05294 fibroblasts were plated at a density of 3.5×10^5 cells/plate in 60-mm dishes the day before transduction. After being allowed to adhere overnight, cells received fresh media and were transduced with adenoviruses encoding HA.Colα1(I) and c-Myc.Colα1(I)N1365Q, as indicated. The next day, cells received fresh media. Two days post-transduction, cells were split 1:4 into 12-well plates. Collagen-I expression was induced with ascorbate the following day as described for Western blotting, and cells were treated with DMSO or 750 nM thapsigargin (Tg). After 24 h, the media was collected, separated on 4/8% Tris-Gly SDS–PAGE gels, and probed for HA or c-Myc.

Western Blotting Analysis and SDS Solubilization. Samples of medium were denatured by being boiled in Laemmli buffer (supplemented with 167 mM 1,4-dithiothreitol; DTT), separated on homemade 4/8% Tris-Gly SDS–PAGE gels, and transferred to nitrocellulose membranes. Non-reduced samples were boiled in Laemmli buffer without DTT prior to SDS-PAGE. Because of cross-reactivity between antibodies, samples of medium for collagen-I blots were run on duplicate gels, transferred, and then probed separately for either mouse anti-HA or mouse anti-c-Myc.

For Western blots of cell lysates, cells were trypsinized, washed with phosphate-buffered saline (PBS), and lysed in Triton-X lysis buffer with 1.5 mM phenylmethylsulfonyl fluoride (PMSF) and a protease inhibitor tablet (Pierce). Lysates were then cleared by centrifugation at 21.1×10^3 g and 4 °C for 15 min. Total protein in the supernatant was quantified using the bicinchoninic acid assay, and equal amounts of total protein were analyzed for each sample. Each experiment was performed in biological triplicate. Blots were imaged after incubation with appropriate primary and 800CW or 680LT secondary antibodies (LI-COR) by scanning on an Odyssey infrared imager (LI-COR), followed by quantification using ImageJ.

For SDS solubilization, insoluble pellets obtained after the centrifugation step of cell lysis (above) were washed in 1 mL of PBS with vortexing, and then spun at 21.1×10^3 g for 15 min. Pellets were then incubated with 4 volumes of SDS solubilization buffer (1% SDS, 50 mM Tris, 150 mM NaCl, 10 mM $MgCl_2$, pH 7.5), vortexing, at room temperature for 3 h. Pellets were removed from the vortexer and incubated at room temperature overnight. The next day, samples were heated with Laemmli buffer (supplemented with 167 mM DTT), separated on homemade 4/8% Tris-Gly SDS–PAGE gels, and transferred to nitrocellulose membranes.

Castanospermine Treatment. HEK293T cells were plated at a density of 7.5×10^5 cells/well in six-well plates and allowed to adhere overnight. The next day, cells were transfected with the indicated Colα1(I) C-Pro constructs using 1 μg of DNA/well and Lipofectamine3000. The next day, cells were split 1:4 and allowed to adhere overnight. The following day, cells received fresh media and either DMSO or increasing concentrations of castanospermine: 125, 250, 500, or 1000 μM. Media was collected for 6 h before harvesting for immunoblotting.

Funding

References

[1] Gelse, K., Pöschl, E., and Aigner, T. (2003) Collagens—structure, function, and biosynthesis, *Adv. Drug Delivery Rev. 55*, 1531-1546.

[2] Ishikawa, Y., and Bachinger, H. P. (2013) A molecular ensemble in the rER for procollagen maturation, *Biochim. Biophys. Acta 1833*, 2479-2491.

[3] Alvares, K., Siddiqui, F., Malone, J., and Veis, A. (1999) Assembly of the type 1 procollagen molecule: selectivity of the interactions between the alpha 1(I)- and alpha 2(I)-carboxyl propeptides, *Biochemistry 38*, 5401-5411.

[4] Malone, J. P., Alvares, K., and Veis, A. (2005) Structure and assembly of the heterotrimeric and homotrimeric C-propeptides of type I collagen: significance of the alpha2(I) chain, *Biochemistry 44*, 15269-15279.

[5] Sharma, U., Carrique, L., Vadon-Le Goff, S., Mariano, N., Georges, R.-N., Delolme, F., Koivunen, P., Myllyharju, J., Moali, C., Aghajari, N., and Hulmes, D. J. S. (2017) Structural basis of homo- and heterotrimerization of collagen I, *Nat. Commun. 8*, 14671.

[6] DiChiara, A. S., Li, R. C., Suen, P. H., Hosseini, A. S., Taylor, R. J., Weickhardt, A. F., Malhotra, D., McCaslin, D. R., and Shoulders, M. D. (2018) A Cysteine-Based Molecular Code Informs Collagen C-Propeptide Assembly, *Nat. Commun. 9*, In Press.

[7] Ricard-Blum, S. (2011) The collagen family, *Cold Spring Harbor Perspect. Biol. 3*, a004978.

[8] Fitzgerald, J., Lamandé, S. R., and Bateman, J. F. (1999) Proteasomal Degradation of Unassembled Mutant Type I Collagen Pro-α1(I) Chains, *J. Biol. Chem. 274*, 27392-27398.

[9] Lamandé, S. R., and Bateman, J. F. (1995) The Type I Collagen pro1(I) COOH-terminal Propeptide N-Linked Oligosaccharide: Functional Analysis by Site-Directed Mutagenesis, *J. Biol. Chem. 270*, 17858-17865.

[10] Caramelo, J. J., and Parodi, A. J. (2008) Getting in and out from calnexin/calreticulin cycles, *J. Biol. Chem. 283*, 10221-10225.

[11] Bourhis, J. M., Mariano, N., Zhao, Y., Harlos, K., Exposito, J. Y., Jones, E. Y., Moali, C., Aghajari, N., and Hulmes, D. J. (2012) Structural basis of fibrillar collagen trimerization and related genetic disorders, *Nat. Struct. Mol. Biol. 19*, 1031-1036.

[12] Hanson, S. R., Culyba, E. K., Hsu, T.-L., Wong, C.-H., Kelly, J. W., and Powers, E. T. (2009) The core trisaccharide of an N-linked glycoprotein intrinsically accelerates folding and enhances stability, *Proc. Natl. Acad. Sci. U.S.A. 106*, 3131-3136.

[13] Pace, J. M., Kuslich, C. D., Willing, M. C., and Byers, P. H. (2001) Disruption of one intra-chain disulphide bond in the carboxyl-terminal propeptide of the proα1(I) chain of type I procollagen permits slow assembly and secretion of overmodified, but stable procollagen trimers and results in mild osteogenesis imperfecta, *J. Med. Genet. 38*, 443-449.

[14] Dewal, M. B., DiChiara, A. S., Antonopoulos, A., Taylor, R. J., Harmon, C. J., Haslam, S. M., Dell, A., and Shoulders, M. D. (2015) XBP1s Links the Unfolded Protein Response to the Molecular Architecture of Mature N-Glycans, *Chem. Biol. 22*, 1301-1312.

[15] Elbein, A. D. (1991) Glycosidase inhibitors: inhibitors of N-linked oligosaccharide processing, *FASEB J. 5*, 3055-3063.

[16] Keller, S. H., Lindstrom, J., and Taylor, P. (1998) Inhibition of Glucose Trimming with Castanospermine Reduces Calnexin Association and Promotes Proteasome Degradation of the α-Subunit of the Nicotinic Acetylcholine Receptor, *J. Biol. Chem. 273*, 17064-17072.

[17] Saul, R., Molyneux, R. J., and Elbein, A. D. (1984) Studies on the mechanism of castanospermine inhibition of alpha- and beta-glucosidases, *Arch. Biochem. Biophys. 230*, 668-675.

[18] Saito, A., Ochiai, K., Kondo, S., Tsumagari, K., Murakami, T., Cavener, D. R., and Imaizumi, K. (2011) Endoplasmic reticulum stress response mediated by the PERK-

eIF2(alpha)-ATF4 pathway is involved in osteoblast differentiation induced by BMP2, *J. Biol. Chem. 286*, 4809-4818.

[19] Hisanaga, S., Miyake, M., Taniuchi, S., Oyadomari, M., Morimoto, M., Sato, R., Hirose, J., Mizuta, H., and Oyadomari, S. (2018) PERK-mediated translational control is required for collagen secretion in chondrocytes, *Sci. Rep. 8*, 773.

[20] Cameron, T. L., Gresshoff, I. L., Bell, K. M., Pirog, K. A., Sampurno, L., Hartley, C. L., Sanford, E. M., Wilson, R., Ermann, J., Boot-Handford, R. P., Glimcher, L. H., Briggs, M. D., and Bateman, J. F. (2015) Cartilage-specific ablation of XBP1 signaling in mouse results in a chondrodysplasia characterized by reduced chondrocyte proliferation and delayed cartilage maturation and mineralization, *Osteoarthritis Cartilage 23*, 661-670.

[21] Myllyla, R., Majamaa, K., Gunzler, V., Hanauske-Abel, H. M., and Kivirikko, K. I. (1984) Ascorbate is consumed stoichiometrically in the uncoupled reactions catalyzed by prolyl 4-hydroxylase and lysyl hydroxylase, *J. Biol. Chem. 259*, 5403-5405.

[22] de Jong, L., Albracht, S. P., and Kemp, A. (1982) Prolyl 4-hydroxylase activity in relation to the oxidation state of enzyme-bound iron. The role of ascorbate in peptidyl proline hydroxylation, *Biochim. Biophys. Acta 704*, 326-332.

[23] Li, J.-H., Huang, W., Lin, P., Wu, B., Fu, Z.-G., Shen, H.-M., Jing, L., Liu, Z.-Y., Zhou, Y., Meng, Y., Xu, B.-Q., Chen, Z.-N., and Jiang, J.-L. (2016) N-linked glycosylation at Asn152 on CD147 affects protein folding and stability: promoting tumour metastasis in hepatocellular carcinoma, *Sci. Rep. 6*, 35210.

[24] Chessler, S. D., and Byers, P. H. (1993) BiP binds type I procollagen pro alpha chains with mutations in the carboxyl-terminal propeptide synthesized by cells from patients with osteogenesis imperfecta, *J. Biol. Chem. 268*, 18226-18233.

[25] Parfitt, A. M., Simon, L. S., Villanueva, A. R., and Krane, S. M. (1987) Procollagen type I carboxy-terminal extension peptide in serum as a marker of collagen biosynthesis in bone. Correlation with Iliac bone formation rates and comparison with total alkaline phosphatase, *J. Bone Miner. Res. 2*, 427-436.

[26] Laderach, D. J., Compagno, D., Toscano, M. A., Croci, D. O., Dergan-Dylon, S., Salatino, M., and Rabinovich, G. A. (2010) Dissecting the signal transduction pathways triggered by galectin-glycan interactions in physiological and pathological settings, *IUBMB Life 62*, 1-13.

[27] Jurgensen, H. J., Madsen, D. H., Ingvarsen, S., Melander, M. C., Gardsvoll, H., Patthy, L., Engelholm, L. H., and Behrendt, N. (2011) A novel functional role of collagen glycosylation: interaction with the endocytic collagen receptor uparap/ENDO180, *J. Biol. Chem. 286*, 32736-32748.

[28] Fuhrmann, U., Bause, E., Legler, G., and Ploegh, H. (1984) Novel mannosidase inhibitor blocking conversion of high mannose to complex oligosaccharides, *Nature 307*, 755-758.

[29] Avezov, E., Frenkel, Z., Ehrlich, M., Herscovics, A., and Lederkremer, G. Z. (2008) Endoplasmic reticulum (ER) mannosidase I is compartmentalized and required for N-glycan trimming to Man5-6GlcNAc2 in glycoprotein ER-associated degradation, *Mol. Biol. Cell 19*, 216-225.

[30] Price, J. L., Powers, D. L., Powers, E. T., and Kelly, J. W. (2011) Glycosylation of the enhanced aromatic sequon is similarly stabilizing in three distinct reverse turn contexts, *Proc. Natl. Acad. Sci. U.S.A. 108*, 14127-14132.

[31] Yim, S. H., Everley, R. A., Schildberg, F. A., Lee, S. G., Orsi, A., Barbati, Z. R., Karatepe, K., Fomenko, D. E., Tsuji, P. A., Luo, H. R., Gygi, S. P., Sitia, R., Sharpe, A. H., Hatfield, D. L., and Gladyshev, V. N. (2018) Role of Selenof as a Gatekeeper of Secreted Disulfide-Rich Glycoproteins, *Cell Rep. 23*, 1387-1398.

[32] Labunskyy, V. M., Yoo, M. H., Hatfield, D. L., and Gladyshev, V. N. (2009) Sep15, a thioredoxin-like selenoprotein, is involved in the unfolded protein response and differentially regulated by adaptive and acute ER stresses, *Biochemistry 48*, 8458-8465.

[33] DiChiara, A. S., Taylor, R. J., Wong, M. Y., Doan, N. D., Rosario, A. M., and Shoulders, M. D. (2016) Mapping and exploring the collagen-I proteostasis network, *ACS Chem. Biol. 11*, 1408-1421.

[34] Campeau, E., Ruhl, V. E., Rodier, F., Smith, C. L., Rahmberg, B. L., Fuss, J. O., Campisi, J., Yaswen, P., Cooper, P. K., and Kaufman, P. D. (2009) A versatile viral system for expression and depletion of proteins in mammalian cells, *PLoS One 4*, e0006529.

In the course of research on protein folding, it can be easy to think that proteostasis is chiefly concerned with how much of a protein exists in folded form. The challenge of maintaining proteostasis in the cell thus becomes equivalent to recruiting chaperones, tuning translation and degradation rates, and ensuring that a given protein has sufficient time, space, and assistance to assume its native fold, while avoiding off-pathway oligomer and aggregate formation. This perspective is useful for thinking about protein folding in cells, and lends itself well to interpreting data from biochemical techniques that reveal how much of a protein is made, degraded, or stable. What it often fails to capture, however, is the importance of location. Where a protein ends up, and the extent to which this location is in line with where it is intended to be, arguably has as much significance for function as how much folded protein exists at all.

Because proteins rarely exist in isolation outside the test tube, the location of a protein plays a large part in deciding what neighbors it has, and what processes it can or cannot access. Subcellular compartments have distinct stress responses, designated isoforms of key chaperones, and different metal and redox conditions.[1-4] It is not surprising, then, that cells have innate safeguards against misplaced proteins, from localization sequences and cargo sorting proteins to pH-dependent recycling systems that determine protein trafficking pathways and organellar protein composition.[5, 6]

This thesis has considered two examples of methods cells can use to direct the localization of specific proteins. In the first case (Chapters II and III), cells regulate the export of collagen-I by controlling its membrane carriers. ER-resident (Tango1, Hsp47, cTAGE5[7-11]) and cytosolic (Hsp90β, and possibly NudCL; this work, as well as KLHL12, CUL3, ALG2, and PEF1[12, 13]) proteins are recruited to collagen-I at ER exit sites (ERES) and assist in transporting it from the ER. The inability of collagen-I to fit into conventional packaging vesicles, and the inclusion of chaperones like Hsp47 in the secretion machinery, allows cells to monitor collagen-I folding and (mostly) release it only when appropriate. The second case involves modification of the proteins themselves. Glycosylation of specific residues on proteins can influence their recruitment to microvesicles or provide epitopes for recognition by lectins and receptors,[14, 15] perhaps in response to cell-wide transcriptional changes (Chapter V). Alternatively, the presence of a single conserved N-glycan might enable proteins to engage lectin chaperones that facilitate their folding and/or secretion, as we observed in Chapter VI.

Moving forward, there is much work left to do. One area of particular interest is the proposed post-translational modification used to initiate membrane and/or COP-II protein recruitment.[12] The monoubiquitin moiety installed on Sec31a by a KLHL12-CUL3 complex bears similarities to a polyubiquitin motif required for post-Golgi anterograde, but not retrograde, trafficking. In this instance, a KLHL20-CUL3 complex ubiquitinates the protein coronin 7 and targets it to the trans-Golgi network. Time-lapse confocal microscopy revealed that KLHL20-CUL3 activity enables formation and elongation of cargo carriers from the TGN, likely by stabilizing F-actin during carrier coat assembly.[16] The export of representative secretory cargo was impaired in KLHL20-depleted cells, pointing to non-degradative ubiquitylation as a means of regulating secretion from both the ER and the TGN.[17, 18]

The existence of a common, ubiquitin-based signal for carrier growth and protein export might also indicate a more integrated picture of membrane utilization. It is possible, for instance, to envision cellular membrane distribution determining the balance between protein production (by expanding the ER and contributing to carrier growth) and degradation (via autophagosome formation or reducing ER volume). While speculative, this framework finds support in recent studies that directly link membrane availability to collagen secretion. An RNAi screen for *Drosophila* genes involved in collagen IV production identified dynamin, a GTPase required for endocytosis, as a screening hit.[19] The dynamin-deficient phenotype of autonomous, pericellular

collagen IV deposits could be rescued by depleting lipids from growth medium and thereby reducing excess plasma membrane. Such deposits are distinct from those observed upon Tango1 knockdown, which are solely intracellular. And although the basement membrane components nidogen (~139 kDa, 24 nm in radius) and perlecan (~470 kDa, 170 ± 20 nm in length) accumulated with collagen IV in dynamin-deficient adipocytes, laminin (140–400 kDa, depending on the chain; 130 nm in length) was not detected in the collagen-IV deposits; the results therefore suggest that protein identity, rather than size, determines susceptibility to pericellular trapping by collagen IV in the presence of excess plasma membrane.

The collagen and coat protein adapter Tango1 has been proposed to recruit ERGIC-derived membranes to ERES, driving expansion of collagen-containing COP-II vesicles.[9, 20] Consistent with this hypothesis, knockdown of Tango1 in *Drosophila* causes intracellular retention of the large protein Dumpy and the collagen gene Viking, blocking smaller cargoes from exiting the ER. However, Tango1 depletion also disrupted ER-Golgi morphology even in the absence of Dumpy, suggesting that the protein may have roles in organelle organization outside of vesicle expansion.[21] Thus, while specific signaling mechanisms (whether through post-translational modification or through designated cargo-receptor interactions) are needed to initiate secretion of large cargo, the dynamin and Tango1 deficiency phenotypes suggest that controlling membrane abundance constitutes another key mode of regulating secretion. Knockdown of syntaxins 5 and 18, membrane-associated proteins that drive vesicle fusion, impacted collagen VII secretion;[20] it therefore seems likely that other transmembrane proteins beyond these may prove to be involved as well.

We must note, however, that the signaling mechanisms for initiating collagen secretion are not yet defined. While biochemical, high-resolution imaging, and *in vitro* reconstitution assays support the hypothesis that collagen is exported from the ER in discrete vesicles that contain monoubiquitylated Sec31a,[12, 13, 22] imaging and genetic knockdown experiments by other groups have suggested the possibility of a continuous, membrane-bound carrier that remains associated with the ER.[23, 24] At the time of writing, preliminary work by another group has argued against transport of collagen-I in COP-II coated vesicles, presenting data that appears to be more consistent with contiguous carriers that extend from the ER to eventually fuse directly with the Golgi.[25] Resolution of the debate thus awaits additional evidence, perhaps through application of other methods such as proximity labeling, which can address these alternate theories.

Yet even in the absence of a complete mechanism, the assays and small molecule modulators originating from studies of collagen-I secretion provide tools for asking a new set of questions. Collagen types III and V, for instance, are fibrillar collagens that co-assemble with collagen-I to regulate fibrillogenesis and fibril diameter, respectively, in the extracellular matrix.[26, 27] What secretion pathways these collagens use, and to what extent the process is jointly regulated with secretion of collagen-I, could provide insight into EDS disease phenotypes associated with mutations in these collagens[28, 29] while identifying shared (or type-specific) features of collagen secretion mechanisms. A natural starting point for these studies might be examining collagen-III and -V secretion by cells treated with small molecule modulators of collagen-I secretion (Chapter II), or transduced with shRNA constructs against genes identified as collagen-I interactors.[30] Annexin A2 and Golim4, in particular, are strong candidates for initial shRNA studies, as both proteins localize to the secretory pathway and were required for endogenous collagen-I secretion by Saos-2 cells. While *in vitro* studies of collagen secretion have predominantly focused on either type VII or type I, Tango1 KO mice showed defective secretion of types I – IV, VII, and IX.[31] Such findings might suggest that types I and III (and II and IX, which co-assemble in cartilage[32]) share the same secretion pathway, but that their respective binding partners collagen-V and collagen-XI are exported through a different mechanism.

Another area for future work involves transcriptional regulation of collagen secretion. Besides the classical UPR stress sensors IRE1, PERK, and ATF6, five additional ATF6-like transcription factors have been identified. These transcription factors show low (~30%) overall sequence identity with ATF6, but contain key features of ATF6 protein structure, including a transcriptional activation domain, a bZIP domain, transmembrane region, and a motif for cleavage by site 1 protease (S1P).[33, 34] Among these new proteins, the osteoblast-specific OASIS and the chondrocyte-specific BBF2H7 target genes such as *COL1A1* and *SEC23A*, respectively, and are required for bone (OASIS) or cartilage (BBF2H7) development.[35, 36] Dual regulation of secretion-related genes appears possible, as ATF6 and BBF2H7 knockouts both showed reduced expression of Sec23a and Sec13. ATF6 activity, however, is not sufficient for secretion of extended collagen-II fibrils.[37]

OASIS[-/-] mice accumulate collagen-I in the rough ER of osteoblasts, and chondrocytes from BBF2H7[-/-] mice contain aggregated collagen-II in the rough ER, indicating that these transcription factors likely regulate collagen secretion.[35, 36] Surprisingly, overexpression of the active *N*-terminal portion of BBF2H7 causes embryonic lethality in the medaka fish, and studies of BBF2H7 function in this system required transient expression at an early stage of development, well before formation of the notochord and secretion of type II collagen.[37] However, doxycycline-responsive versions of OASIS and BBF2H7 have been stably integrated into MEFs and used to show that both transcription factors are regulated in a mechanism distinct from that of ATF6.[38] It would therefore be interesting to introduce an orthogonal, small molecule-regulated version of ATF6[39] to explore effects of controlled joint activation of these transcription factors, and perhaps obtain insight into the distinct transcriptional requirements for the secretion of large proteins such as collagen-I and II.

Finally, we have found that the transcription factor XBP1s is capable of globally altering the structures of *N*-glycans on cell surface and secreted proteins (Chapter V). Such work builds on previous reports that XBP1s activity can increase glycosylation by controlling biosynthesis of precursors.[40] We also show that, at least for misfolding variants of the collagen-I C-terminal propeptide domain, that a conserved *N*-glycan facilitates secretion of the protein, possibly by allowing it to access the calnexin/calreticulin lectin chaperones (e.g., Chapter VI). The presence of *N*-glycans has been found to both enhance protein stability and promote folding,[41] suggesting that a similar mechanism may motivate conservation of the *N*-glycan on the collagen-I propeptide.

Increasing evidence suggests that *O*-glycosylation, however, also acts to regulate protein secretion. *O*-GlcNAcylation stabilizes nucleoporins against proteasomal degradation, and has begun to gain acceptance as an intracellular signal akin to phosphorylation.[42, 43] Removal of *O*-GlcNac was first found to promote phosphorylation of the Sec24 protein family, impairing membrane recruitment and initiating a temporary trafficking block during mitosis.[44] A subsequent study of heterozygous and homozygous mice deficient for the glycosyltransferase Galnt1 showed that loss of *O*-glycosylation impaired secretion of collagen IV in the developing submandibular gland and activated both XBP1s and ATF6.[45] Notably, ablation of *N*-glycosylation by treatment with tunicamycin not only caused global ER stress, but also caused mistargeting of the non-basement membrane component E-cadherin, indicating distinct roles for *N*- and *O*-linked sugar modifications.

Since then, *O*-GlcNAcylation has been identified on several components of the COP-II machinery. Modification of Sec31a on the residue Ser964 decreases its affinity for ALG2 and increases membrane association of Sec31a, supporting ERES formation and protein secretion.[46] *O*-GlcNAcylation by the glycosyltransferase *PGANT4* protects Tango1 from furin proteolysis in *Drosophila*, allowing for protein secretion and maintenance of Golgi structure; *O*-glycosylation sites have also been identified on mammalian Tango1, and so proteolysis of non-glycosylated Tango1 may prove to be a conserved mode of regulating protein activity.[47] Meanwhile, a recent proteomics-based approach found *O*-GlcNAcylation sites on Sec23a,

Sec24b, Sec24c, Sec31a, and TFG, which were shown to media protein-protein interactions in mammalian cells.[48] Expression of the Ser184Ala mutant of Sec23a caused collagen retention in the ER of chondrosarcoma cells, and a nonglycosylatable mutant of Sec23a was only partially able to rescue collagen secretion in the Sec23a-depleted *crusher* zebrafish model. *O*-GlcNAcylation levels have been found to correlate with both chondrocyte[49] and osteoblast differentiation,[50] which could link the observed *O*-GlcNAc modification of COP-II proteins with BBF2H7 and OASIS signaling. Furthermore, the substrate specificity of the *O*-GlcNAc transferase seems to depend on UDP-GlcNAc concentration, which points to possible cross-talk between *N*-glycosylation and *O*-GlcNAcylation via XBP1s-mediated changes to UDP-GlcNAc availability.[51] That *N*-glycosylation sites have also been assigned to Tango1 further suggests that *N*- and *O*-glycosylation may coordinately regulate protein secretion, the details of which may prove relevant for our understanding of both development and disease.[52]

References

[1] Ron, D., and Walter, P. (2007) Signal integration in the endoplasmic reticulum unfolded protein response, *Nat. Rev. Mol. Cell Biol. 8*, 519-529.

[2] Vabulas, R. M., Raychaudhuri, S., Hayer-Hartl, M., and Hartl, F. U. (2010) Protein Folding in the Cytoplasm and the Heat Shock Response, *Cold Spring Harbor Perspect. Biol. 2*, a004390.

[3] Görlach, A., Klappa, P., and Kietzmann, D. T. (2006) The Endoplasmic Reticulum: Folding, Calcium Homeostasis, Signaling, and Redox Control, *Antioxid. Redox Signaling 8*, 1391-1418.

[4] Taipale, M., Jarosz, D. F., and Lindquist, S. (2010) HSP90 at the hub of protein homeostasis: emerging mechanistic insights, *Nat. Rev. Mol. Cell Biol. 11*, 515-528.

[5] Barlowe, C., and Helenius, A. (2016) Cargo Capture and Bulk Flow in the Early Secretory Pathway, *Annu. Rev. Cell Dev. Biol. 32*, 197-222.

[6] Oecal, S., Socher, E., Uthoff, M., Ernst, C., Zaucke, F., Sticht, H., Baumann, U., and Gebauer, J. M. (2016) The pH-dependent Client Release from the Collagen-specific Chaperone HSP47 Is Triggered by a Tandem Histidine Pair, *J. Biol. Chem. 291*, 12612-12626.

[7] Saito, K., Chen, M., Bard, F., Chen, S., Zhou, H., Woodley, D., Polischuk, R., Schekman, R., and Malhotra, V. (2009) TANGO1 facilitates cargo loading at endoplasmic reticulum exit sites, *Cell 136*, 891-902.

[8] Saito, K., Yamashiro, K., Ichikawa, Y., Erlmann, P., Kontani, K., Malhotra, V., and Katada, T. (2011) cTAGE5 mediates collagen secretion through interaction with TANGO1 at endoplasmic reticulum exit sites, *Mol. Biol. Cell 22*, 2301-2308.

[9] Santos, A. J., Raote, I., Scarpa, M., Brouwers, N., and Malhotra, V. (2015) TANGO1 recruits ERGIC membranes to the endoplasmic reticulum for procollagen export, *eLife 4*, e10982.

[10] Ma, W., and Goldberg, J. (2016) TANGO1/cTAGE5 receptor as a polyvalent template for assembly of large COPII coats, *Proc. Natl. Acad. Sci. U.S.A. 113*, 10061-10066.

[11] Ishikawa, Y., and Bachinger, H. P. (2013) A molecular ensemble in the rER for procollagen maturation, *Biochim. Biophys. Acta 1833*, 2479-2491.

[12] Jin, L., Pahuja, K. B., Wickliffe, K. E., Gorur, A., Baumgartel, C., Schekman, R., and Rape, M. (2012) Ubiquitin-dependent regulation of COPII coat size and function, *Nature 482*, 495-500.

[13] McGourty, C. A., Akopian, D., Walsh, C., Gorur, A., Werner, A., Schekman, R., Bautista, D., and Rape, M. Regulation of the CUL3 Ubiquitin Ligase by a Calcium-Dependent Co-adaptor, *Cell 167*, 525-538.

[14] Liang, Y., Eng, W. S., Colquhoun, D. R., Dinglasan, R. R., Graham, D. R., and Mahal, L. K. (2014) Complex N-linked glycans serve as a determinant for exosome/microvesicle cargo recruitment, *J. Biol. Chem. 289*, 32526-32537.

[15] Varki, A. (2017) Biological roles of glycans, *Glycobiology 27*, 3-49.

[16] Yuan, W. C., Lee, Y. R., Lin, S. Y., Chang, L. Y., Tan, Y. P., Hung, C. C., Kuo, J. C., Liu, C. H., Lin, M. Y., Xu, M., Chen, Z. J., and Chen, R. H. (2014) K33-Linked Polyubiquitination of Coronin 7 by Cul3-KLHL20 Ubiquitin E3 Ligase Regulates Protein Trafficking, *Mol. Cell 54*, 586-600.

[17] Lu, A., and Pfeffer, S. R. (2014) A CULLINary ride across the secretory pathway: more than just secretion, *Trends Cell Biol. 24*, 389-399.

[18] Giovannone, A. J., Reales, E., Bhattaram, P., Fraile-Ramos, A., Weimbs, T., and Mostov, K. E. (2017) Monoubiquitination of syntaxin 3 leads to retrieval from the basolateral plasma membrane and facilitates cargo recruitment to exosomes, *Mol. Biol. Cell 28*, 2843-2853.

[19] Zang, Y., Wan, M., Liu, M., Ke, H., Ma, S., Liu, L. P., Ni, J. Q., and Pastor-Pareja, J. C. (2015) Plasma membrane overgrowth causes fibrotic collagen accumulation and immune activation in Drosophila adipocytes, *eLife 4*, e07187.

[20] Nogueira, C., Erlmann, P., Villeneuve, J., Santos, A. J., Martinez-Alonso, E., Martinez-Menarguez, J. A., and Malhotra, V. (2014) SLY1 and Syntaxin 18 specify a distinct pathway for procollagen VII export from the endoplasmic reticulum, *eLife 3*, e02784.

[21] Rios-Barrera, L. D., Sigurbjornsdottir, S., Baer, M., and Leptin, M. (2017) Dual function for Tango1 in secretion of bulky cargo and in ER-Golgi morphology, *Proc. Natl. Acad. Sci. U.S.A. 114*, e10389-e10398.

[22] Gorur, A., Yuan, L., Kenny, S. J., Baba, S., Xu, K., and Schekman, R. (2017) COPII-coated membranes function as transport carriers of intracellular procollagen I, *J. Cell Biol. 216*, 1745-1759.

[23] Malhotra, V., and Erlmann, P. (2015) The pathway of collagen secretion, *Annu. Rev. Cell Dev. Biol. 31*, 109-124.

[24] Mironov, A. A., Mironov, A. A., Jr., Beznoussenko, G. V., Trucco, A., Lupetti, P., Smith, J. D., Geerts, W. J., Koster, A. J., Burger, K. N., Martone, M. E., Deerinck, T. J., Ellisman, M. H., and Luini, A. (2003) ER-to-Golgi carriers arise through direct en bloc protrusion and multistage maturation of specialized ER exit domains, *Dev. Cell 5*, 583-594.

[25] McCaughey, J., Stevenson, N., Cross, S., and Stephens, D. (2018) ER-to-Golgi trafficking of procollagen in the absence of large carriers, *bioRxiv*.

[26] Liu, X., Wu, H., Byrne, M., Krane, S., and Jaenisch, R. (1997) Type III collagen is crucial for collagen I fibrillogenesis and for normal cardiovascular development, *Proc. Natl. Acad. Sci. U.S.A. 94*, 1852-1856.

[27] Kadler, K. E., Hill, A., and Canty-Laird, E. G. (2008) Collagen fibrillogenesis: fibronectin, integrins, and minor collagens as organizers and nucleators, *Curr. Opin. Cell Biol. 20*, 495-501.

[28] Smith, L. T., Schwarze, U., Goldstein, J., and Byers, P. H. (1997) Mutations in the COL3A1 Gene Result in the Ehlers–Danlos Syndrome Type IV and Alterations in the Size and Distribution of the Major Collagen Fibrils of the Dermis, *J. Invest. Dermatol. 108*, 241-247.

[29] Wenstrup, R. J., Florer, J. B., Davidson, J. M., Phillips, C. L., Pfeiffer, B. J., Menezes, D. W., Chervoneva, I., and Birk, D. E. (2006) Murine model of the Ehlers-Danlos syndrome. col5a1 haploinsufficiency disrupts collagen fibril assembly at multiple stages, *J. Biol. Chem. 281*, 12888-12895.

[30] DiChiara, A. S., Taylor, R. J., Wong, M. Y., Doan, N. D., Rosario, A. M., and Shoulders, M. D. (2016) Mapping and exploring the collagen-I proteostasis network, *ACS Chem. Biol. 11*, 1408-1421.

[31] Wilson, D. G., Phamluong, K., Li, L., Sun, M., Cao, T. C., Liu, P. S., Modrusan, Z., Sandoval, W. N., Rangell, L., Carano, R. A., Peterson, A. S., and Solloway, M. J. (2011) Global defects in collagen secretion in a Mia3/TANGO1 knockout mouse, *J. Cell Biol. 193*, 935-951.

[32] Mendler, M., Eich-Bender, S. G., Vaughan, L., Winterhalter, K. H., and Bruckner, P. (1989) Cartilage contains mixed fibrils of collagen types II, IX, and XI, *J. Cell Biol. 108*, 191-197.

[33] Kondo, S., Saito, A., Hino, S., Murakami, T., Ogata, M., Kanemoto, S., Nara, S., Yamashita, A., Yoshinaga, K., Hara, H., and Imaizumi, K. (2007) BBF2H7, a novel transmembrane bZIP transcription factor, is a new type of endoplasmic reticulum stress transducer, *Mol. Cell Biol. 27*, 1716-1729.

[34] Kondo, S., Saito, A., Asada, R., Kanemoto, S., and Imaizumi, K. (2011) Physiological unfolded protein response regulated by OASIS family members, transmembrane bZIP transcription factors, *IUBMB Life 63*, 233-239.

[35] Saito, A., Hino, S.-i., Murakami, T., Kanemoto, S., Kondo, S., Saitoh, M., Nishimura, R., Yoneda, T., Furuichi, T., Ikegawa, S., Ikawa, M., Okabe, M., and Imaizumi, K. (2009) Regulation of endoplasmic reticulum stress response by a BBF2H7-mediated Sec23a pathway is essential for chondrogenesis, *Nat. Cell Biol. 11*, 1197-1204.

[36] Murakami, T., Saito, A., Hino, S., Kondo, S., Kanemoto, S., Chihara, K., Sekiya, H., Tsumagari, K., Ochiai, K., Yoshinaga, K., Saitoh, M., Nishimura, R., Yoneda, T., Kou, I., Furuichi, T., Ikegawa, S., Ikawa, M., Okabe, M., Wanaka, A., and Imaizumi, K. (2009) Signalling mediated by the endoplasmic reticulum stress transducer OASIS is involved in bone formation, *Nat. Cell Biol. 11*, 1205-1211.

[37] Ishikawa, T., Toyama, T., Nakamura, Y., Tamada, K., Shimizu, H., Ninagawa, S., Okada, T., Kamei, Y., Ishikawa-Fujiwara, T., Todo, T., Aoyama, E., Takigawa, M., Harada, A., and Mori, K. (2017) UPR transducer BBF2H7 allows export of type II collagen in a cargo- and developmental stage–specific manner, *J. Cell Biol. 216*, 1761-1774.

[38] Kondo, S., Hino, S. I., Saito, A., Kanemoto, S., Kawasaki, N., Asada, R., Izumi, S., Iwamoto, H., Oki, M., Miyagi, H., Kaneko, M., Nomura, Y., Urano, F., and Imaizumi, K. (2012) Activation of OASIS family, ER stress transducers, is dependent on its stabilization, *Cell Death Differ. 19*, 1939-1949.

[39] Shoulders, M. D., Ryno, L. M., Genereux, J. C., Moresco, J. J., Tu, P. G., Wu, C., Yates, J. R., 3rd, Su, A. I., Kelly, J. W., and Wiseman, R. L. (2013) Stress-independent activation of XBP1s and/or ATF6 reveals three functionally diverse ER proteostasis environments, *Cell Rep. 3*, 1279-1292.

[40] Wang, Z. V., Deng, Y., Gao, N., Pedrozo, Z., Li, D. L., Morales, C. R., Criollo, A., Luo, X., Tan, W., Jiang, N., Lehrman, M. A., Rothermel, B. A., Lee, A. H., Lavandero, S., Mammen, P. P., Ferdous, A., Gillette, T. G., Scherer, P. E., and Hill, J. A. (2014) Spliced X-box binding protein 1 couples the unfolded protein response to hexosamine biosynthetic pathway, *Cell 156*, 1179-1192.

[41] Hanson, S. R., Culyba, E. K., Hsu, T.-L., Wong, C.-H., Kelly, J. W., and Powers, E. T. (2009) The core trisaccharide of an N-linked glycoprotein intrinsically accelerates folding and enhances stability, *Proc. Natl. Acad. Sci. U.S.A. 106*, 3131-3136.

[42] Zhu, Y., Liu, T.-W., Madden, Z., Yuzwa, S. A., Murray, K., Cecioni, S., Zachara, N., and Vocadlo, D. J. (2016) Post-translational O-GlcNAcylation is essential for nuclear pore integrity and maintenance of the pore selectivity filter, *J. Mol. Cell Biol. 8*, 2-16.

[43] Wells, L., Vosseller, K., and Hart, G. W. (2001) Glycosylation of Nucleocytoplasmic Proteins: Signal Transduction and O-GlcNAc, *Science 291*, 2376-2378.

[44] Dudognon, P., Maeder-Garavaglia, C., Carpentier, J. L., and Paccaud, J. P. (2004) Regulation of a COPII component by cytosolic O-glycosylation during mitosis, *FEBS Lett. 561*, 44-50.

[45] Tian, E., Hoffman, M. P., and Ten Hagen, K. G. (2012) O-glycosylation modulates integrin and FGF signalling by influencing the secretion of basement membrane components, *Nat. Commun. 3*, 869.

[46] Cho, H. J., and Mook-Jung, I. (2018) O-GlcNAcylation regulates endoplasmic reticulum exit sites through Sec31A modification in conventional secretory pathway, *FASEB J.* fj201701523R.

[47] Zhang, L., Syed, Z. A., van Dijk Hard, I., Lim, J. M., Wells, L., and Ten Hagen, K. G. (2014) O-glycosylation regulates polarized secretion by modulating Tango1 stability, *Proc. Natl. Acad. Sci. U.S.A. 111*, 7296-7301.

[48] Cox, N. J., Unlu, G., Bisnett, B. J., Meister, T. R., Condon, B. M., Luo, P. M., Smith, T. J., Hanna, M., Chhetri, A., Soderblom, E. J., Audhya, A., Knapik, E. W., and Boyce, M. (2018) Dynamic Glycosylation Governs the Vertebrate COPII Protein Trafficking Pathway, *Biochemistry 57*, 91-107.

[49] Andrés-Bergós, J., Tardio, L., Larranaga-Vera, A., Gómez, R., Herrero-Beaumont, G., and Largo, R. (2012) The increase in O-GlcNAc protein modification stimulates chondrogenic differentiation both in vitro and in vivo, *J. Biol. Chem. 287*, 33615-33628.

[50] Koyama, T., and Kamemura, K. (2015) Global increase in O-linked N-acetylglucosamine modification promotes osteoblast differentiation, *Exp. Cell Res. 338*, 194-202.

[51] Zachara, N. A., Y.; Hart, G.W. (2017) The O-GlcNAc Modification, In *Essentials of Glycobiology* (Varki, A. e. a., Ed.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

[52] Chen, R., Jiang, X., Sun, D., Han, G., Wang, F., Ye, M., Wang, L., and Zou, H. (2009) Glycoproteomics analysis of human liver tissue by combination of multiple enzyme digestion and hydrazide chemistry, *J. Proteome Res. 8*, 651-661.

Supporting Experimental Procedures

RNA Extraction and Real-Time qPCR. HEK^{XBP1s} and HeLa^{XBP1s} cells were plated and treated as for lectin microarray analysis (below). After 72 h of XBP1s activation with 1 μg/mL dox (or 24 h of treatment with 0.1% DMSO or 750 nM Tg), cells were washed with PBS and RNA was extracted using the Omega E.Z.N.A. Total RNA extraction kit. cDNA was prepared from 500 ng RNA, normalized for all samples in each run, using an Applied Biosystems Reverse Transcriptase cDNA Kit in a BioRad Thermocycler. Samples were run on a Light Cycler 480 II Real Time PCR Instrument in the MIT BioMicro Center using previously described primers and data were analyzed as described previously.[1, 2] For qPCR arrays, HEK^{XBP1s} and HeLa^{XBP1s} cells were treated for 48 h with 1 μg/mL dox or 0.1% DMSO prior to harvesting. RNA was then extracted using the Qiagen RNeasy kit, cDNA was prepared from equal amounts of RNA using the Qiagen RT2 First Strand Kit in a BioRad Thermocycler, and samples were loaded on an RT^2 Profiler PCR Array for Human Glycosylation (Qiagen PAHS-046Z). Analyses were performed in a Light Cycler 480 II Real Time PCR Instrument in the MIT BioMicro Center. Data from three biological replicates were analyzed using the ΔΔCt method. A list of glycogenes was manually curated from the HEK^{XBP1s} microarray[2] and HeLa^{XBP1s} RNA-Seq data, and then combined with detected glycogenes from the PCR array (Table S8). Significance cut-offs used were FDR or *p*-value ≤ 0.05 and fold-change ≥ 1.5.

RNA-Seq. HeLa^{XBP1s} cells were plated in biological triplicate at a density of 6 × 10^5 cells per well in 6-well plates and allowed to adhere overnight. Cells were then treated for 48 h with vehicle or 1 μg/mL dox, or for 24 h with 750 nM Tg. Cells were harvested and RNA was extracted using the RNeasy Plus Mini Kit (Qiagen). RNA quality was confirmed using a Fragment Analyzer (Advanced Analytical). RNA samples were then loaded on a HiSeq cartridge as a 50 base single-end run with 6 + 6 nucleotide indexes. *H. sapiens* RNA-Seq reads were aligned to hg19 with bowtie version 1.0.1[3] and expression was summarized using rsem version 1.2.26[4] using the ensemble gencode annotation release 75. Differential expression analysis was done with deseq2 version 1.10.0[5] running under R version 3.2.3. Default options were selected for deseq2 runs, except Cooks Cutoff and Independent Filtering were both set to false during results preparation. Gene Set Enrichment Analysis[6] Java command-line version 2.3.0 beta was run in both pre-ranked and standard mode using the stat output from deseq2 to order genes. Custom gene sets and a selected subset from MSigDB version 5.2 were analyzed; these gene sets are provided in Table S1E. Hierarchical clustering was performed with Spotfire 7.6.1, using counts of transcripts with highest expression for each gene.

Membrane Proteome Preparation. HeLa^{XBP1s} or HEK293^{XBP1s} cells were plated in 10 cm dishes at a density of 1 × 10^6 cells per plate and allowed to adhere overnight. The next day, XBP1s expression was induced by treatment with 1 μg/mL dox. Vehicle and Tg-treated plates received fresh media and either DMSO or 750 nM Tg for 24 h before harvesting. After 72 h of induction, cells were harvested by scraping in 1× PBS + 1 mM EDTA, sonicated to disrupt cells, and then centrifuged at 35k RPM at 4 °C for 1 h to pellet the membrane fraction. Pellets were resuspended in 100 μL PBS, homogenized with 21G and 27G needles, and protein concentrations were measured via A280 on a NanoDrop ND-1000 spectrophotometer.

Secretome Preparation from HEK^{XBP1s} Cells. Prior to plating cells, 15 cm dishes were coated with 0.05 mg/mL of poly-D-lysine hydrobromide (Sigma P6407) for 10 min at room temperature. Each dish was washed three times with PBS before seeding with 10 × 10^6 HEK^{XBP1s} cells in complete medium. After 24 h, media were changed to add the appropriate compounds: 1 μg/mL dox for XBP1s activation or 0.1% DMSO for control. After another 24 h, media were removed and cells were washed with PBS (containing Ca^{2+} and Mg^{2+}) three times. Cells were then incubated in Freestyle^{TM} medium (Gibco) with either 1 μg/mL dox, 0.1% DMSO, or 750 nM of Tg. Conditioned media were harvested from cells after a total of 48 h with dox or DMSO, or 28 h with Tg. Medium samples were filtered through 0.2 μm PES membranes (VWR) and concentrated in 3 kDa MWCO centrifugal units (Millipore Amicon Ultra). For immunoblotting analysis of the secretome, 6 × 10^5 HEK^{XBP1s} cells were plated per well on poly-D-lysine coated 6-well plates and treated with dox or vehicle as described above. After 24 h, cells were washed with PBS (containing Ca^{2+} and Mg^{2+}) three times and then incubated in DMEM (+10% FBS, L-glutamine) or Freestyle^{TM} media, with or without 1 μg/mL dox. Conditioned media were collected after 48 h, spun at 1.5k RPM at 4 °C for 5 min to pellet cell debris, and then run on 4/8% SDS-PAGE gels. Proteins were transferred to nitrocellulose membranes and incubated with an α-KDEL (recognizes both Grp78 and Grp94, Enzo ADI-SPA-827) antibody. Blots were developed with the appropriate 800CW secondary antibody (LI-COR) prior to scanning on an Odyssey infrared imager (LI-COR).

Lectin Flow Cytometry. HeLa^{XBP1s} cells were treated with or without dox for 72 h, then very briefly trypsinized. Cells were washed twice with 2% FBS in PBS, and 1 × 10^6 cells/sample were incubated for 30 min at 4 °C with HHL-biotin (Vector Labs) at 5 μg/mL. Competitive sugars (200 mM α-methylmannoside, Sigma-Aldrich) were pre-incubated 30 min with lectins at room temperature for inhibitory controls. Cells were washed and resuspended in 20 μg/mL of Cy5-streptavidin (Thermo Fisher Scientific) for 20 min at 4 °C in the dark. Cells were resuspended in 1 mL of 2% FBS/PBS, passed through cell strainers (Falcon), and analyzed on a BD Accuri C6 flow cytometer.

CellTiter-Glo Assay. HeLa^{XBP1s} cells were cultured with or without doxycycline (1 μg/mL, 72 h) or 1-DMM (80 μg/mL, 24 h). Cells were counted and plated at 1 × 10^4 cells/well 24 h prior to carrying out the assay according to the manufacturer's instructions (Promega). In brief, cells were cultured in 100 μL of culture media and an equal volume of reagent was added to lyse cells for 2 min. After incubation at room temperature for 10 min, luminescence measurements were taken.

Resazurin Assay. HeLa^{XBP1s} cells were cultured with or without doxycycline (1 μg/mL, 72 h) or 1-DMM (80 μg/mL, 24 h). Cells were counted and plated at 1 × 10^4 cells/well 24 h prior to measurements. Resazurin was added to cells at a final concentration of 0.03 mg/mL for 60 min at 37 °C. Fluorescence readings were taken at 560 nm excitation, 590 nm emission.

Secretome Proteomics Analysis. HEK^{XBP1s} cells were treated with either vehicle or doxycycline for 16 h, washed well, and incubated in DMEM without serum for 6 h. Proteins were then precipitated from conditioned media with 10% TCA for 1.5 h at 4 °C and resolublized using 1% aqueous Rapigest (Waters) in 100 mM HEPES pH 8.0. Proteins were reduced with 5 mM TCEP (Sigma) for 30 min at 37 °C, alkylated with 10 mM iodoacetamide (Sigma) for 30 min in the dark at room temperature, digested with 0.5 μg trypsin overnight at 37 °C and 600 rpm, labeled with appropriate isotopic TMT reagent in 40% CHC$_3$N for 1 h, quenched with 0.4% ammonium bicarbonate for 1 h, pooled, evaporated, resuspended in buffer A (5% CHC$_3$N, 0.1% formic

acid), brought to pH < 2 with formic acid, heated at 37 °C to precipitate Rapigest, and stored at -80 °C prior to LC/LC-MS/MS analysis.

Nanopure water and mass spectrometry grade solvents were used for all preparations. MudPIT loading columns were prepared by briefly dipping 250 μm inner diameter (ID) undeactivated fused silica capillaries (Agilent) in 3:1 Kasil 1624 (PQ Corp) : formamide, curing overnight at 100 °C, and trimming the frit to 1 mm. The columns were then rinsed with MeOH, loaded under pressure with 2.5 cm strong cation exchange resin (SCX Luna, 5 μm diameter, 125 Å pore size, Partisphere), and loaded with another 2.5 cm reversed phase resin (C18 Aqua, 5 μm diameter, 125 Å pore size, Phenomenex). The columns were rinsed well with methanol and buffer A prior to loading the peptide digest, and further washed with buffer A. Analytical columns were prepared by pulling 100 μm ID-fused capillaries to a 5 μm ID on a P-2000 tip puller (Sutter Instrument Co.) and loaded with 15 cm of reversed phase resin, followed by methanol rinsing and equilibration in buffer A. For LC-MS/MS, each loading column was connected to the analytical column by a zero-dead-volume union, and connected to the HPLC through a tee junction that allowed connection to the 2.5 kV ESI voltage. A flow rate of 300 nL/min was maintained through a 1:1000 split flow line from an Agilent 1200 pump. MuDPIT experiments were performed where each step corresponds to 0, 10, 20, 30, 40, 50, 60, 70, 80, 90, and 100% buffer C being run for 4 min at the beginning of each gradient of buffer B. Electrospray was performed directly from the analytical column by applying the ESI voltage at a tee (150 mm ID, Upchurch Scientific) directly downstream of a 1:1000 split flow used to reduce the flow rate to 300 nL/min through the columns. Electrospray directly from the LC column was done at 2.5 kV with an inlet capillary temperature of 275 °C.

Data-dependent acquisition of MS/MS spectra were performed with the following settings: eluted peptides were scanned from 300 to 1600 m/z with resolution 30000 and the mass spectrometer in a data dependent acquisition mode. The top ten peaks for each full scan were fragmented by HCD using a normalized collision energy of 45%, a 100 ms activation time, and a resolution of 7500. Dynamic exclusion parameters were 1 repeat count, 30 ms repeat duration, 500 exclusion list size, 120 s exclusion duration, and exclusion width between 0.51 and 1.51. Protein and peptide identification and protein quantitation were done with the Integrated Proteomics Pipeline - IP2 (Integrated Proteomics Applications, Inc., San Diego, CA. http://www.integratedproteomics.com/). Tandem mass spectra were extracted from raw files using Raw Xtractor 1.9.13 and were searched against a database containing 20245 human sequences (longest entry for the IPI database for each protein) with reversed sequences using ProLuCID. Carbamidomethylation (+57.02146 Da) of cysteine and TMT tagging of N-termini and of lysine residues (+229.1629 Da) were considered as static modifications. Peptide candidates were filtered using DTASelect2 (version 2.0.27) for a false positive (decoy) peptide ratio of ~1%. Quantitation was performed using Census7, followed by deconvolution of isotopic impurity as reported in the lot analysis supplied by Thermo Fisher, and finally normalization of ratio values based on the mode. Redundant peptides were generally assigned to all proteins. The heavy to light ratio was quantified by Census.[7]

Supporting Tables

Table S1: GSEA results and gene sets (see .xls file in Supporting Reference [9]).
Table S1A (Sheet 1). GSEA results for Tg treatment using c5bp (MSigDB), shown in Figure S1B.
Table S1B (Sheet 2). GSEA results for Tg treatment using c5mf (MSigDB), shown in Fig. S1B.
Table S1C (Sheet 3). GSEA results for XBP1s activation using c5bp (MSigDB), shown in Figure S1B.
Table S1D (Sheet 4). GSEA results for XBP1s activation using c5mf (MSigDB), shown in Figure S1B.
Table S1E (Sheet 5). Full list of glycogenes in custom "lipid-linked oligosaccharide biosynthesis" and "trimming and elaboration" gene sets used for generating enrichment plots shown in Figure S8A and B.

Table S2. List of lectins used on the microarrays (see .xls file in Supporting Reference [9]).
Table S2A (Sheet 1). List of lectins used for HEK293[XBP1s] microarray experiments. Lectin print concentration, print sugar, and glycan epitope are indicated. Duplicate lectins obtained from different sources are noted in the last column.
Table S2B (Sheet 2). List of lectins used for HeLa[XBP1s] microarray experiments. Lectin print concentration, print sugar, and glycan epitope are indicated. Duplicate lectins obtained from different sources are noted in the last column.

Table S3. Processed lectin microarray data (see .xls file in Supporting Reference [9]). Data represent median normalized log2 values of sample to pooled-sample reference ratios.
Table S3A (Sheet 1). HEK[XBP1s] membrane data.
Table S3B (Sheet 2). HeLa[XBP1s] membrane data.
Table S3C (Sheet 3). HEK[XBP1s] media data.

Table S4: GC-MS linkage analysis of partially methylated alditol acetates (PMAA) from vehicle, XBP1s-activated, and thapsigargin-treated membrane (HEK[XBP1s] and HeLa[XBP1s]) and secreted (HEK[XBP1s]) glycoproteins. *N*-Linked glycans were permethylated, hydrolyzed, reduced, acetylated and analyzed by GC-MS (see Experimental Procedures).

Quantification[a]	Identification[b]	Residue	Rt (min)	HEK membrane[c] Vehicle	XBP1s	Tg	HeLa membrane[c] Vehicle	XBP1s	Tg	HEK secretome[c] Vehicle	XBP1s	Tg
101, 102, 115, 118, 131, 162, 175	142, 187	t-Fuc	13.88	30.1	9.6	9.2	20.6	25.9	36.5	84.8	95.5	121.5
102, 118, 129, 145, 161, 162, 205	71, 87, 113, 174	t-Man	15.32	163.9	189.7	205.4	134.4	151.8	369.8	20.2	25.3	29.4
102, 118, 129, 145, 161, 162, 205	71, 87, 113, 174	t-Gal	15.59	48.6	25.3	23.4	58.7	92.5	49.1	133.1	154.9	76.9
100, 101, 129, 130, 161, 190	71, 87, 88	2-Man	16.44	276.9	297.5	335.4	321.0	300.9	460.4	190.1	222.0	308.9
130, 190, 205	87, 88, 100, 129, 145, 161, 174, 205	2-Gal	16.71	1.2	0.3	0.7	0.0	0.7	0.0	2.9	2.3	1.0
118, 143	87, 101, 129, 174, 202, 217, 234	3-Gal	16.77	29.3	6.7	4.0	26.5	25.3	11.9	56.2	36.1	61.4
102, 118, 129, 159, 162	71, 99, 143, 173, 189, 233	6-Gal	17.24	23.6	4.0	4.1	0.0	4.1	0.0	98.9	63.8	100.1
130, 190, 233	87, 88, 172, 173	2,4-Man	17.58	13.7	6.5	7.3	8.2	16.6	5.3	36.1	25.8	25.6
100, 129, 130, 189, 190	87, 88, 99, 143, 159, 173, 174	2,6-Man	17.91	30.5	11.0	9.6	16.2	26.5	9.4	105.8	71.1	36.0
118, 129, 189, 234	74, 87, 101, 160, 174, 202, 245	3,6-Man	18.15	98.6	99.6	99.2	100.0	100.0	100.0	95.3	97.7	96.3
118, 129, 189, 234	74, 87, 101, 160, 174, 202	3,6-Gal	18.32	n.d.[d]	n.d.	n.d.	11.9	16.4	6.7	n.d.	n.d.	n.d.
118, 139	87, 97, 129, 160, 202	3,4,6-Man	18.56	1.4	0.4	0.8	n.d.	n.d.	n.d.	4.7	2.3	3.7
117, 129, 143, 145, 159, 203, 205	75, 187, 217	t-GlcNAc	18.98	2.4	1.8	4.2	1.9	2.1	0.0	4.8	2.7	1.3
117, 159, 233	75, 99, 143, 171, 203	4-GlcNAc	19.83	163.1	56.0	70.1	66.0	206.7	35.3	415.7	108.7	89.1
117, 142, 159	75, 99, 171, 231, 244	3,4-GlcNAc	20.62	1.5	0.9	0.8	1.2	2.8	10.0	5.8	5.2	2.2
117, 159	75, 99, 124, 127, 142, 143, 201, 245	4,6-GlcNAc	21.05	6.6	1.8	3.7	5.4	10.6	3.1	30.3	11.9	10.9

[a] Electron impact fragment ions used for quantification of the PMAA residues. For relative abundance, in order to minimize interference from the baseline and/or contamination, all fragment ions greater than 100 (>100) were used for the extracted ion current (XIC) chromatogram, as indicated above.

[b] Electron impact fragment ions used for identification of the PMAA residues, in addition to the ions used for quantification.

[c] Relative abundances were obtained by normalizing the integrated area peak of the extracted ion current (XIC) chromatogram of a specific residue to the summed integrated areas of the XICs of 3,6-linked mannose and 3,4,6-linked mannose.

[d] *n.d.* = not detected.

160

Table S5: Summary of observed changes to the HEK^XBP1s and HeLa^XBP1s N-glycome with XBP1s activation.

Observed Change with XBP1s	HEK^XBP1s Membrane	HEK^XBP1s Secretome	HeLa^XBP1s Membrane
Increased high-mannose		✓	✓
Decreased sialylation	✓	✓	
Decreased bisecting GlcNAc	✓	✓	
Increased core fucosylation			✓
Increased tetra-antennary glycans			✓

Table S6: Summary of observed changes to the HEK^XBP1s and HeLa^XBP1s N-glycome with global UPR activation by treatment with Tg.

Observed Change with Tg	HEK^XBP1s Membrane	HEK^XBP1s Secretome	HeLa^XBP1s Membrane
Increased high-mannose			✓
Decreased sialylation	✓	✓	
Decreased bisecting GlcNAc	✓	✓	
Increased core fucosylation			
Decreased branching		✓	✓
Increased core 1/3 O-glycans	✓		

Table S7. MS analysis of the HEK^XBP1s secretome with XBP1s activation (see .xls file in Supporting Reference [9]). Spectral counts, m/z ratios, and normalized intensities for all proteins detected by TMT-MS with or without XBP1s activation in Replicates 1–3 (Sheets 7A–7C, respectively). Processed data are presented in Sheet 7D.
Table S7A (Sheet 1). Full TMT-MS data for the 6 h secretome of HEK^XBP1s cells (Replicate 1).
Table S7B (Sheet 2). Full TMT-MS data for the 6 h secretome of HEK^XBP1s cells (Replicate 2).
Table S7C (Sheet 3). Full TMT-MS data for the 6 h secretome of HEK^XBP1s cells (Replicate 3).
Table S7D (Sheet 4). Fold-change and significance values for all proteins in the XBP1s secretome. Proteins that met fold-change and significance cut-offs are highlighted; significance and fold-change cut-offs were as follows: p-value ≤ 0.05, fold-change ≥ 1.5-fold. Proteins annotated as N-glycosylated (UniProtKB) are indicated along with the reported number of N-glycan sequons.

Table S8: HEK^XBP1s and HeLa^XBP1s glycogene data (see .xls file in Supporting Reference [9]).
Table S8A (Sheet 1). Gene sets related to N-glycosylation (including glycosyltransferases, glucosidases, and sugar transporters) were selected from MSigDB and used to generate Figure 5.5.
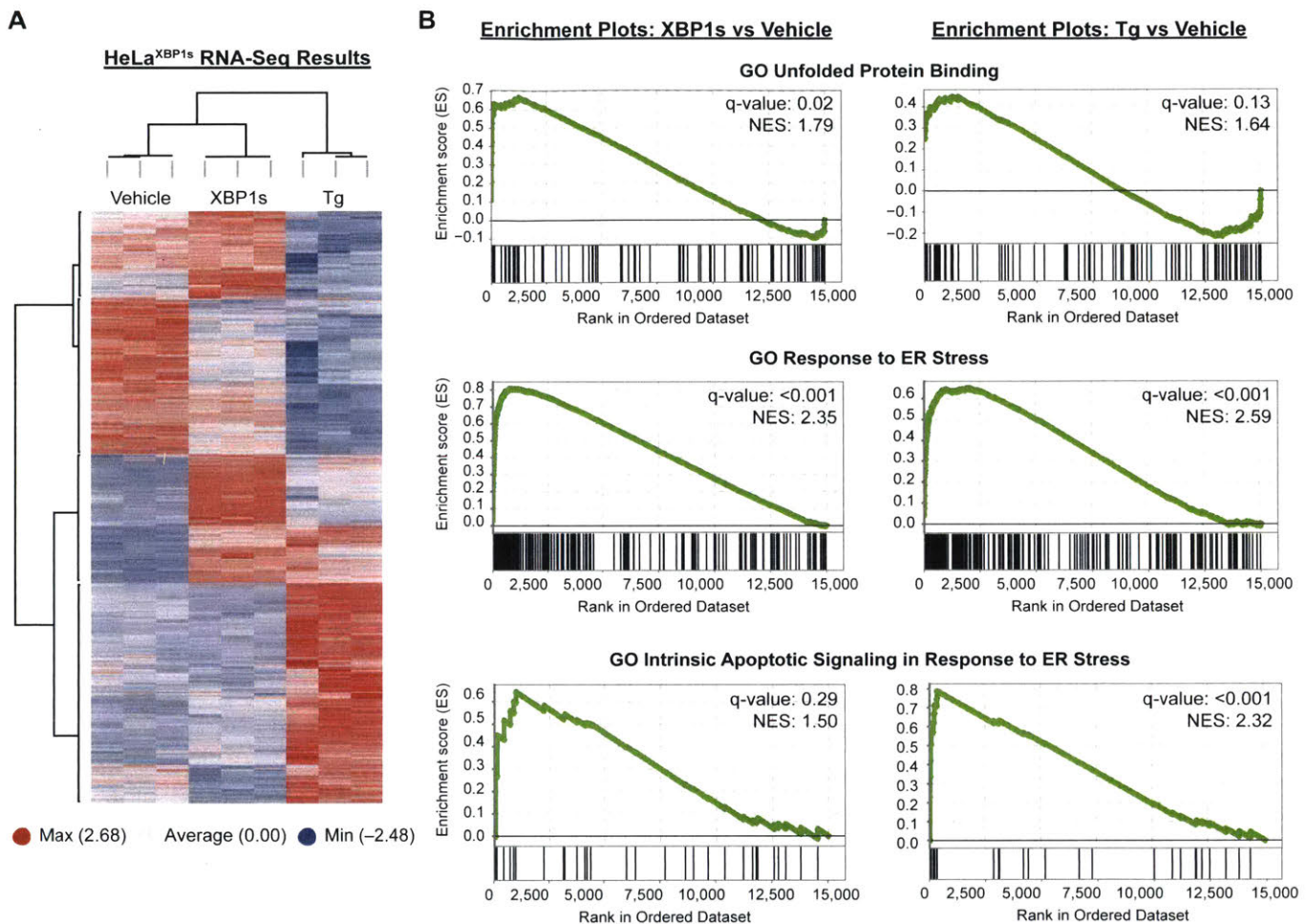Table S8B (Sheet 2). Full list of unique genes contained in the selected gene sets listed in Table S8A.
Table S8C (Sheet 3). Fold-change and significance values for glycogenes in HEK^XBP1s cells with XBP1s activation used to generate Figure 5.5A. Glycosylation-related genes listed in Table S8B were extracted from the HEK^XBP1s microarray data and were integrated with glycosylation PCR array data; bold-faced genes have fold-change and significance values obtained from the qPCR

arrays. Significance and fold-change cut-offs were as follows: FDR or *p*-value ≤ 0.05, fold-changes ≥ 1.5-fold (XBP1s versus vehicle). Genes from the list in Table S8B that were not detected in either data set are not shown.

Table S8D (Sheet 4). Fold-change and significance values for glycogenes in HeLa^{XBP1s} cells with XBP1s activation used to generate Figure 5.5C. Glycosylation-related genes listed in Table S8B were extracted from the HeLa^{XBP1s} RNA-Seq data and integrated with glycosylation PCR array data; bold-faced genes have fold-change and significance values obtained from the qPCR arrays. Significance and fold-change cut-offs were as follows: FDR or *p*-value ≤ 0.05, fold-changes ≥ 1.5-fold (XBP1s versus vehicle). Genes from the list in Table S8B that were not detected in either data set are not shown.

A

HeLa^XBP1s RNA-Seq Results

Max (2.68)   Average (0.00)   Min (−2.48)

B

Enrichment Plots: XBP1s vs Vehicle   **Enrichment Plots: Tg vs Vehicle**

GO Unfolded Protein Binding

q-value: 0.02
NES: 1.79

q-value: 0.13
NES: 1.64

GO Response to ER Stress

q-value: <0.001
NES: 2.35

q-value: <0.001
NES: 2.59

GO Intrinsic Apoptotic Signaling in Response to ER Stress

q-value: 0.29
NES: 1.50

q-value: <0.001
NES: 2.32

Figure S1: *RNA-Seq highlights distinctive effects of stress-mediated global UPR activation versus stress-independent XBP1s activation.* (A) Hierarchical clustering of differentially expressed protein-coding genes. HeLa^XBP1s cells were plated at a density of 6×10^5 cells per well in 6-well plates and allowed to adhere overnight. Cells were then treated with dox (1 μg/mL, 48 h), DMSO (0.1%, 24 h), or thapsigargin (Tg; 750 nM, 24 h), total RNA was extracted with the Qiagen RNeasy Plus Mini Kit, and samples were loaded on a HiSeq cartridge. Supervised hierarchical clustering was performed using Ward's method implemented in Spotfire 7.6.1 with row-centered log2 FPKM data for genes differentially expressed in at least one of the following comparisons: Tg versus vehicle, XBP1s vs vehicle, or XBP1s vs Tg. For the hierarchical clusting, differentially expressed genes were defined as those having absolute log(2) fold-change > 1 and FDR adjusted *p*-values < 0.05. (B) Selected GSEA enrichment plots for XBP1s versus vehicle and Tg versus vehicle showing distinct effects of XBP1s-activation versus global UPR activation with the ER stressor Tg.
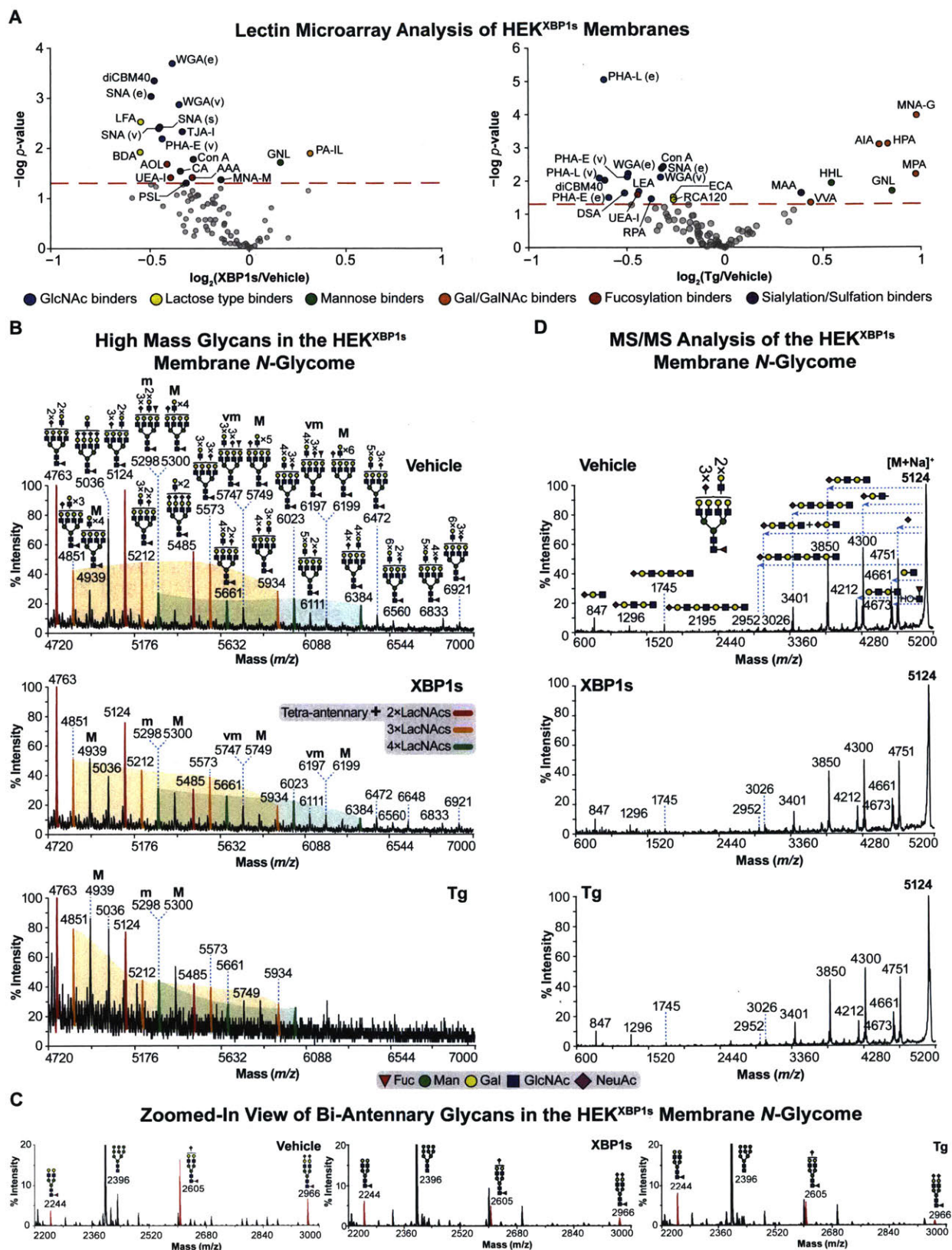
Western Blot - Conditioned Media

	full DMEM		Freestyle™	
XBP1s:	−	+	−	+
α-KDEL (low exposure)				— 80 kDa
α-KDEL (high exposure)				— 80 kDa
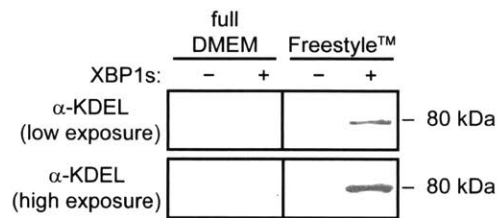
Figure S2: Freestyle^TM and complete DMEM media Western blot. Western blot of conditioned media from HEK^XBP1s cells cultured in complete DMEM or Freestyle™ media, shown at both low and high exposure. HEK^XBP1s cells were plated at a density of 6×10^5 cells per well on poly-D-lysine coated 6-well plates and allowed to adhere overnight. Cells were then treated with or without dox (1 μg/mL, 24 h), after which plates were washed three times with PBS (containing Ca^{2+} and Mg^{2+}) prior to treatment with either full DMEM (+10% FBS, L-glutamine) or Freestyle™ media, with or without 1 μg/mL dox. Conditioned media was collected for 48 h, spun at 1.5k RPM at 4 °C for 5 min to pellet cell debris, and then run on 4/8% SDS-PAGE polyacrylamide gels. Grp78 signal was detectable in Freestyle™ media samples, but only upon XBP1s activation.

A Lectin Microarray Analysis of HEK^XBP1s Membranes

B High Mass Glycans in the HEK^XBP1s Membrane *N*-Glycome

D MS/MS Analysis of the HEK^XBP1s Membrane *N*-Glycome

**C** Zoomed-In View of Bi-Antennary Glycans in the HEK^XBP1s Membrane *N*-Glycome

165

Figure S3: *Analysis of the HEK^{XBP1s} membrane glycoproteome.* (A) Lectin microarray analysis of membrane proteomes isolated from HEK^{XBP1s} cells showed decreased sialylation and a loss of bisecting GlcNAc upon XBP1s activation (left panel). Tg treatment (right panel) also showed decreases in sialylation and bisecting GlcNAc, along with an increase in core 1/3 O-glycans. Spot colors correspond to lectin sugar specificity, as indicated; the dotted line represents a significance cutoff of p-value $\leq$ 0.05 across three biological replicates. (B) MALDI-TOF MS spectra of high mass N-glycans. Red, yellow and green peaks correspond to tetra-antennary N-glycans with 2, 3 and 4 additional LacNAc repeats, respectively, with various levels of sialylation. Yellow and green shaded areas, manually inserted, highlight the shift towards fewer NeuAc residues. Structures outside a bracket have not been unequivocally defined, and "M," "m," and "vm," designations indicate major, minor and very minor abundances, respectively. Putative structures were assigned based on glycan composition, tandem mass spectrometry data, and knowledge of biosynthetic pathways. (C) Zoomed-in MALDI-TOF MS spectra indicating bi-antennary N-glycans (red peaks) with varying levels of sialylation. Putative structures were assigned based on glycan composition, tandem mass spectrometry data, and knowledge of biosynthetic pathways. (D) MALDI-TOF/TOF MS/MS spectra (selected from corresponding panels in Fig. S3B) for vehicle, XBP1s-activated, and Tg-treated HEK^{XBP1s} membrane proteomes. Spectra depict fragments of the $[M+Na]^+$ molecular ion found at m/z 5124. Structures above each horizontal dashed arrow indicate loss of the designated N-glycan sequence from the molecular ion; vertical dashed lines indicate the corresponding fragment ion peak. All molecular and fragment ions are $[M+Na]^+$. Cartoon structures were drawn according to the Consortium for Functional Glycomics (http://www.functionalglycomics.org) guidelines. For the molecular ion at m/z 5124, note the fragment ions at m/z 4300, 3850, 3401 and 2952 that correspond to loss of sialylated-LacNAc, sialylated-LacNAc$_2$, sialylated-LacNAc$_3$ and sialylated-LacNAc$_4$ respectively. The presence of such fragment ions suggests the existence of structural isomers corresponding mainly to tetra-antennary (and to a lesser extent tri-antennary) N-glycans. The relative abundance of these fragment ions did not shift between vehicle-treated, XBP1s-activated and Tg-treated HEK^{XBP1s} cells, further suggesting that the ratio of structural isomers remained relatively constant. Similar results were obtained for other ions (data not shown).
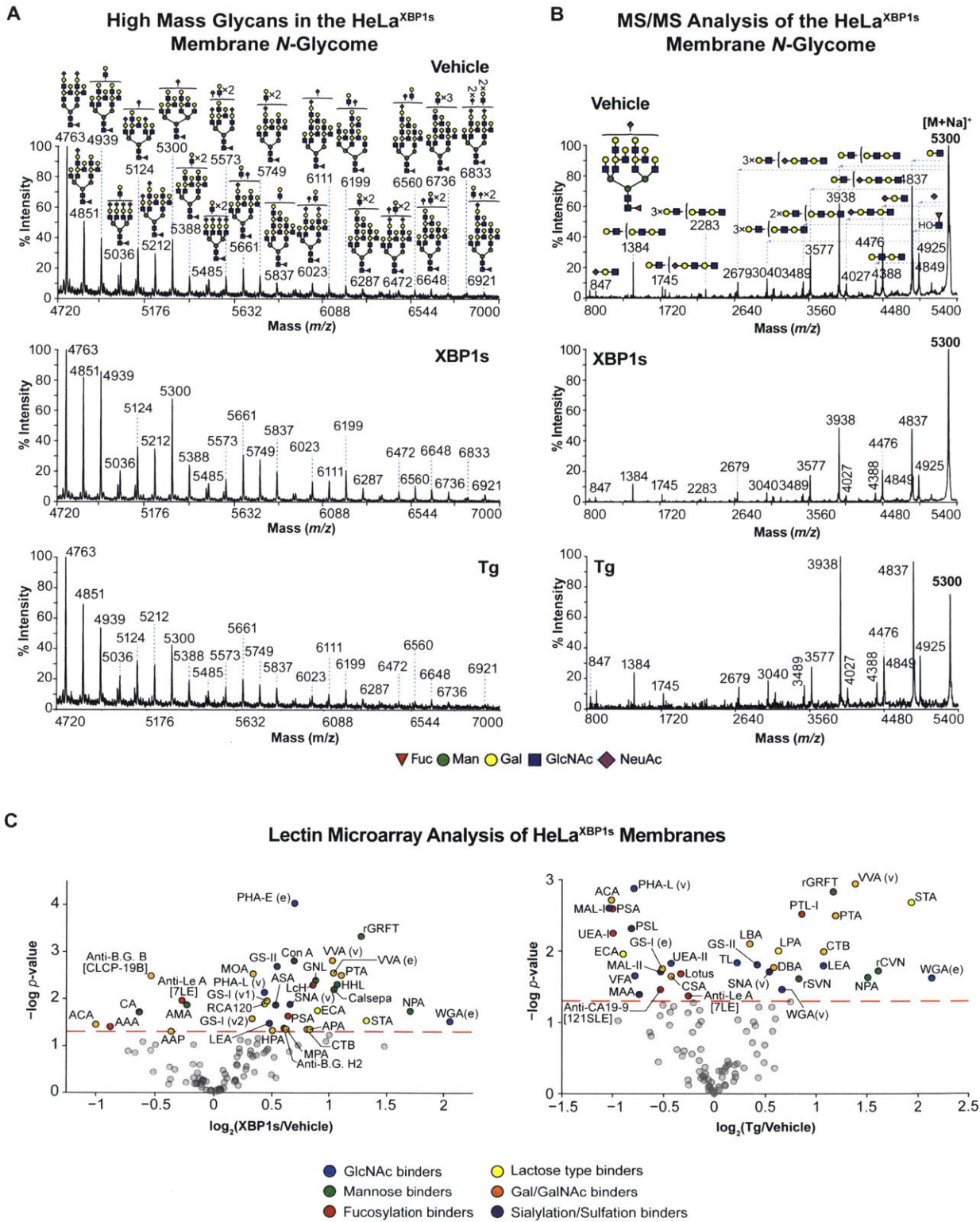
Figure S4

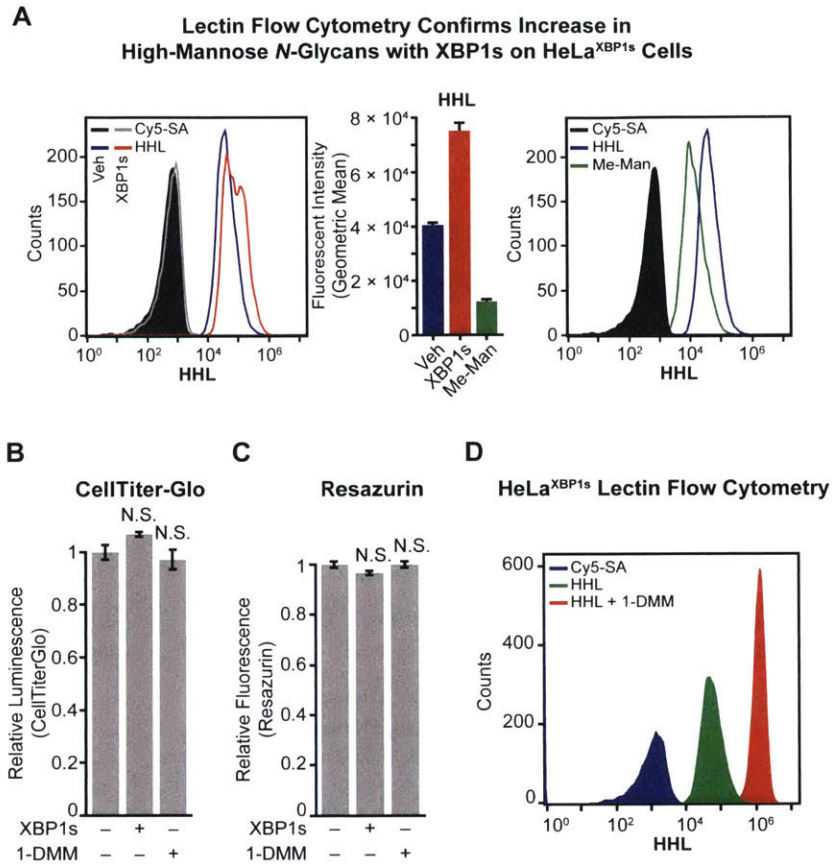**Lectin Microarray Analysis of Baseline Cellular *N*-Glycomes**

167
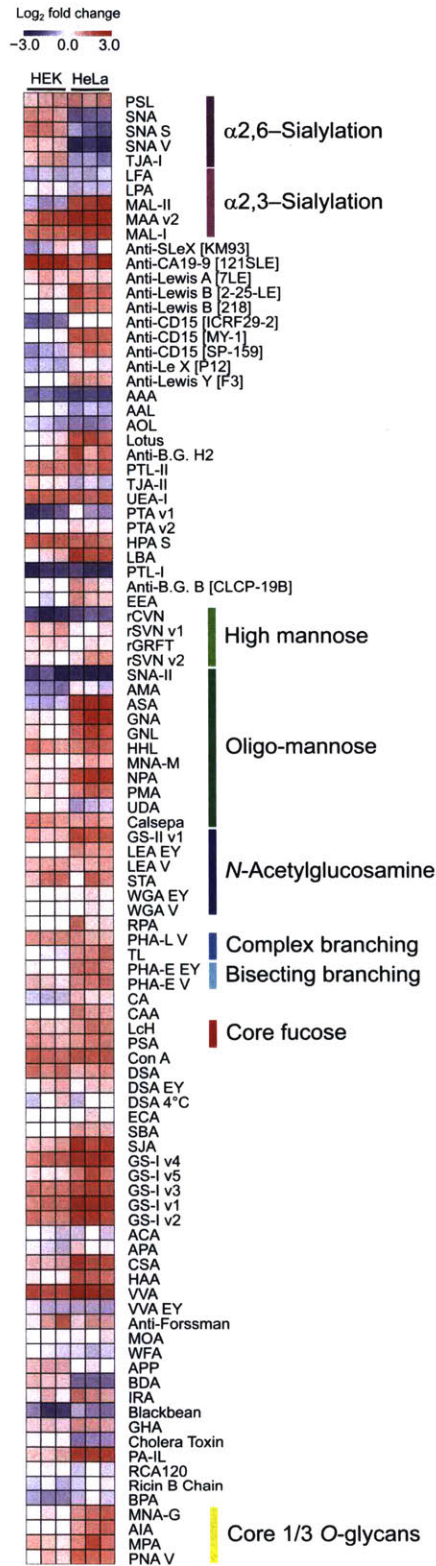
Figure S4: *Differences in baseline HEK^{XBP1s} and HeLa^{XBP1s} membrane glycoproteomes.* Heat map of lectin microarray data generated from isolated HEK^{XBP1s} and HeLa^{XBP1s} membrane glycoproteomes under vehicle-treated conditions. Samples were run on the same lectin microarray slide. Color intensity represents normalized log2 ratio data relative to a pooled (in this case, multi-cell line) sample reference. Each column represents one biological replicate of the indicated sample.

A High Mass Glycans in the HeLa^{XBP1s} Membrane *N*-Glycome

B MS/MS Analysis of the HeLa^{XBP1s} Membrane *N*-Glycome

▼ Fuc ● Man ○ Gal ■ GlcNAc ◆ NeuAc

C Lectin Microarray Analysis of HeLa^{XBP1s} Membranes

● GlcNAc binders ○ Lactose type binders
● Mannose binders ● Gal/GalNAc binders
● Fucosylation binders ● Sialylation/Sulfation binders

169

Figure S5: *Analysis of the HeLa^{XBP1s} membrane N-glycoproteome.* (A) MALDI-TOF MS spectra of high mass N-glycans and (B) MALDI-TOF/TOF MS/MS spectra (selected from corresponding panels in Fig. S5A) for vehicle, XBP1s-activated, and Tg-treated HeLa^{XBP1s} membrane proteomes. MALDI-TOF/TOF MS/MS spectra depict fragments of the [M+Na]$^+$ molecular ion found at *m/z* 5300. (C) Lectin microarray analysis of membrane proteomes isolated from HeLa^{XBP1s} cells showed an increase in high-mannose N-glycans upon XBP1s activation (left panel), as well as increases in core fucosylation and tetra-antennary N-glycans (β1,6-GlcNAc). Tg activation (right panel) showed a similar gain of high-mannose N-glycans, but not increased β1,6-GlcNAc branching. Spot colors correspond to lectin sugar specificity; the dotted line represents a significance cutoff of *p*-value ≤ 0.05 across three biological replicates.
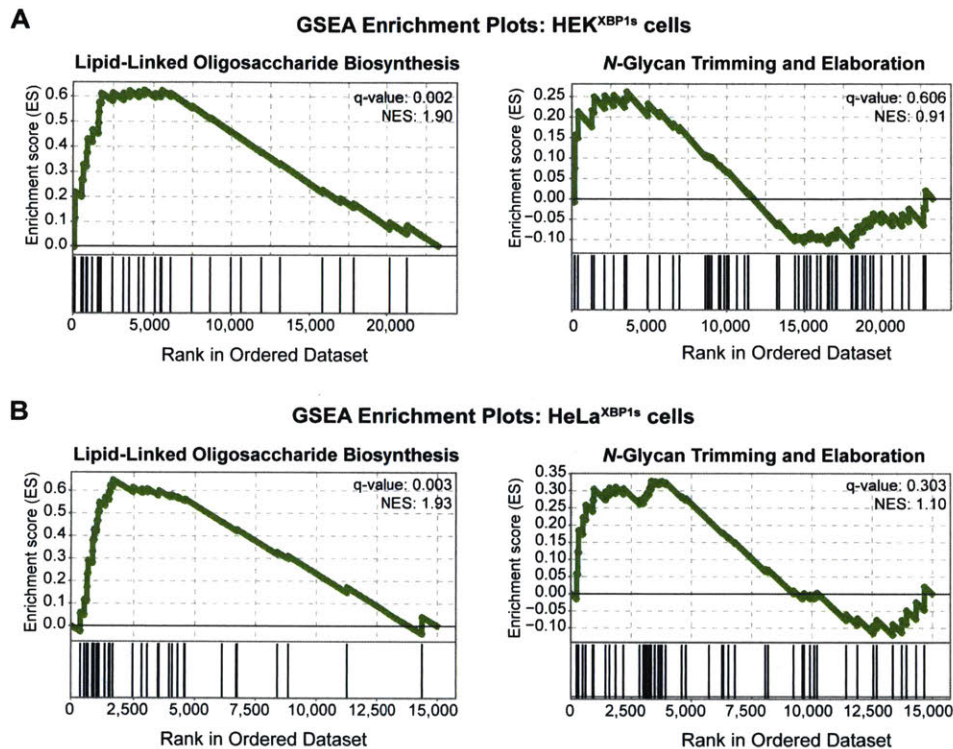
A

**Lectin Flow Cytometry Confirms Increase in
High-Mannose *N*-Glycans with XBP1s on HeLa^XBP1s Cells**

B CellTiter-Glo

C Resazurin

D HeLa^XBP1s Lectin Flow Cytometry

Figure S6: *Lectin flow cytometry confirms that XBP1s activation increases high-mannose N-glycans on HeLa^XBP1s cells.* (A) Lectin flow cytometry on live HeLa^XBP1s cells confirmed the increase in levels of high-mannose *N*-glycans on the cell surface (left panel, histogram; middle panel, quantification). Inhibitory sugar controls (right panel) were performed with methyl-mannose. Cell viability analyzed by CellTiter-Glo (B) or resazurin assay (C) was not significantly affected by treatment with dox (1 μg/mL, 72 h) or 1-deoxymannojirimycin (1-DMM, Sigma-Aldrich; 80 μg/mL, 24 h). 1×10^4 HeLa^XBP1s cells were seeded for each well 24 h prior to measurements. Error bars represent SEM from four replicates. (D) To verify that HHL recognized high-mannose glycans, HeLa^XBP1s cells were treated with 80 μg/mL of 1-DMM for 12 h to inhibit mannosidase-I and thus increase high-mannose N-glycan levels prior to analysis.

Lectin Microarray Analysis of the HEK^XBP1s Secretome
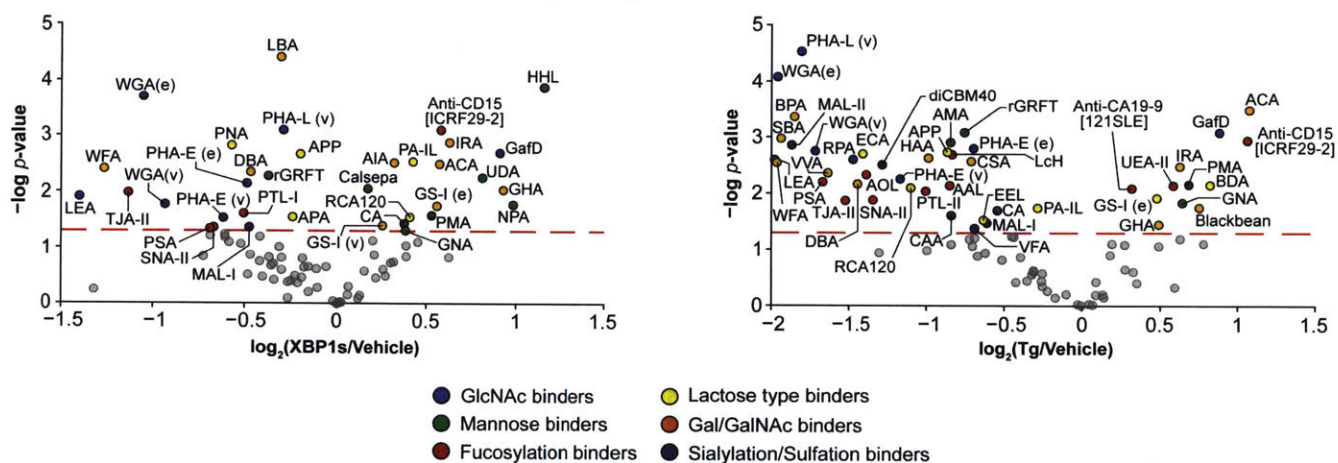
Figure S7: Analysis of the HEK^XBP1s secreted glycoproteome. Lectin microarray analysis of the HEK^XBP1s secretome showed a decrease in bisecting GlcNAc and an increase in early processed *N*-glycans upon XBP1s activation (left panel). Tg treatment (right panel) showed a decrease in multi-antennary *N*-glycans with no gain in high-mannose structures. Spot colors correspond to lectin sugar specificity; the dotted line represents a significance cutoff of *p*-value ≤ 0.05 across three biological replicates.

A

GSEA Enrichment Plots: HEK^{XBP1s} cells

Lipid-Linked Oligosaccharide Biosynthesis

q-value: 0.002
NES: 1.90

N-Glycan Trimming and Elaboration

q-value: 0.606
NES: 0.91

B

GSEA Enrichment Plots: HeLa^{XBP1s} cells

Lipid-Linked Oligosaccharide Biosynthesis

q-value: 0.003
NES: 1.93

N-Glycan Trimming and Elaboration

q-value: 0.303
NES: 1.10

Figure S8: *XBP1s impacts expression of genes involved in lipid-linked oligosaccharide biosynthesis.* (A) GSEA enrichment plots for HEK^{XBP1s} microarray data upon XBP1s activation using the custom gene sets listed in Table S1E. Standard GSEA was run using signal-to-noise ratio as a ranking metric. (B) Enrichment plots for HeLa^{XBP1s} RNA-Seq data upon XBP1s activation using the custom gene sets listed in Table S1E. Standard GSEA was run using signal-to-noise ratio as a ranking metric.

173
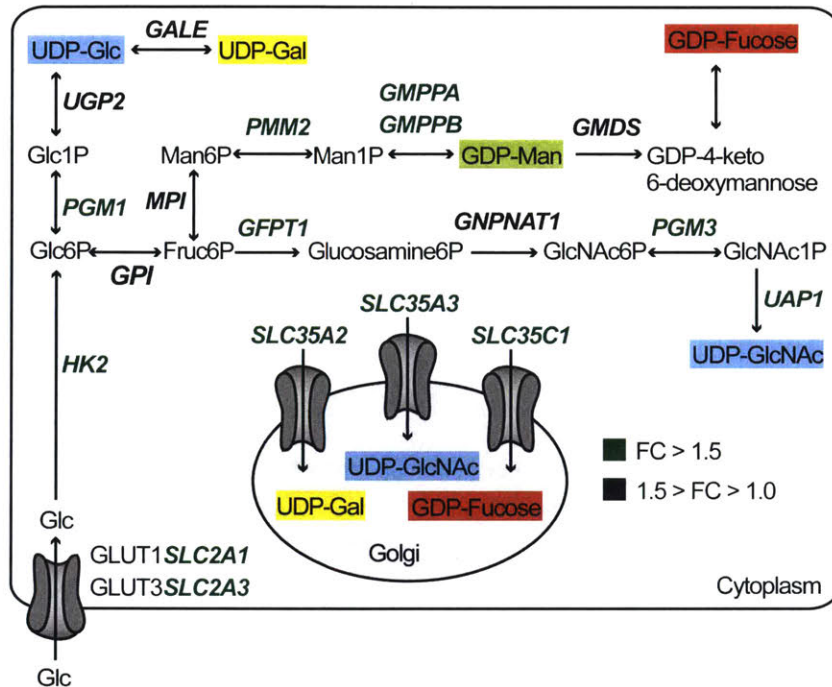
Figure S9: *XBP1s impacts expression of genes involved in nucleotide sugar donor biosynthesis.* Simplified biosynthetic pathway of nucleotide sugars depicting the enzymes/transporters that are significantly upregulated upon XBP1s activation in HeLa[XBP1s] cells. Transcripts shown in green (all with significant expression changes based on FDR or *p*-value ≤ 0.05) encode enzymes or transporters with a fold-change ≥ 1.5, while genes in black encode enzymes with fold-change between 1.0–1.5-fold. Note that *GNPDA1*, which together with *GNPDA2* catalyzes the reverse reaction of *GFPT1*, was significantly decreased (0.79-fold) in HeLa[XBP1s] cells (Table S8D). Biosynthetic pathway adapted from Ref. [8] and the KEGG database (*Homo sapiens* nucleotide sugar metabolism, http://www.kegg.jp).

Supporting References

[1] Dewal, M. B., DiChiara, A. S., Antonopoulos, A., Taylor, R. J., Harmon, C. J., Haslam, S. M., Dell, A., and Shoulders, M. D. (2015) XBP1s Links the Unfolded Protein Response to the Molecular Architecture of Mature N-Glycans, *Chem. Biol. 22*, 1301-1312.

[2] Shoulders, M. D., Ryno, L. M., Genereux, J. C., Moresco, J. J., Tu, P. G., Wu, C., Yates, J. R., 3rd, Su, A. I., Kelly, J. W., and Wiseman, R. L. (2013) Stress-independent activation of XBP1s and/or ATF6 reveals three functionally diverse ER proteostasis environments, *Cell Rep. 3*, 1279-1292.

[3] 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.

[4] Li, B., and Dewey, C. N. (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome, *BMC Bioinformatics 12*, 323.

[5] Love, M. I., Huber, W., and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2, *Genome Biol. 15*, 550.

[6] Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S., and Mesirov, J. P. (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles, *Proc. Natl. Acad. Sci. U.S.A. 102*, 15545-15550.

[7] Park, S. K., Venable, J. D., Xu, T., and Yates, J. R., 3rd. (2008) A quantitative analysis software tool for mass spectrometry-based proteomics, *Nat. Methods 5*, 319-322.

[8] Freeze, H. H., Hart, G. W., and Schnaar., R. L. (2017) Glycosylation Precursors, In *Essentials of Glycobiology* (Varki A, C. R., Esko JD, et al., Ed.) 3 ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY).

[9] Wong M.Y., Chen, K., Antonopoulos A., Kasper B.T., Dewal, M.B., Taylor, R.J., Whittaker, C.A., Hein, P.P., Dell, A., Genereux, J.C., Haslam, S.M., Mahal, L.K., and Shoulders, M.D. (2018) XBP1s activation can globally remodel *N*-glycan structure distribution patterns. Submitted.