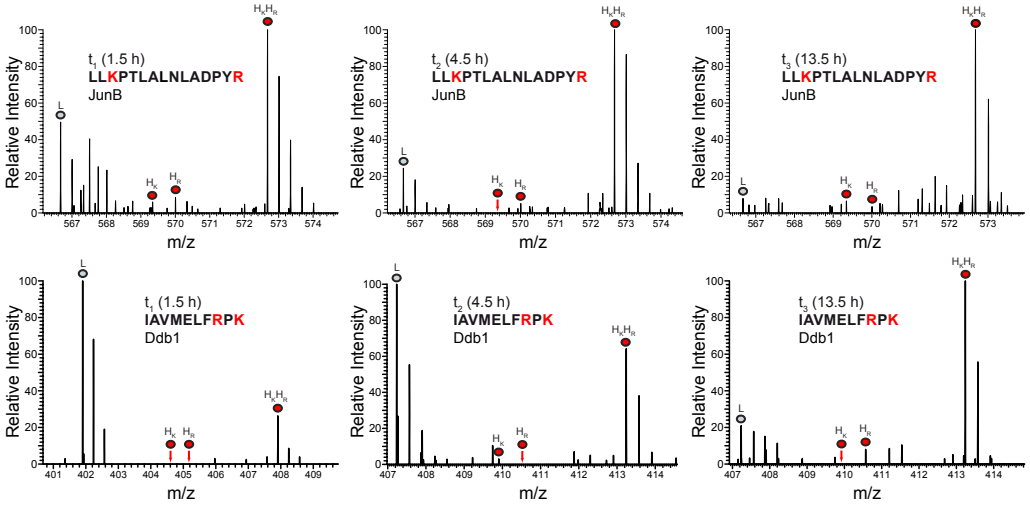
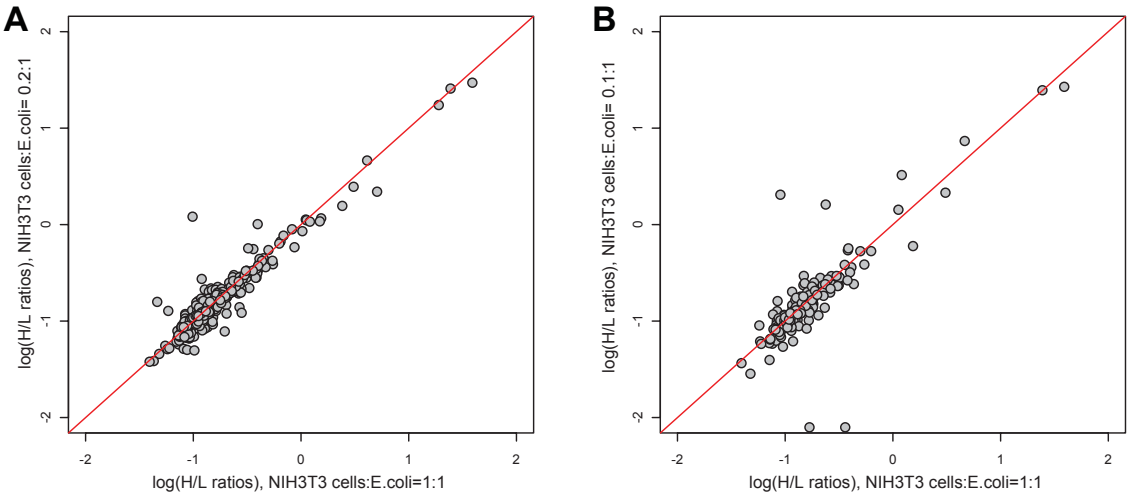
doi:10.1038/nature10098

Fig. S1



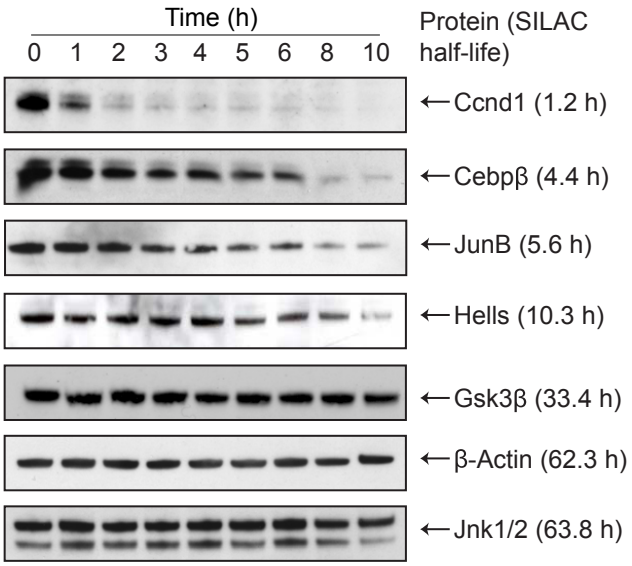
Protein turnover can be under-estimated due to the presence of light (L) amino acids that remain in the precursor pool after transferring cells to heavy (H) SILAC medium (e.g. by recycling). The degree of L amino acids incorporated after transferring cells to H medium can be evaluated by inspection of mass spectra of missed cleaved tryptic peptides which contain two labelled amino acids (R and K). In case of complete labelling newly synthesized peptides should contain only heavy K and heavy R (H_KH_R). Additional peaks containing only one heavy amino acid (H_K or H_R) are indicative of light amino acids incorporated after medium exchange. The figure shows mass spectra of missed-cleaved peptides from a high and a medium turnover protein at all three time points. In some cases, partially heavy labeled peptides are observed. However, these peaks are very small compared to the completely labeled peak and do not appreciably affect measured protein half-lives.

Fig. S2



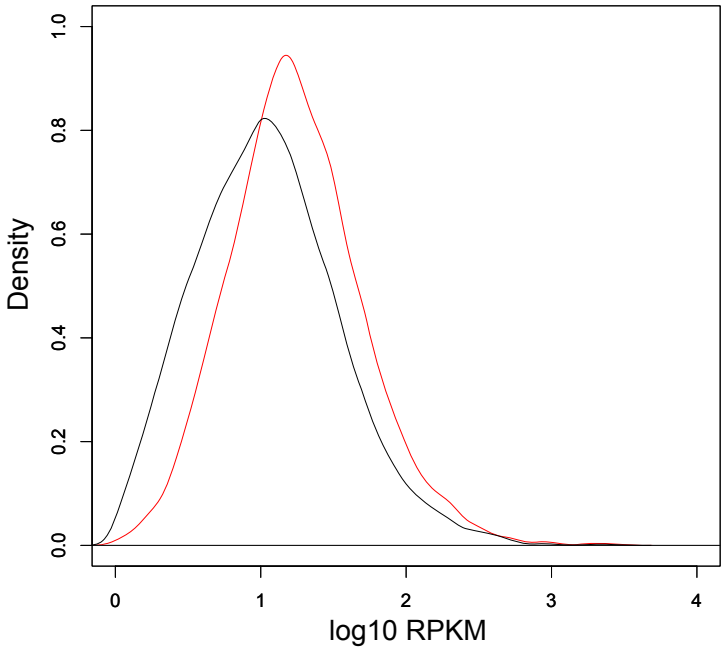
Effect of protein abundance on H/L ratios measured by mass spectrometry. NIH3T3 cells were pulsed-labelled with heavy amino acids for 13.5 h. Different amounts of this NIH3T3 lysate were mixed with the same amount of E.coli lysate to simulate different protein abundance. Comparing the H/L ratios for NIH3T3 proteins in the 1:1 mixture with the 0.2:1 (A) and 0.1:1 (B) mixture clearly revealed that H/L ratio measurements are not affected by protein abundance.

Fig. S3



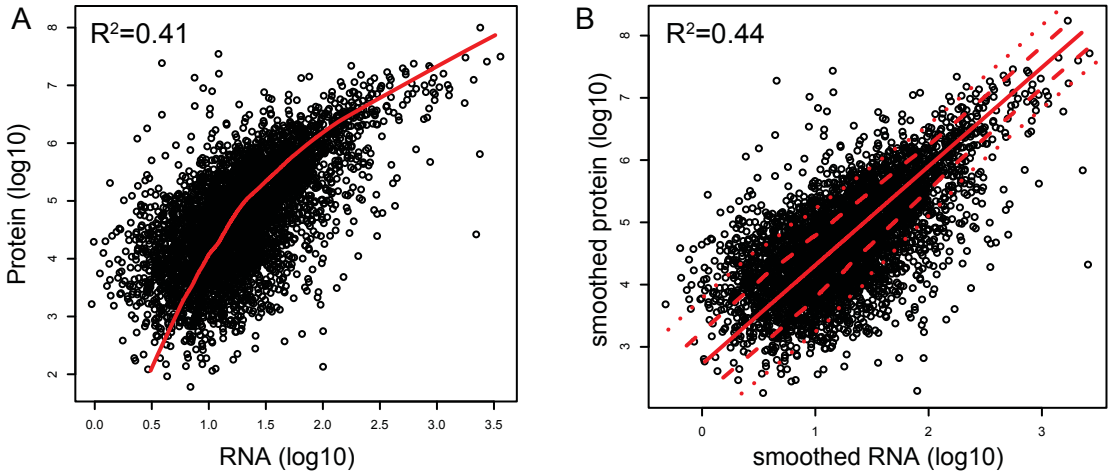
Comparison of protein half-lives measured by SILAC and traditional cycloheximide-chase experiments. Half-lives measured by SILAC (values are given in parentheses) are in good agreement with Western Blots in all cases.

Fig. S4



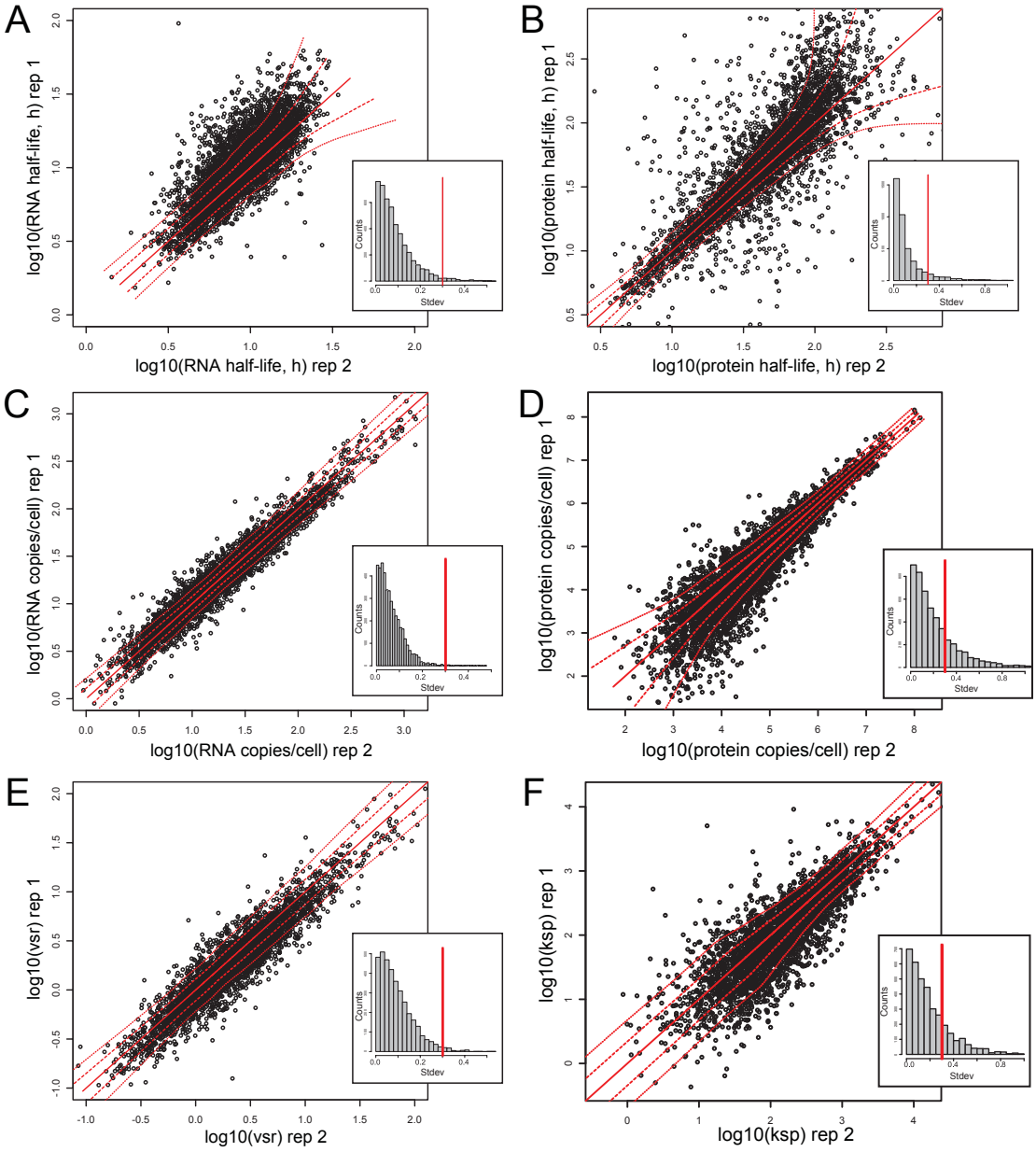
Kernel density estimates of mRNA levels (given as RPKM values) for all genes (black) and genes detected at the protein level (red). The slight shift indicates a mild detection bias against lowly expressed proteins. All subsequent calculations were therefore done on the set of genes detected at both the protein and mRNA level.

Fig. S5



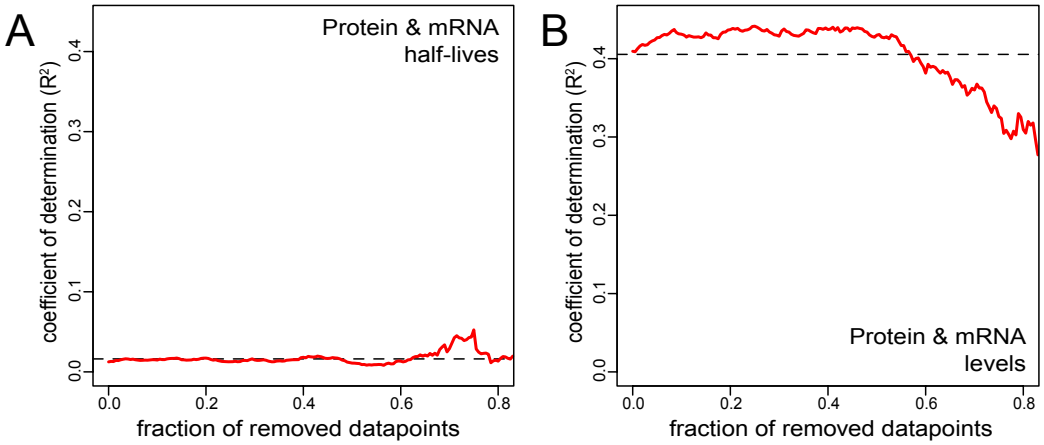
Effect of LOWESS (locally weighted scatterplot smoothing) transformation on the correlation of mRNA and protein levels. (A) Determination of non-linear transfer function by employing LOWESS smoothing (LOWESS function as implemented in R-statistics framework with parameter f set to 0.25). Procedure as suggested by Terry Speed (http://www.stat.berkeley.edu/tech-reports/589.pdf). (B), data after transformation: convincing linear relation and rather stable variances are achieved by applying lowess transformation. Even this very general transformation results in rather modest improvement of correlation as indicated by a coefficient of determination of R^2=0.44. Solid lines indicate averaged means. Dashed and dotted lines indicate one and two standard deviations, respectively.

Fig. S6



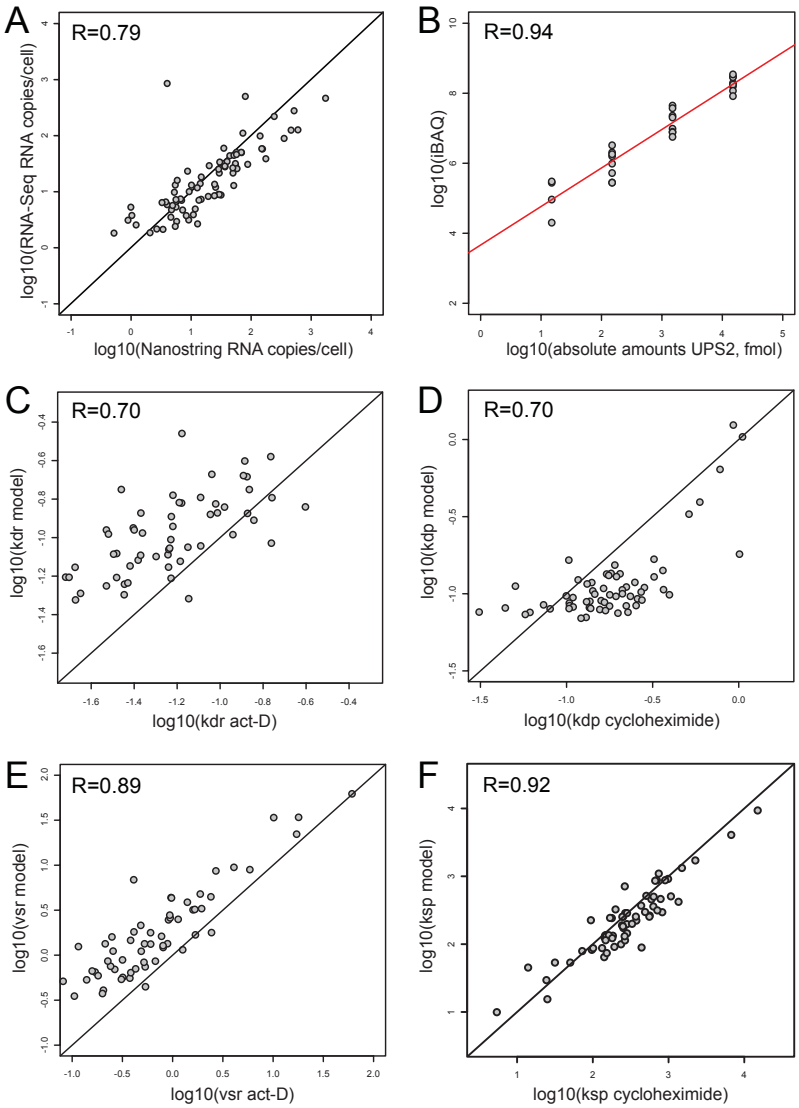
Reproducibility of biological replicates. (A) mRNA half-lives, (B) protein half-lives, (C) mRNA copies, (D) protein copies, (E) transcription rates and (F) translation rate constants. Except for calibrating RNA and protein abundance data as described in Supplementary Methods, we did not use any smoothing transformation (like e.g. lowess regression) for reducing variability in our data. Each pair of values is simply mapped to its mean and its standard deviation. As a guide to the eye we report averaged means (solid lines) and standard deviations (dashed and dotted lines for one and two stdevs, respectively). Lines were obtained using the lowess scatter smoother on the point-wise stdevs giving us a local pooled estimate. The small insets show the distribution of stdevs in a histogram, and the line indicates the proportion of standard deviations below a factor two. For more details we report standard deviations in three typical signal ranges (low, medium, high) in Supplementary Table 1.

Fig. S7



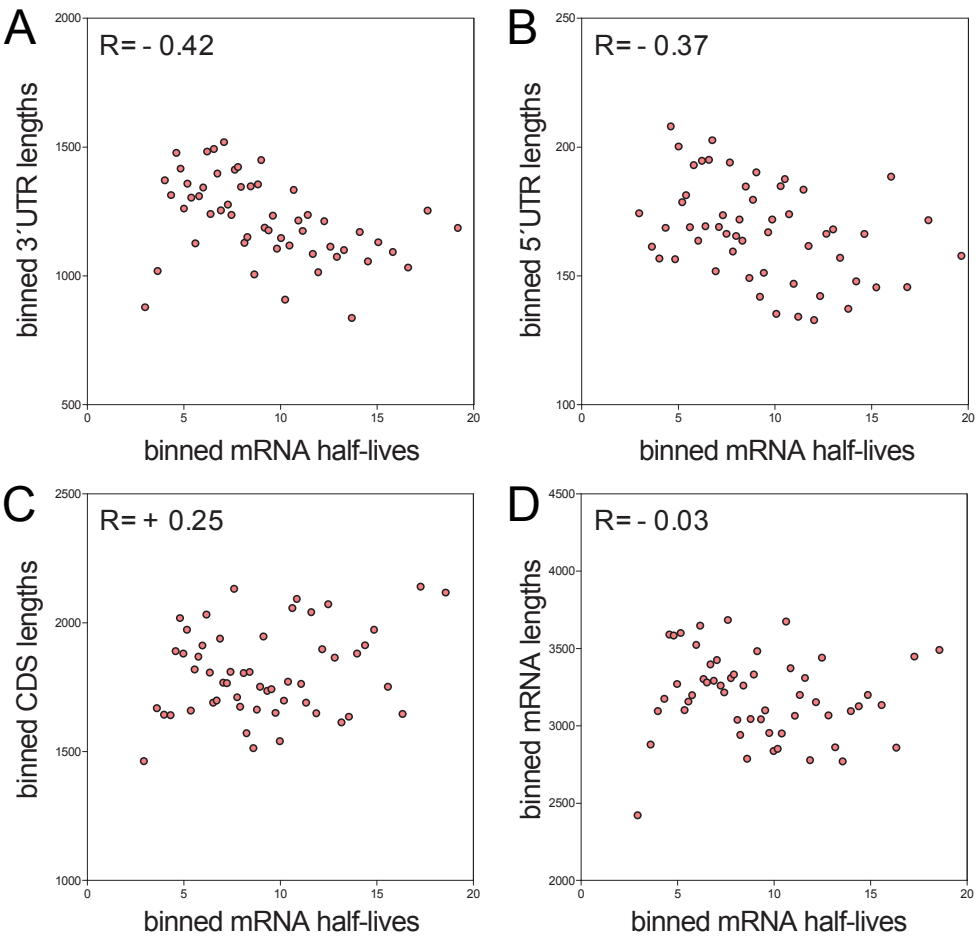
Impact of experimental noise on the correlation between mRNAs and proteins. To test if techical noise affects the observed correlation we consecutively discarded genes with the highest variability between both replicates (i.e. outliers in Fig. S6 A-D). For the remaining fraction we investigated the correlation between mRNA and protein half-lives (A) and levels (B) again. Removal of less reproducible data points did not significantly increase correlations of mRNA and protein half-lives (A) or levels (B).

Fig. S8



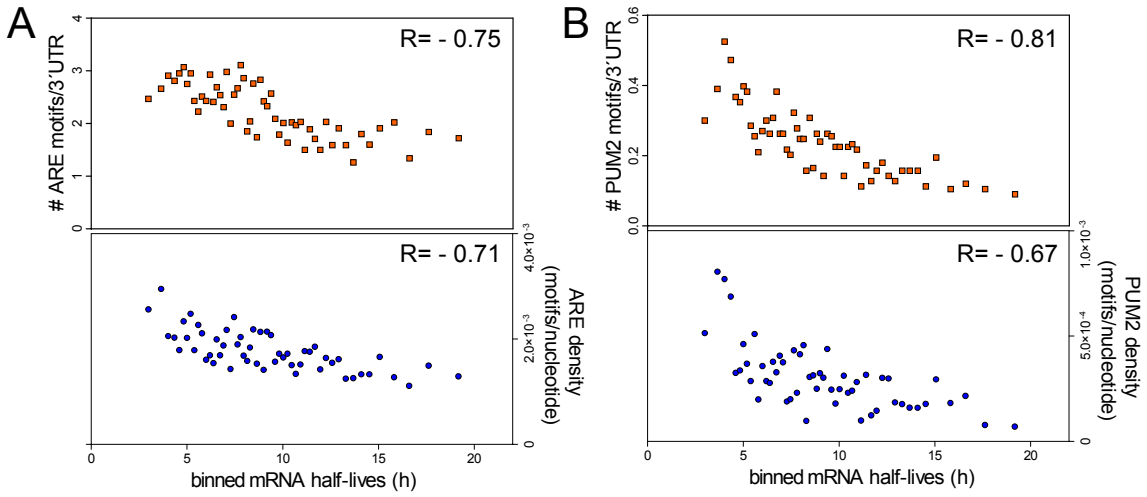
Validation of mRNA and protein levels, degradation and synthesis rates. (A) A total of 79 mRNAs were independently quantified using the NanoString technology and plotted against calculated mRNA levels based on the RNA sequencing results. The linear correlation holds true over the entire range of calculated mRNA copy numbers. (B) Correlation between absolute amounts of the spiked UPS2 standard proteins and the iBAQ approach. The red line denotes linear regression (y = 1.10 x + 3.67) used to scale the iBAQ values (see Supplementary Methods for details). Validation of mRNA degradation (C, kdr) and transcription rates (E, vsr) using actinomycin D chase analysis to block de novo mRNA synthesis (see Supplementary Methods for details). Differences in mRNA abundance upon actinomycin D treatment were compared to the untreated control and used for determination of rates. Similarly, protein degradation rates (D, kdp) and translation rate constants (F, ksp) were independently measured via cycloheximide-chase analysis and conventional SILAC (see Supplementary Methods for details). Note that these results are compromised by toxic side effects of the drugs (C-F). Nevertheless, for short-lived mRNAs and proteins, degradation rates based on drug treatment agreed well with rates determined by pulse labeling (C, D). Overall, correlation of transcription rates was good but had a notable systematic shift (E). This may be due to stress-induced mRNA stabilization during actinomycin D treatment. Translation rate constants obtained by pulse labeling and cycloheximide-chase agreed well (F). Only rates of proteins with half-lives < 10 h are shown.

Fig. S9



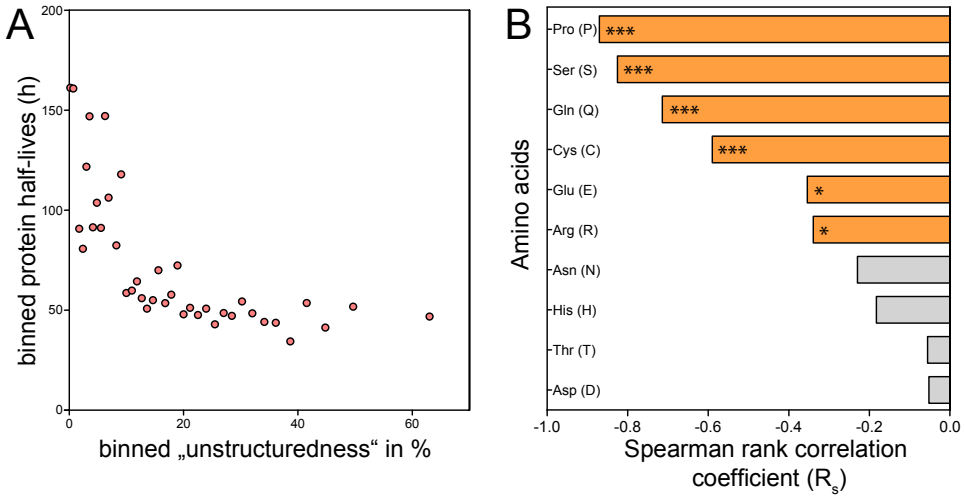
Correlation of mouse (A) 3´ untranslated region (UTR), (B) 5´UTR, (C) coding sequence (CDS) and (D) mRNA lengths with mRNA stability. The average lengths of respective sequences (NCBI Build 37/mm9) in bins of 100 genes were plotted against averaged mRNA half-lives of the same bins. While 3´UTR and 5´UTR lengths show a clear negative correlation to mRNA stability, the CDS length is positively correlated. Plotting overall mRNA lengths yield the net outcome of these opposing effects.

Fig. S10



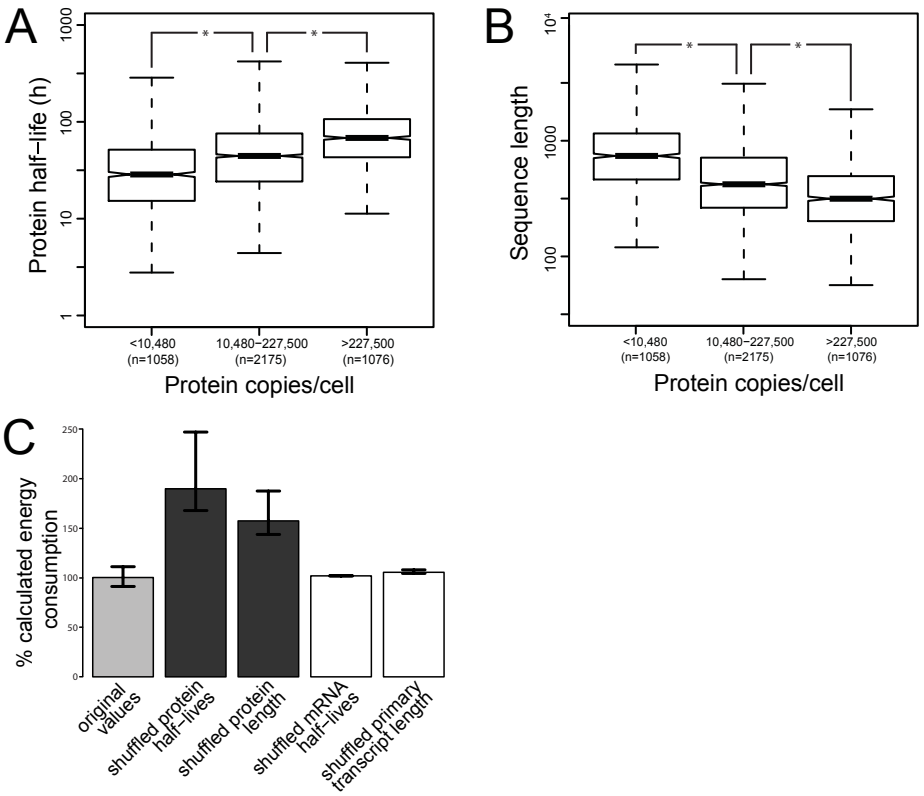
Correlation of sequence motifs with mRNA stability. The number of (A) AU-rich elements (AREs) and (B) Pumilio2 binding sites in 3´UTR sequences is higher in unstable mRNAs (upper panels). To account for stochastic differences in motif occurrence due to vastly different 3´UTR lengths, motif densities (motifs/nucleotide) were calculated by dividing the number of motifs by 3´UTR sequence length (lower panels).

Fig. S11



Protein sequence features with an impact on protein stability. (A) The degree of disorder (see supplemenatary methods) of each protein (in bins of 100) was plotted against corresponding, binned protein half-lives. Highly structured proteins are on average more stable. (B) The bar plot shows to what degree relative fractions of amino acids correlate with protein stability. For instance, the amino acid proline frequently occurs in unstable proteins, resulting in a highly significant negative Spearman rank correlation coefficient (* indicates p-values < 0.05 but ≥ 0.01; *** indicates p-values < 0.001)

Fig. S12



Evidence for energy constraints in mammalian cells. (A, B) Abundant proteins are significantly more stable and shorter than less abundant ones ($p<10^{-15}$, Wilcoxon test). (C) Theoretical energy consumption of gene expression. Randomizing protein half-lives or lengths enhances energy costs. Shuffling mRNA half-lives or lengths has little impact due to their minor contribution to total energy costs. Error bars show 95% confidence intervals determined by multiple randomizations and bootstrapping.