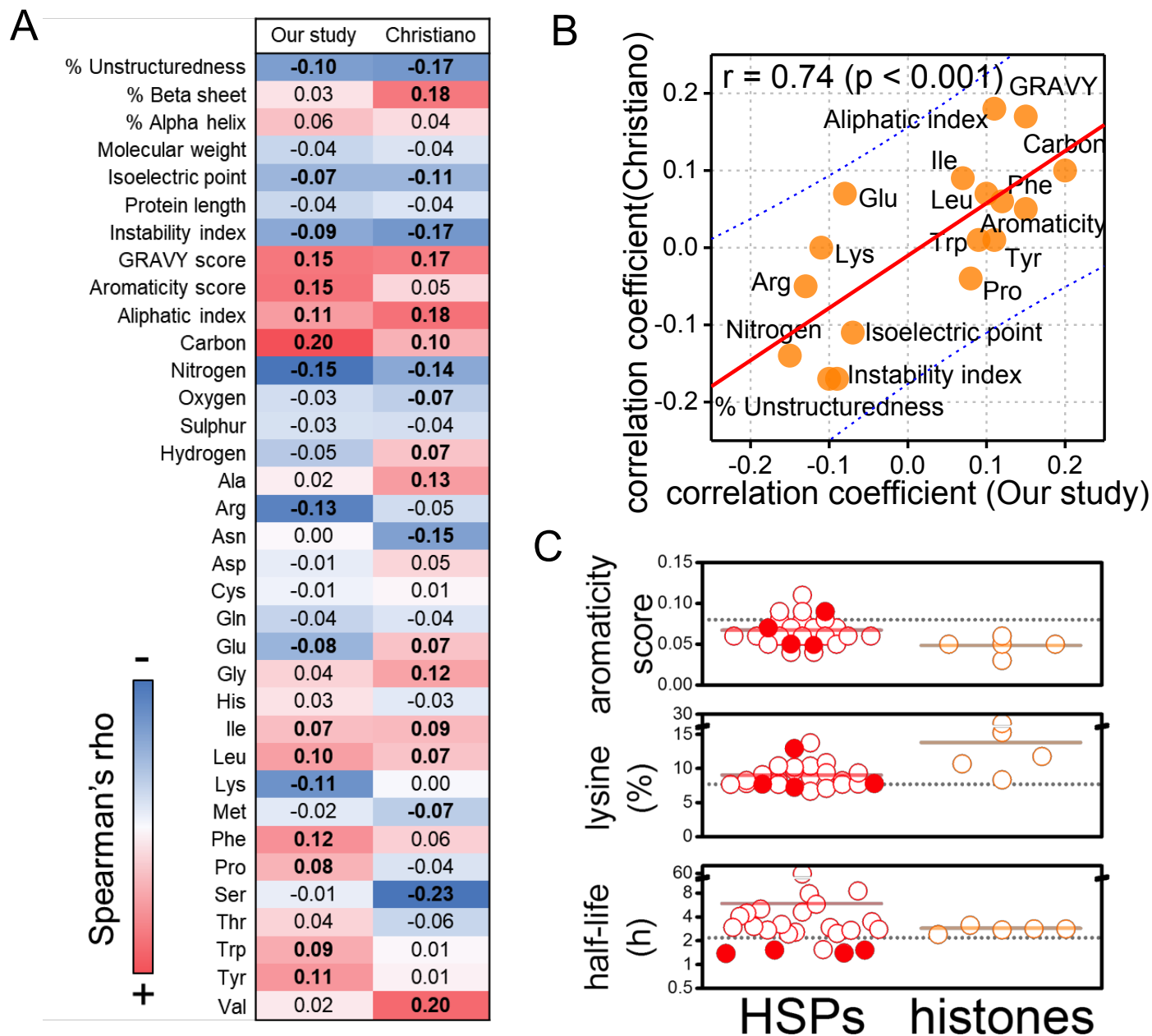
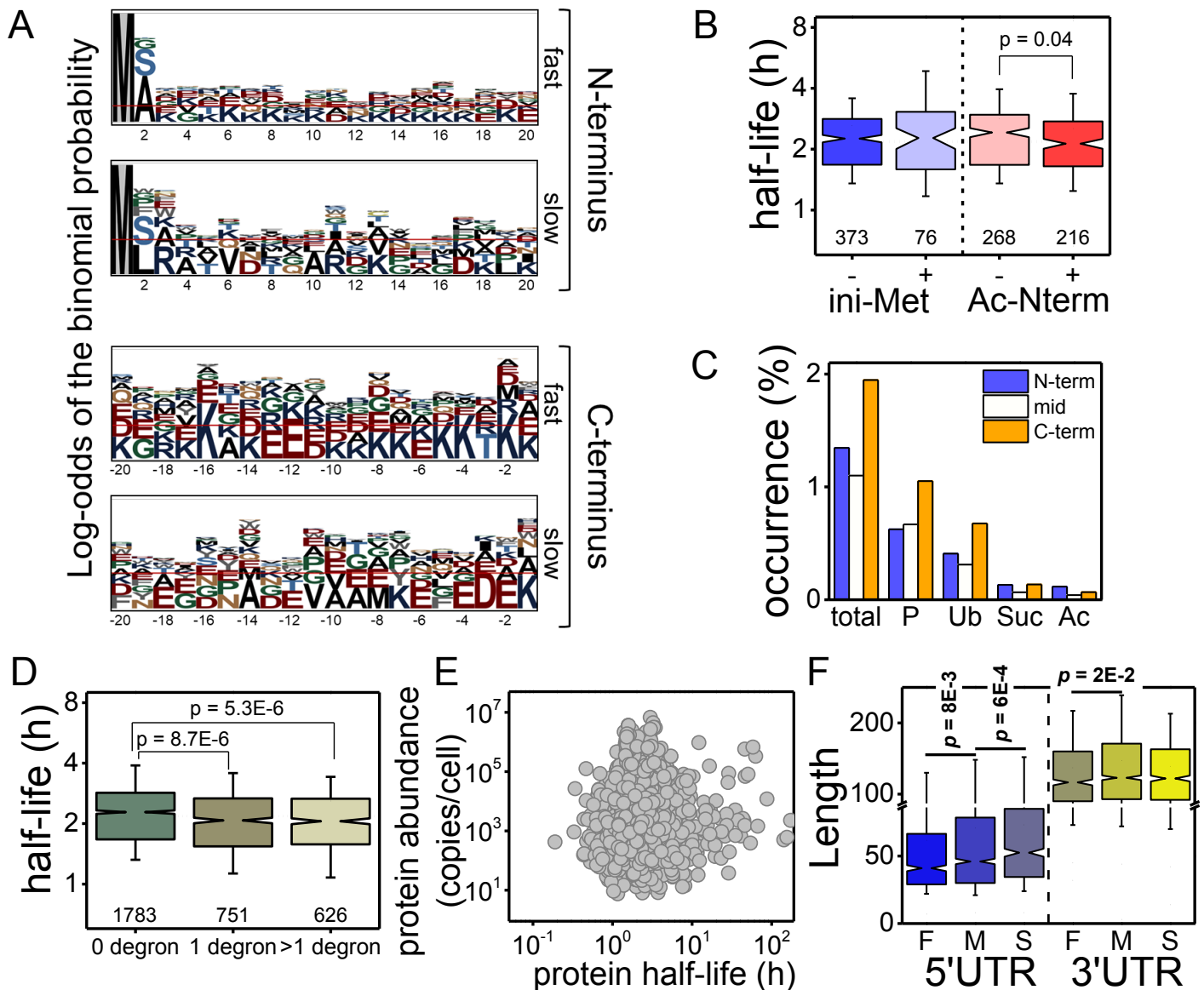


**Figure S1. Reproducibility of protein abundance and half-life measurements. Related to Figure 1.** (A) Distribution of CV of the  $\log_{10}$  transformed protein abundances (copies/cell) across the different timepoints of the study in the replicate 2 ( $CV < 0.2$  in  $>90\%$  of proteins). Inner graph depicts the reproducibility of protein abundance measurements in copy number per cell between replicates ( $n = 3097$ ). Identity function is represented by a grey dotted line. Pearson's correlation values are indicated on the top left corner of the plot. (B) Profile plot of protein abundances present in every time point of the experiment replicate 2. Purple, blue and dark yellow lines represent the evolution of 100 proteins from the top, medium and bottom abundance range. The rest of proteins are plotted in grey. (C) Reproducibility of half-life calculation using all (i.e. curve fitting and single-point measurements;  $n = 2904$ ) or only curve-fitting ( $n = 1978$ ) measurements. (D) Correlation between half-life estimations based on curve fitting and single-point (sgpt) analysis in replicate 2 ( $n = 2203$ ; time-point at '3h' was used for single point measurements). (E) Reproducibility of half-life estimations based on single-point measurements ( $n = 2239$ ; time-point at '3h' was used in both replicates). Pearson's correlation values are indicated on the top left corner of the plot. (F) Abundance distribution of proteins with half-life estimates obtained by curve-fitting or single-point method.

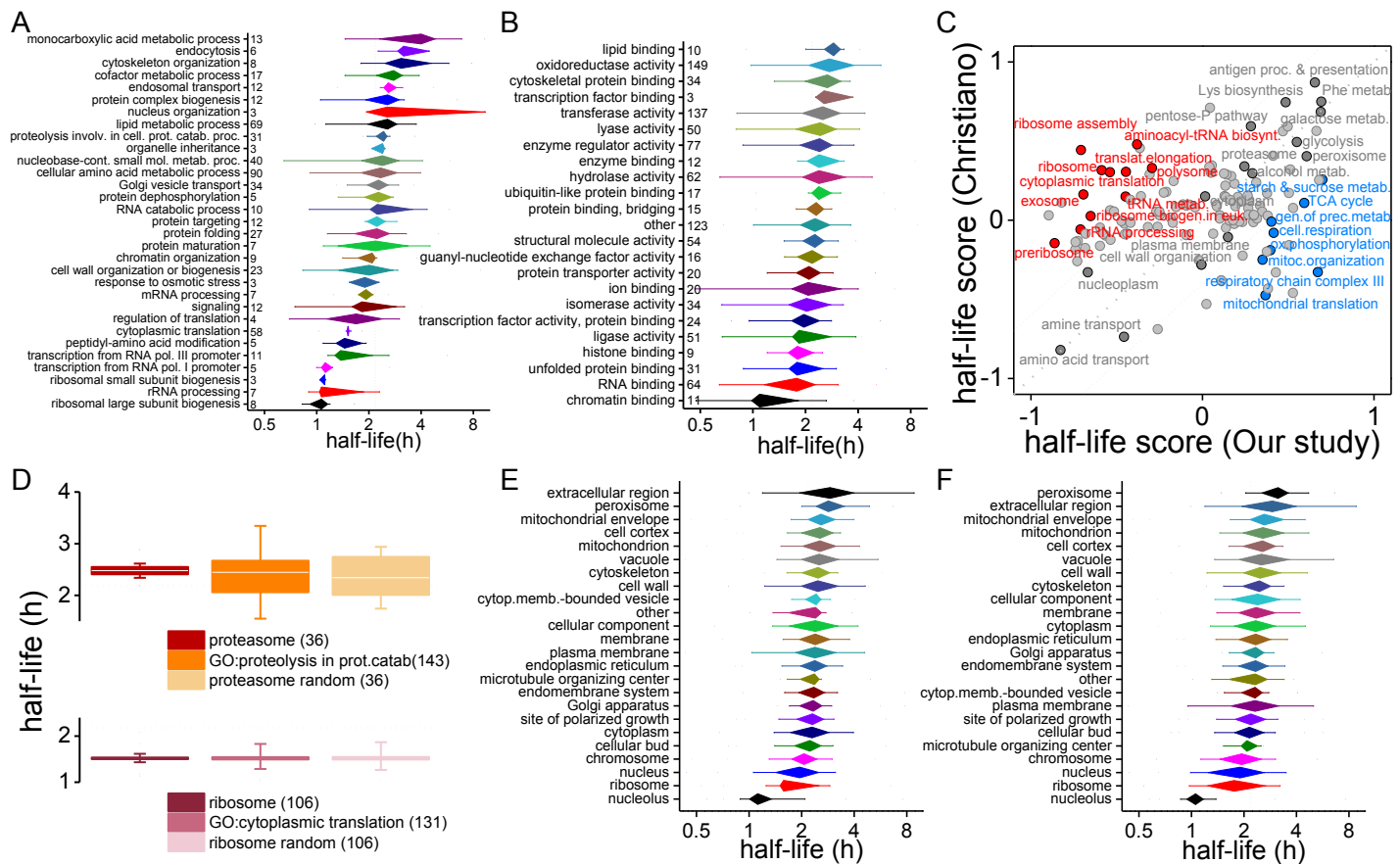




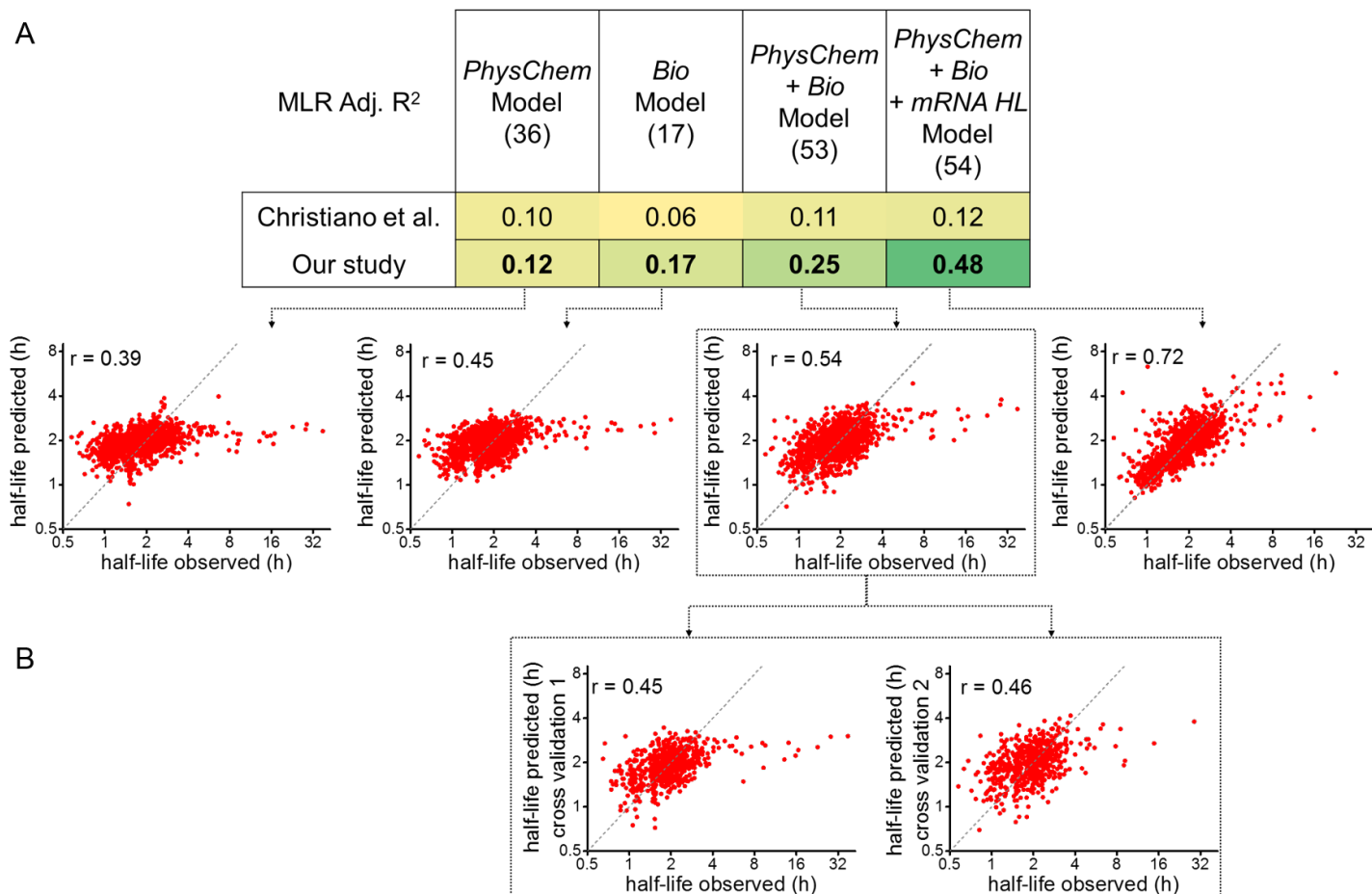
**Figure S3. Complete analysis of physicochemical properties and comparison with previous turnover studies in yeast. Related to Figure 1.** (A) Heat map of Spearman rank coefficients from correlations between protein physicochemical properties and half-lives calculated in our study and others (Christiano et al. 2014). Significant correlations at 99% confidence level ( $p$ -value < 0.01) using the Bonferroni correction are highlighted in bold. (B) Comparison of correlation coefficients between our study and others (Christiano et al. 2014). Only the coefficients for physicochemical properties having significant correlations with half-life in our study are plotted. (C) Physicochemical description of heat shock proteins (HSPs) and histones in yeast. Distribution of aromaticity score, lysine content and protein half-life for members of both protein families found in our study. Filled red circles represent ribosome associated HSPs. Average values are represented by bold lines while median proteome values are represented by a dotted grey line.



**Figure S4. Effect of other physicochemical and biological properties on protein turnover. Related to Figure 1 and Figure 4.** (A) Sequence analysis of the N and C protein termini in stable and unstable protein groups. The first and the last 20 amino acids of the protein sequence, corresponding to the N and C termini, respectively, are shown. N-terminus and C-terminus from all ORFs present in the SGD proteome were used as background (letter size larger than the red line indicate significantly enriched amino acids for that position;  $p < 0.05$ ). (B) Protein half-life distributions of proteins with or without initiator methionine (left panel), and proteins with acetylated or nonacetylated N-terminus (right panel). Data of experimentally determined protein N-termini in yeast obtained from Mommen et al. 2012. Dotted grey line represents the median half-life of the proteome. (C) Relative occurrence in % (i.e. events per hundred amino acids) of main PTMs in the yeast proteome. N-term and C-term account for the first and last 20 amino acids of the protein sequence respectively, while mid account for the sequence in between terms (P: phosphorylation; Ub: ubiquitylation; Suc: succinylation; Ac: acetylation). (D) Comparison of protein half-life distributions between proteins containing none, only one or more than one degron motifs within the same molecule. Dotted grey line represents the median half-life of the proteome. (E) Scatter plot showing no relationship between protein abundances and half-lives measured in exponentially growing yeast cells from our study. (F) Comparison of median 5'UTR and 3'UTR lengths in unstable, intermediate and stable proteins obtained from our study. UTR length data was obtained from Pelechano et al. 2013.



**Figure S5. GO term mapping of protein half-lives. Related to Figure 2.** (A, B) Distribution of protein half-lives according to biological process (A) and molecular function (B) GO terms (proteins exclusively included within the designated GO terms were used and only terms with 3 or more proteins are shown). (C) 2D annotation enrichment analysis (FDR<0.05) comparing protein half-lives measurements from Christiano et al. 2014 and our study. Identity function and 20% confidence intervals are represented by diagonal grey dashed and dotted lines, respectively. (D) Comparison of half-lives distributions between members of the proteasome (top graph) and ribosome (bottom graph) complexes, their corresponding GO biological process (i.e. ‘proteolysis involved in cellular protein catabolic process’ and ‘cytoplasmic translation’, respectively) and a set of randomly selected proteins within these terms matching the same number of complex members identified. Boxes represent 25 and 75 percentiles while whiskers represent the standard deviation. The number of proteins in each boxplot is indicated in brackets within the legend. (E, F) Protein half-lives distribution within the different GO cellular component terms when proteins belonging to complexes are included (E) or not (F).



**Figure S6. Predictability of protein half-life using multiple linear regression (MLR) analysis. Related to Figure 5.** (A) Adjusted R<sup>2</sup> values from MLR predictive analysis of protein half-lives from our study and Christiano et al. study. The predictive models employed use as independent variables: physicochemical parameters (*PhysChem* model, including all 36 variables shown in Figure S3); biological parameters (*Bio* model, including 17 biological parameters: PTMs (phosphorylation, ubiquitylation, acetylation and succinylation) occupancy; degron (PEST, Dbox and KEN) and protease (calpain, caspase 3, 6 and 8) motif density; number of protein interactions; essentiality; 5' and 3' UTR lengths; protein and mRNA abundances); both physicochemical plus biological parameters (*PhysChem+Bio* model); and physicochemical plus biological plus mRNA half-life parameters (*PhysChem+Bio+mRNA HL* model, adding mRNA half-lives from (Geisberg et al., 2014) to the previous model). The number of independent variables employed in each model is indicated in brackets. All variables used in the models were calculated based on protein sequence or derived from external sources, and were already included in the study except for the protein abundance data which was obtained from the integrated PaxDB database. Only protein half-lives calculated by curve fitting in both replicates of our study and that were also present in Christiano et al. study were used for the modelling analysis ( $n = 1540$  proteins). Scatter plots between observed and predicted half-lives for each of the models are shown at the bottom (Pearson's correlations are indicated in the upper left corner of each plot and dotted lines indicates identity function). (B) Cross-validation of the physicochemical plus biological model as a test of overfitting. The data was randomly divided in two halves and each half was used to cross validate the regression derived from the other half of the data (Pearson's correlations are indicated in the upper left corner of each plot and dotted lines indicates identity function). A correlation shrinkage is observed due to a loss of precision and increase of variability in the estimates when fitting on a smaller training set, but they still predict a similar fraction of the variance in the test set.