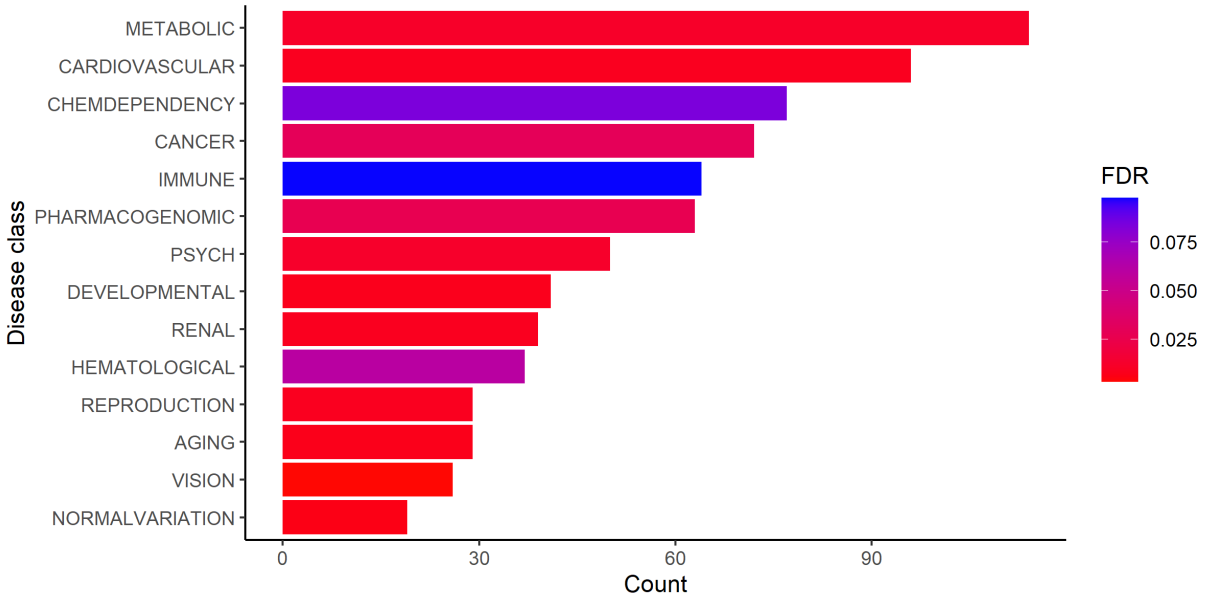
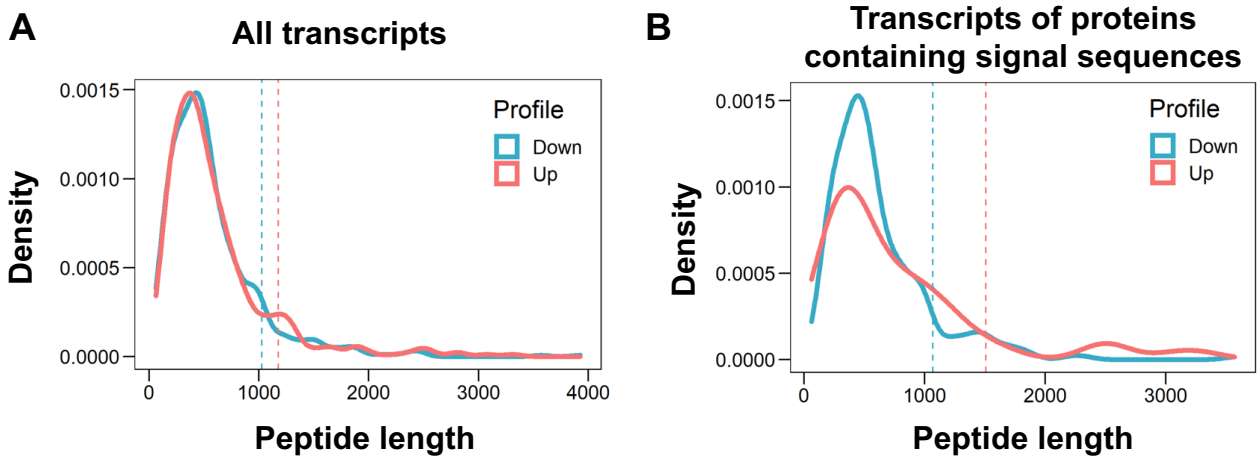
Supplementary Figures

Supplementary Figure S1. Molecular functions of the proteins encoded by differently expressed transcripts in SRP54 knockdown cells. Significantly downregulated transcripts in SRP54 depleted cells show enrichment in molecular processes associated with extracellular proteins including cell adhesion, response to growth factors, extracellular matrix and structure organization, and others (left panel). Significantly upregulated transcripts are associated with protein folding modulators, ubiquitination, and others (right panel). P-value adjusted by Benjamini-Hochberg method (p.adjust). Transcript ratio is the percentage of total differentially expressed transcript in the given gene ontology term. All transcripts expressed in HeLa cells were used as background (universe). Dot plots were generated using the enrichGO function of the R package clusterProfiler (version 3.0.4).



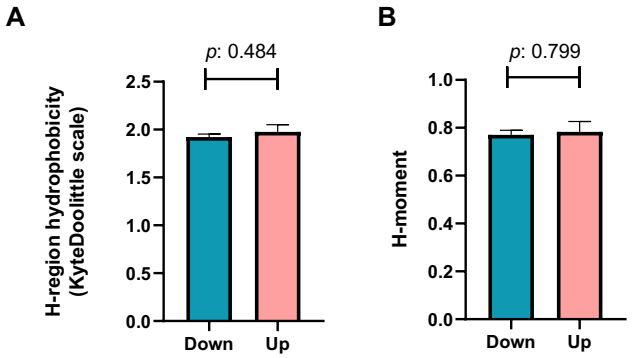
Supplementary Figure S2. Majority of downregulated transcripts encoding proteins with signal sequences in SRP54 knockdown cells are associated with human diseases. DAVID 6.8 bioinformatic resource (https://david.ncifcrf.gov/) was used to identify the number of transcripts (count, X-axis) per associated disease class (Y-axis). The analysis was based on the Genetic Association Database (GAD). False discovery rate (FDR) is represented as gradient scale. Disease classes are highlighted by the level of significance from less (blue) to more (red) significant. Analysis demonstrates that 236 (75.4%) of the total 313 downregulated transcripts with annotated signal sequences are connected with human diseases.



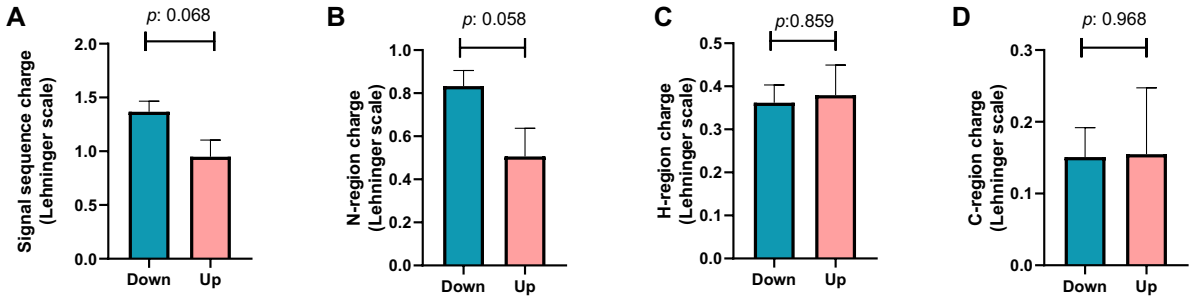
Supplementary Figure S3. Analysis of proteins encoded by differently expressed transcripts in SRP54 knockdown cells regarding their overall protein length. Distributions of differentially expressed transcripts per associated protein length are shown. X-axis represents the protein length reported in UniProt data base. Y-axis represent the abundance of downregulated (Down) or upregulated transcripts (Up) per peptide length as the kernel density score. Blue and red dashed lines mark 90% of downregulated and upregulated transcripts, respectively.

(A) Analysis of total differentially expressed transcripts shows that upregulated and downregulated genes have similar distribution across protein length.

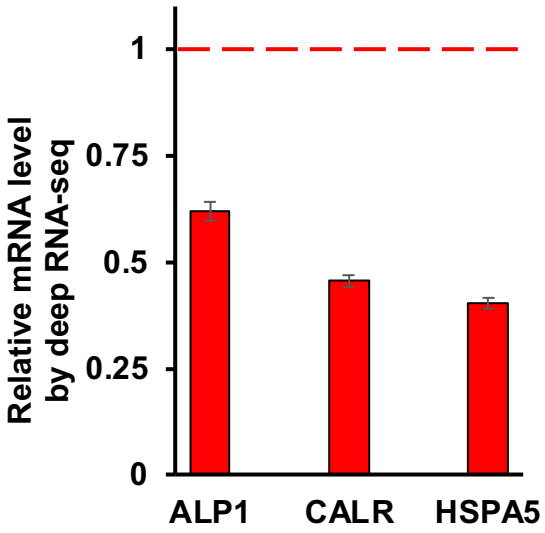
(B) Analysis of differentially expressed transcripts encoding proteins containing signal sequences shows that upregulated transcripts show a wider distribution of the proteins' lengths in comparison with downregulated transcripts.



Supplementary Figure S4. Upregulated and downregulated transcripts encoding proteins with signal sequences in SRP54 depleted cells show similar hydrophobicity. Comparison of hydrophobicity between upregulated (Up) and downregulated (Down) proteins containing signal sequences after SRP54 knockdown shown by the use of two methods.
(A) Hydrophobicity index of the hydrophobic cores of the signal sequences (H-regions) by using Kyte-Doolittle scale. Unpaired T-test two tailed was used, t=0.7001, df=343.
(B) Hydrophobic moment (H-moment scale) for each H-region sequence as a quantitative measure of the amphiphilicity perpendicular to the axis of an a-helix. Unpaired T-test two tailed was used, t=0.2538, df=343.



Supplementary Figure S5. The Signal Sequence N-region of downregulated proteins trends to show a higher positive net charge.
The net charge per amino acid sequence was estimated by using Lehninger scale to compare downregulated (Down) and upregulated (Up) signal sequence containing proteins after SRP54 KD.
(A) Net charge estimated for the whole signal sequences. Unpaired T-test two tailed was used, t=1.830, df=343.
(B) Charge estimated for N-terminal region (N-region) of the signal sequences. Unpaired T-test two tailed was used, t=1.896, df=343.
(C) Net charge estimated for the central hydrophobic domain (H-region) of the signal sequences. Unpaired T-test two tailed was used, t=0.1766, df=343.
(D) Net charge estimated for C-terminal region (C-region) of the signal sequences. Unpaired T-test two tailed was used, t=0.03908, df=343.
Signal sequences were obtained from UniProt data base. The regions of signal sequences were defined by SignalP-6.0 algorithm. P-value (*p)*.



Supplementary Figure S6. Verification of mRNA level of previously identified proteins controlled by RAPP. Relative mRNA levels determined by Deep RNA-seq analysis, shown as a ratio between transcript counts in SRP54 KD and control cells. ALP1 is intestinal alkaline phosphatase, HSPA5 (or BIP) is heat shock protein family A (Hsp70) member 5, CALR is calreticulin.