Supporting Information

Quantitative structural proteomics unveils the conformational changes of proteins in cells under the ER stress

Kejun Yin[1], Ming Tong[1,#], Fangxu Sun[1,#], Ronghu Wu[1*]

[1] School of Chemistry and Biochemistry and the Petit Institute for Bioengineering and Bioscience, Georgia Institute of Technology, Atlanta, Georgia 30332, USA

* Corresponding author: Phone: 404-385-1515; Fax: 404-894-7452;

E-mail: ronghu.wu@chemistry.gatech.edu

Table of contents

Other supporting materials include the following Tables in Excel files:

Table S1. Quantified exposure rates in normal HEK293T cells using Cys-CPP (XLSX).

Table S2. GO enrichment analysis of proteins with differentially exposed cysteines (XLSX).

Table S3. Quantified exposure rates in Tm-treated HEK293T cells using Cys-CPP (XLSX).

Table S4. Comparison of proteins with stable and highly exposed cysteines (XLSX).

Table S5. The cysteine exposure rate changes in different cellular compartments under tunicamycin treatment (XLSX).

Table S6. Quantified exposure rates of newly synthesized proteins in normal HEK293T cells using Cys-CPP and pSILAC (XLSX).

Table S7. GO enrichment analysis comparing differentially unfolded pre-existing and newly synthesized proteins (XLSX).

Supporting Methods

Synthesis and purification of the cysteine-reactive probe

To a solution containing photo-cleavable biotin-tri-polyethylene glycol-N-hydroxysuccinimide ester (PC Biotin-PEG3-NHS) (4.5 mg, 0.0055 mmol) (Sigma-Aldrich) in 25 µL pre-degassed ACN, NEt_3 (0.90 µL, 0.0066 mmol) (Sigma-Aldrich) was added. It was cooled on ice, then further mixed with an ice-cold solution of maleimide-NH_2·TFA (1.3 mg, 0.0055 mmol) (Sigma-Aldrich) in 75 µL pre-degassed ACN. The reaction took place on an end-over-end rotator for 2 hours. The reaction crude was frozen, and then lyophilized in a speed-vac for 4 hours. Further purification was performed using high-pH reversed-phase HPLC. The crude was purified using a 4.6 × 250 mm 5 µm particle reversed-phase column (Waters) with a 30 min gradient of 3–30% ACN with 1% TFA. The eluted fraction containing the cysteine-reactive probe (i.e., photo-cleavable biotin-tri-polyethylene glycol-N-ethylmaleimide, PC Biotin-PEG-NEM) was frozen at -80 ºC and lyophilized. All procedures above were performed in the dark. Quality of the synthesis was monitored by LC-MS analysis of the purified probe using a 30-min gradient on a microcapillary column packed with C18 beads (Magic C18AQ, 3 µm, 200 Å, 75 µm x 16 cm, Michrom Bioresources) using a Dionex WPS-3000TPLRS autosampler 9 (UltiMate 3000 thermostatted Rapid Separation Pulled Loop Wellplate Sampler) with an UltiMate 3000 TPLRS LC coupled with a hybrid LTQ Orbitrap Elite mass spectrometer (Thermo Scientific).

Calculated exposure rate differences and unfolding stoichiometry

Cysteine with $R_{expo} > 1$ was likely caused by the experiment error. Cysteines with $R_{expo} > 1.05$ were discarded, and R_{expo} from 1.00 to 1.05 was assigned as 1.00. The following calculation is based on the assumption that cysteine can exist in completely folded or completely unfolded states (i.e., the two-state model).[1] For a given cysteine, the exposure rate within the folded protein is R_f, the fraction of the folded copies is P_f, the exposure rate within the unfolded protein is R_u, and the fraction of the unfolded copies is $P_u = 1 - P_f$. The exposure rate under normal conditions can be expressed as:

$$R_fP_f + R_uP_u = R_{expo}$$

Assuming that the folding stoichiometry difference for the protein in two different samples is x:

$$R_f(P_f - x) + R_u(P_u + x) = R_{expo}'$$

Combining the above two equations together, the result can be simplified to:

$$x(R_u - R_f) = R_{expo}' - R_{expo} = \Delta R_{expo}$$

For a fully denatured protein, cysteine should be fully exposed, which means $R_u = 1$. Then we have:

$$x(1 - R_f) = \Delta R_{expo}$$

This equation shows that the unfolding stoichiometry change x correlates with the exposure rate difference ΔR_{expo} and the exposure rate of its completely folded states R_f. The folded copy is the dominant species under non-stress conditions for most proteins;[2] thus, R_{expo} should be close to R_f. However, for the cysteine residents in the unstable region, this assumption is not valid.[1] To obtain a proximal implication on unfolding stoichiometry change x, the median exposure rate identified among all quantified cysteines under normal conditions was set as R_f (the median exposure rate = 0.48). Therefore, if we consider 30% unfolding stoichiometry change as a dramatic change (i.e., $x = 30\%$), $\Delta R_{expo} = 0.15$, and when $x = 10\%$ as a threshold value for no obvious changes, $\Delta R_{expo} = 0.05$.

Supporting Figures

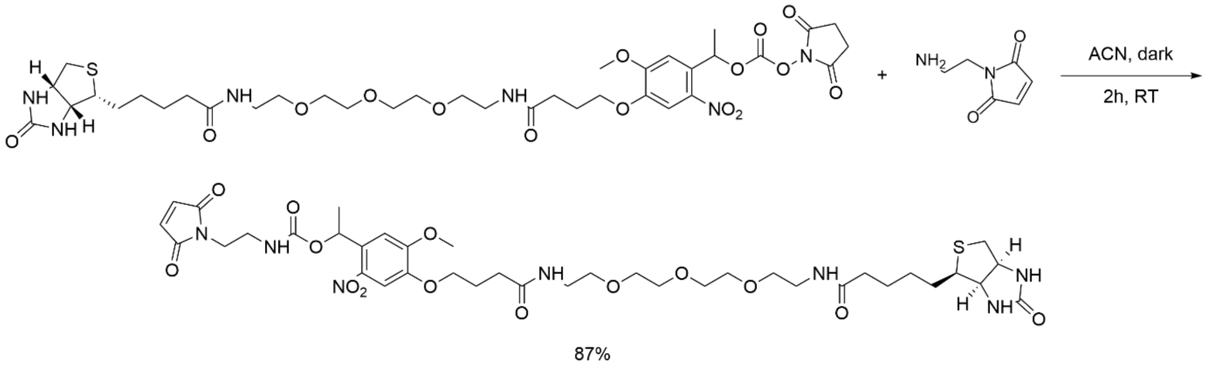

Fig. S1. Synthesis of the cysteine reactive probe.

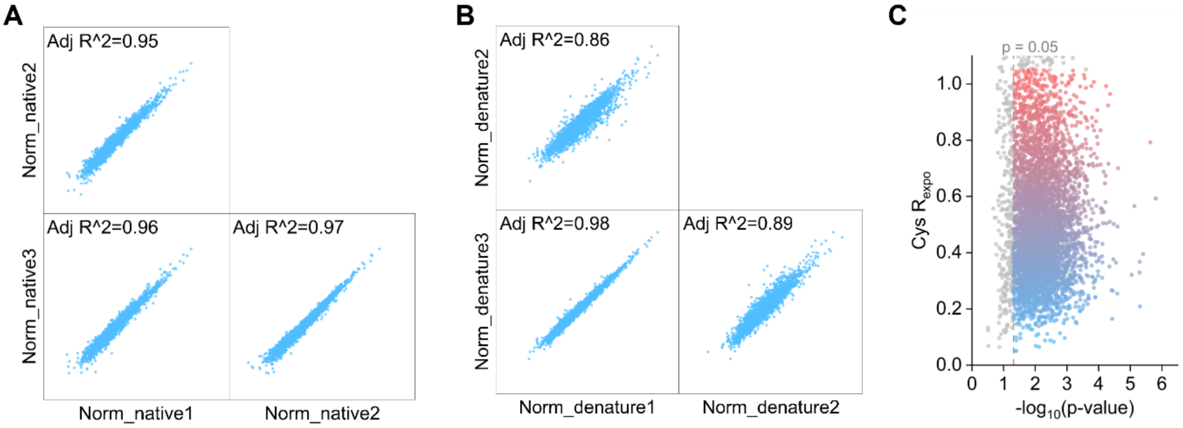

Fig. S2. (A-B) Comparison of the TMT reporter ion intensities in the normal HEK293T samples from different biological replicates. (C) The significantly quantified cysteine exposure rates of proteins in HEK293T ($n = 3$, $p < 0.05$).

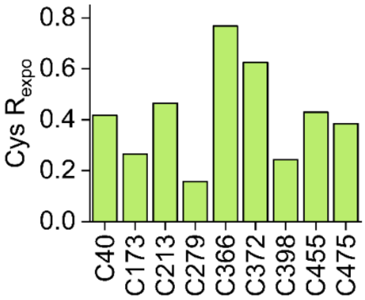

Fig. S3. The exposure rates of cysteines in CCT3.

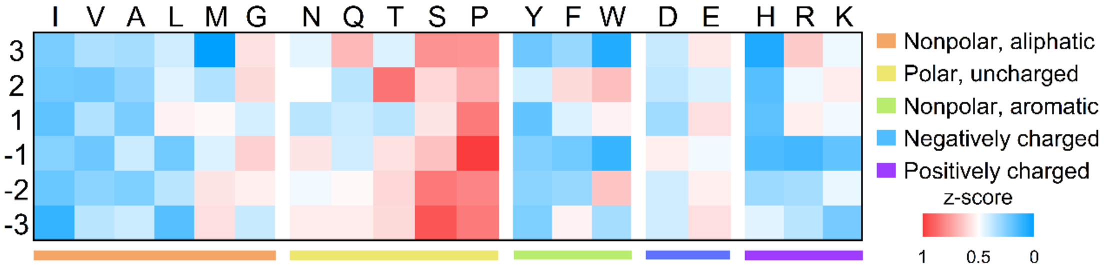

Fig. S4. The z-score transformed values of the cysteine exposure rates with different neighboring residues. Note that cysteine was excluded here.

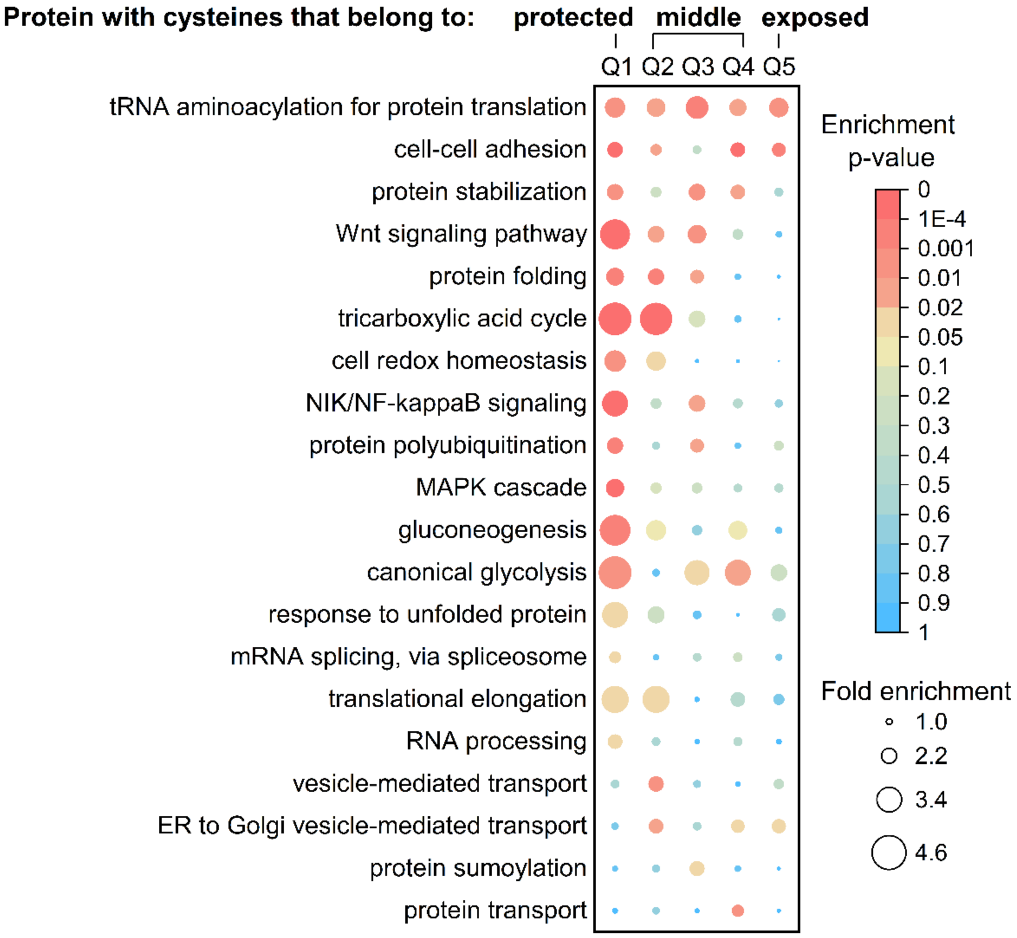

Fig. S5. GO enrichment analysis of proteins with cysteine in each segment.

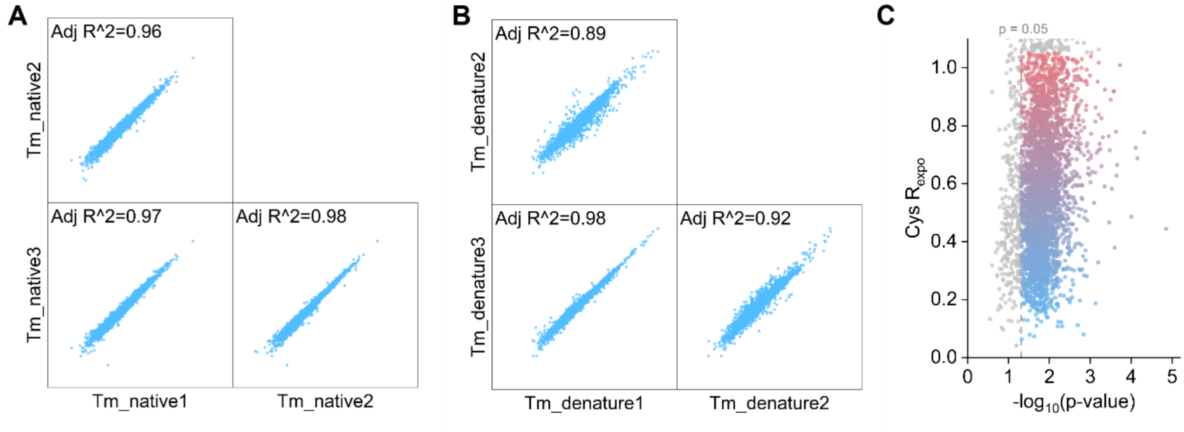

Fig. S6. (A-B) Comparison of the TMT reporter ion intensities in the Tm-treated HEK293T samples from different biological replicates. (C) The significantly quantified cysteine exposure rates of proteins in HEK293T cells treated with Tm ($n = 3$, $p < 0.05$).

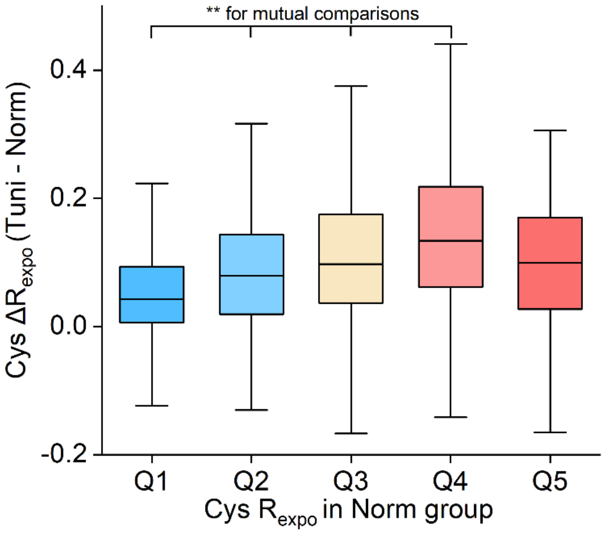

Fig. S7. The cysteine exposure rate changes in different quintile segments (***p < 0.001, **p < 0.01, Kolmogorov–Smirnov test).

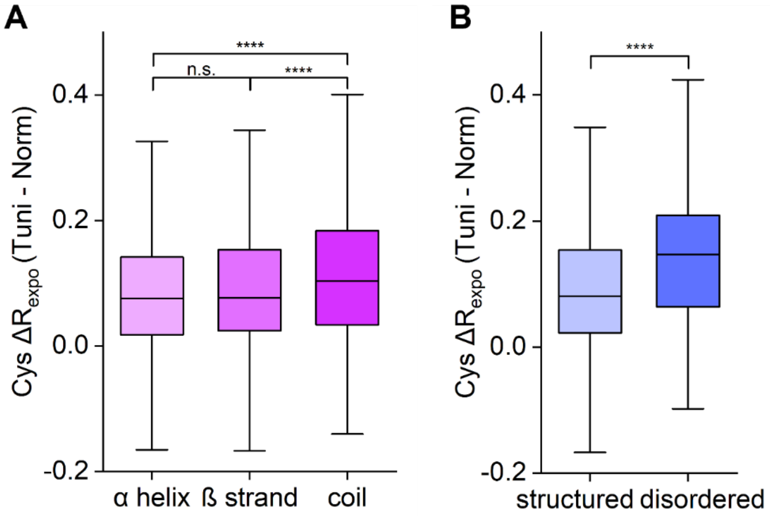

Fig. S8. Comparison of the exposure rates of cysteines in different secondary structures (A) and predicted disorderness (B) (****p < 0.0001, Kolmogorov–Smirnov test).

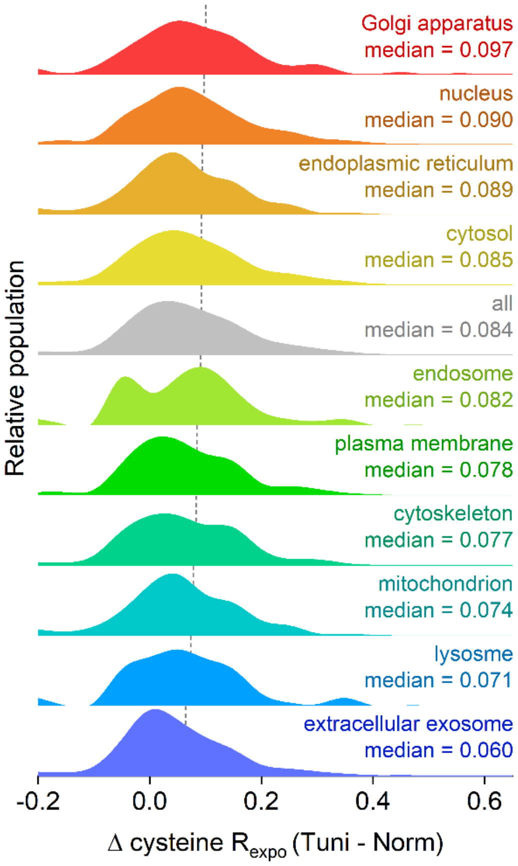

Fig. S9. Comparing the distributions of the protein exposure rate differences in different cellular compartments caused by Tm. The median values are shown as dash lines in the figure.

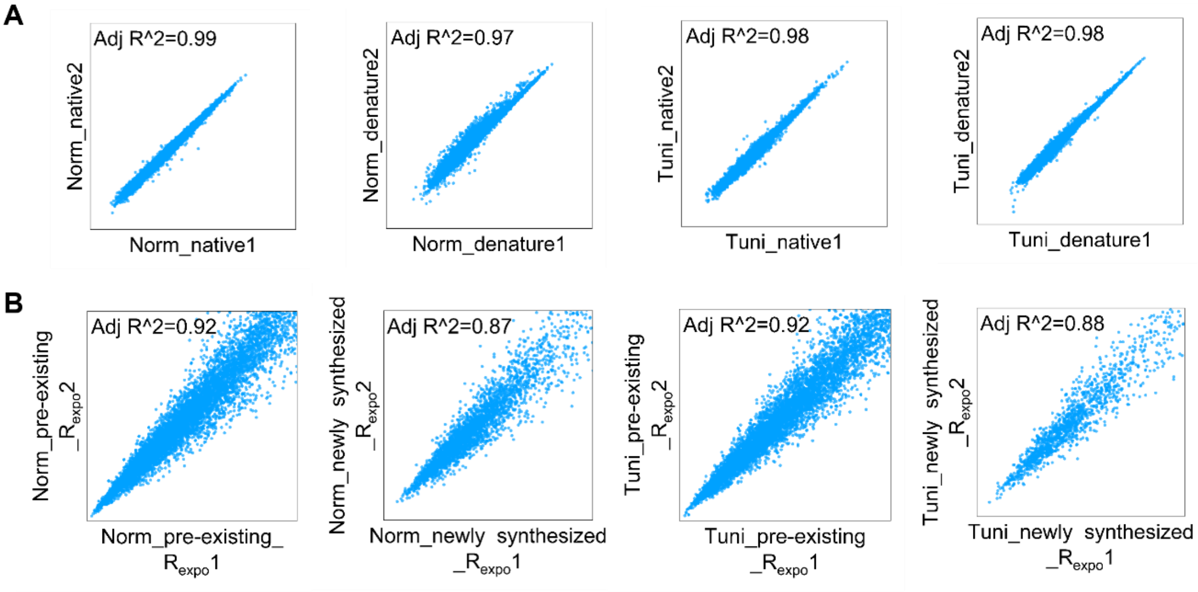

Fig. S10. (A) Comparison of the TMT reporter ion intensities in the pSILAC samples from different biological replicates. (B) Comparison of the calculated exposure rates in the pSILAC samples from different biological replicates.

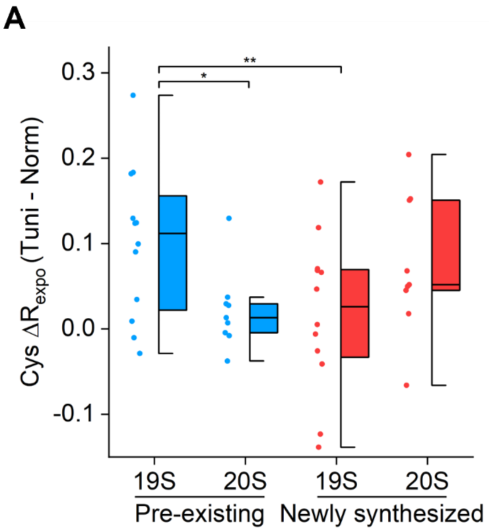

Fig. S11. (A) Comparison of the cysteine exposure rate changes between pre-existing and newly synthesized proteins in the proteasome regulatory (19S) and core (20S) subunits. (**p<0.01, *p < 0.05, Kolmogorov–Smirnov test). (B) Comparison of the cysteine exposure rate changes between the pre-existing cytoplasmic fibrils proteins and other pre-existing proteins in the nuclear pore complex. (**p<0.01, Kolmogorov–Smirnov test).

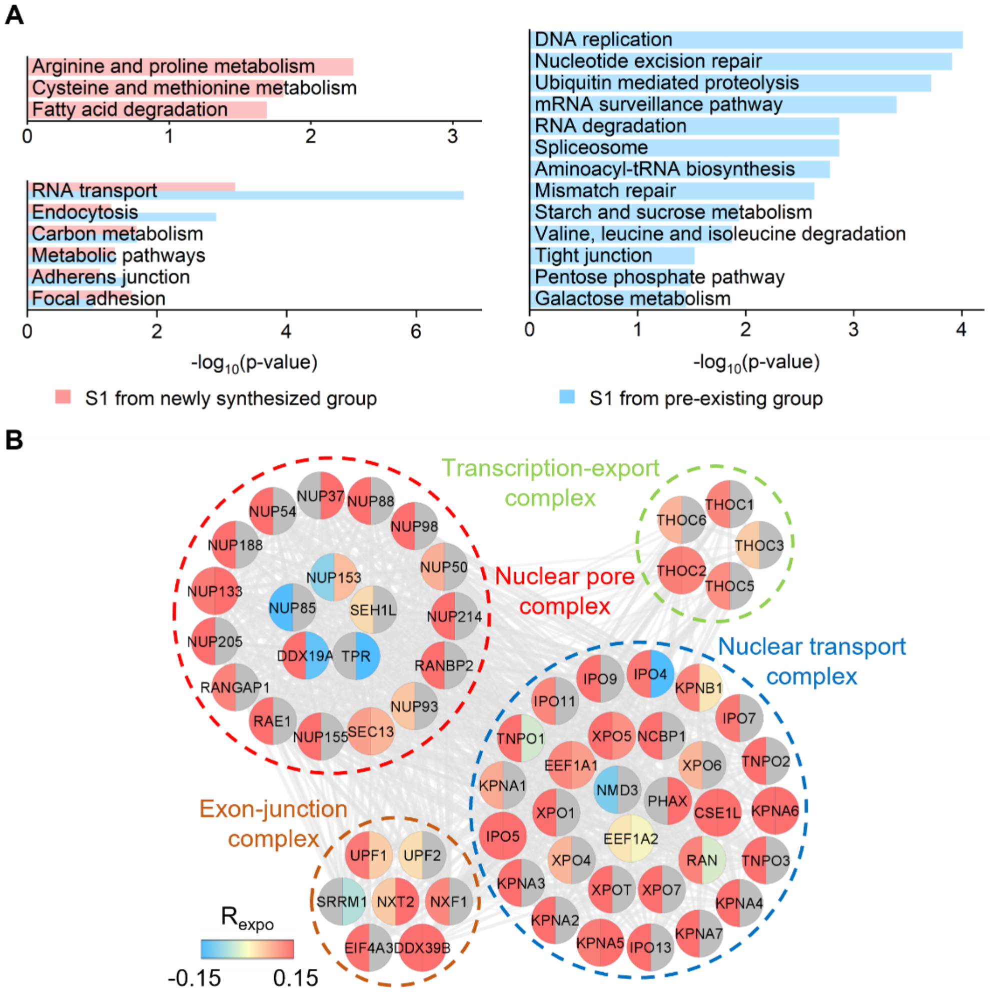

Fig. S12. Pathway analysis for proteins with structure changes under the Tm treatment. (A) Enriched KEGG pathways and their enrichment scores among highly damaged newly synthesized proteins (in red) and pre-existing proteins (in blue) in cells with the Tm-induced stress. (B) Effects of Tm on the structural stability of proteins related to RNA transport. The color represents the segment that the cysteine exposure rates belong to. The left side of each node indicates the maximum cysteine exposure rate change in pre-existing copies, and the right side stands for newly synthesized copies.

[#] Present address: #M.T.: Novo Nordisk research centre, Beijing, 100102, China; #F.S.: Biogen, Cambridge, 02142, MA, USA.

References

(1) Walker, E. J.; Bettinger, J. Q.; Welle, K. A.; Hryhorenko, J. R.; Ghaemmaghami, S., Global analysis of methionine oxidation provides a census of folding stabilities for the human proteome. *Proc. Natl. Acad. Sci. U. S. A.* **2019,** *116* (13), 6081-6090.

(2) Schwanhausser, B.; Busse, D.; Li, N.; Dittmar, G.; Schuchhardt, J.; Wolf, J.; Chen, W.; Selbach, M., Global quantification of mammalian gene expression control. *Nature* **2011,** *473* (7347), 337-42.