Electronic Supplementary Material (ESI) for Molecular BioSystems

Supplementary Information

Escherichia coli achieves faster growth by increasing catalytic and translation rates of proteins

Kaspar Valgepea, Kaarel Adamberg, Andrus Seiman and Raivo Vilu

Table of Contents

Supplementary Figures

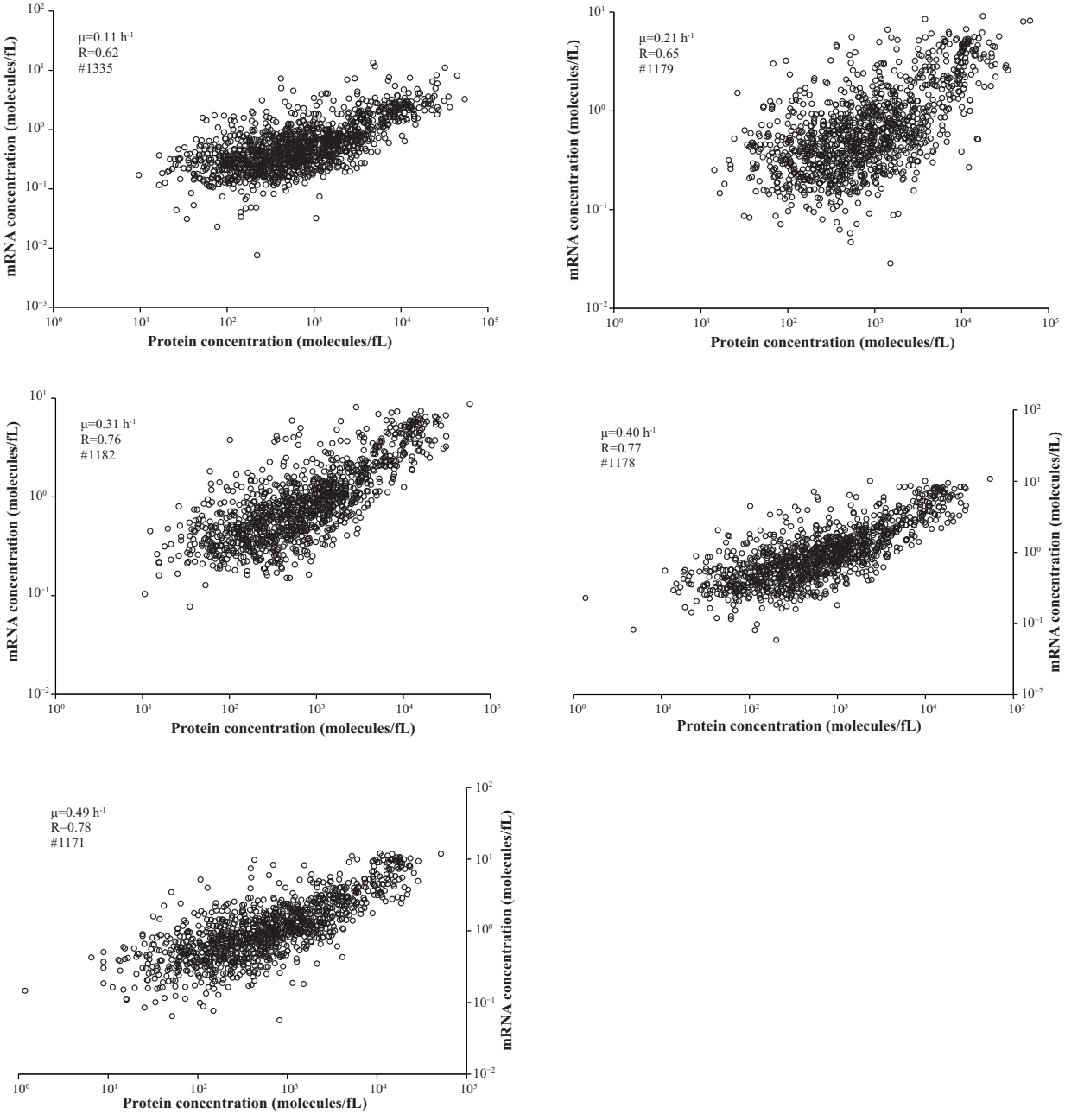


Fig. S1 μ-dependent correlation of mRNA and protein concentrations. R, PCC; #, number of compared genes.

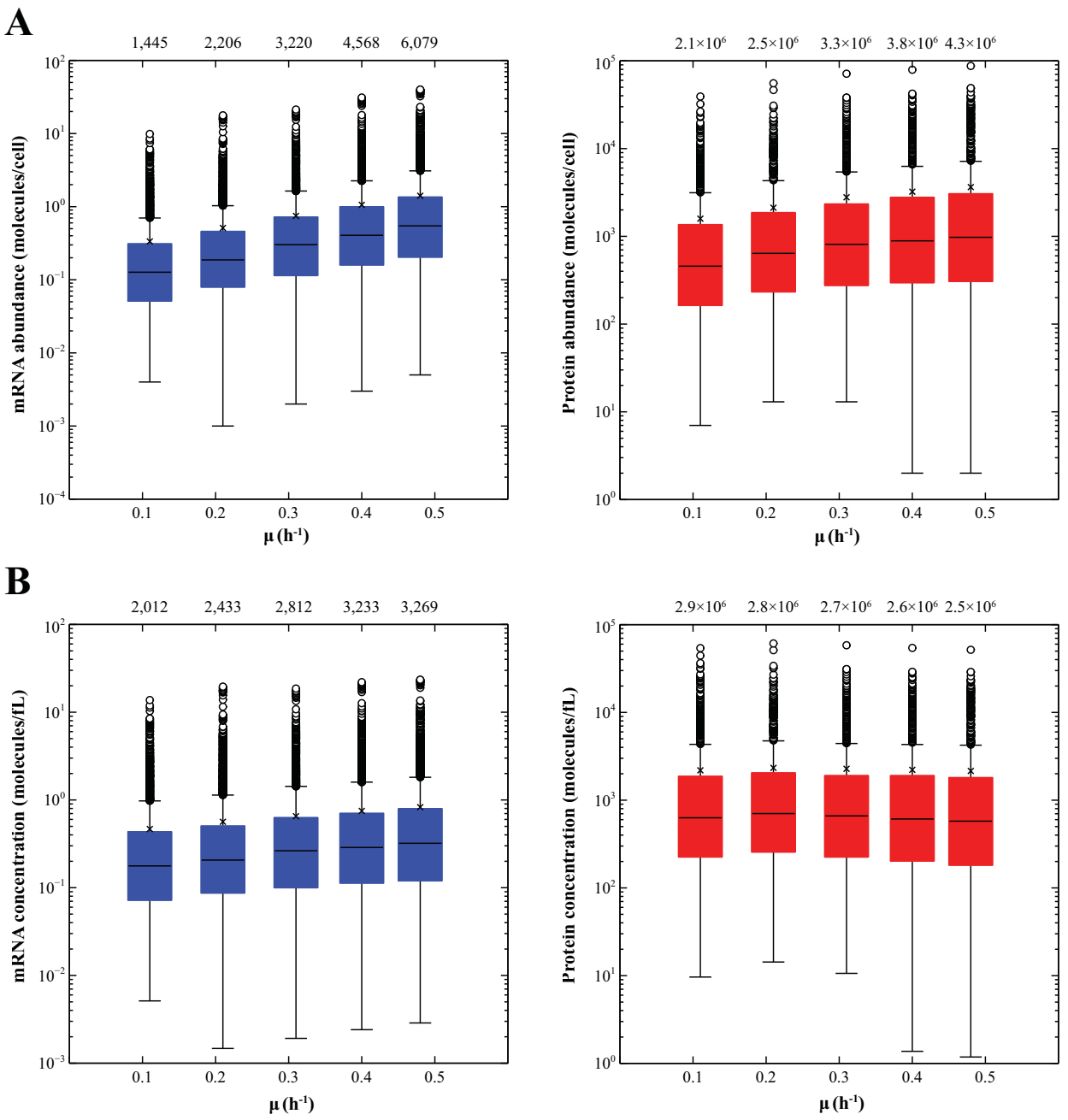


Fig. S2 μ-dependent absolute mRNA and protein dynamics. (A) mRNA and protein abundance distribution box plots. (B) mRNA and protein concentration distribution box plots. Sum of each μ's respective parameters is noted above the box plot. Colored box plot horizontal borders denote 25th and 75th and black horizontal bar 50th (median) percentiles, and whiskers represent 1.5 interquartile ranges.

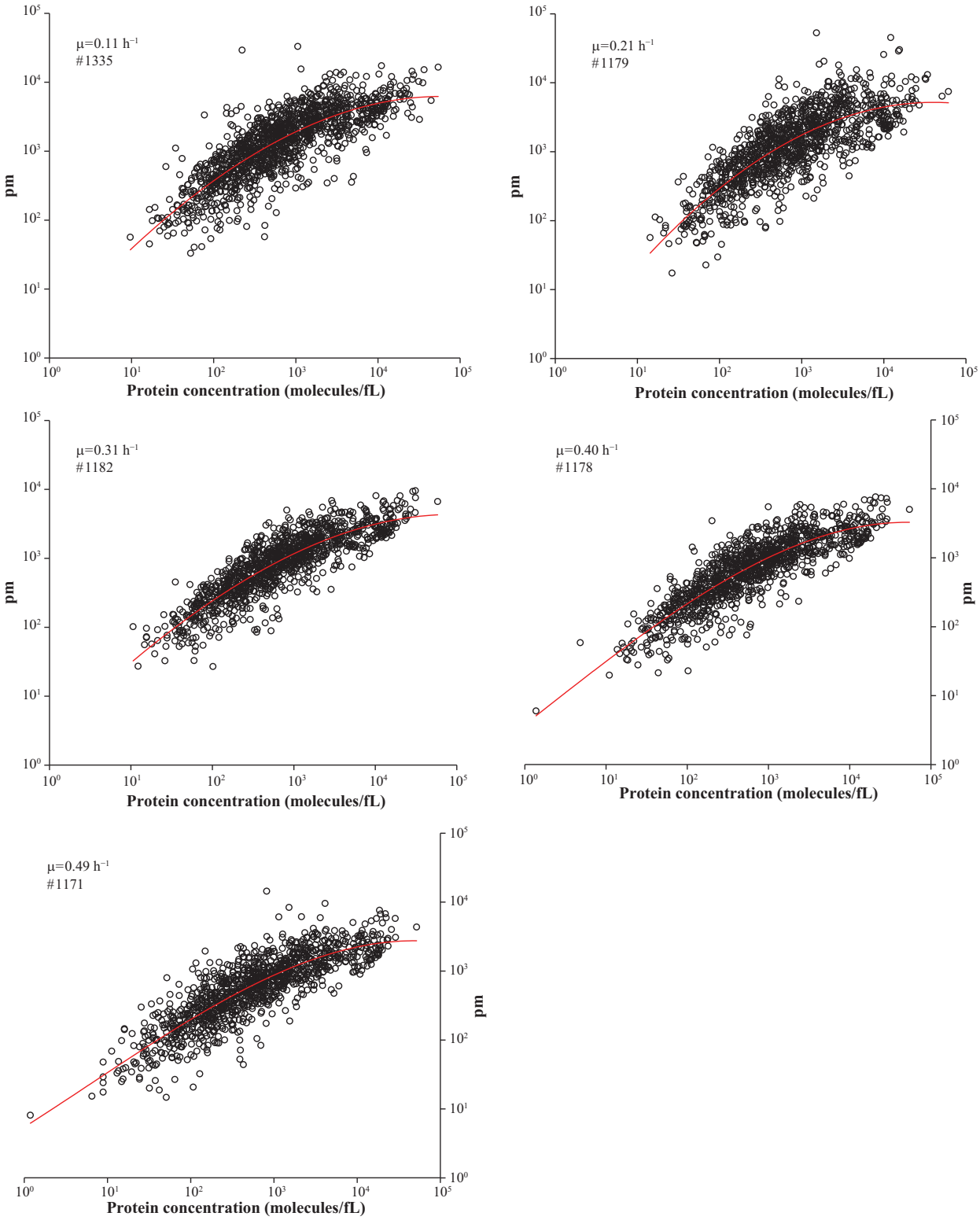


Fig. S3 Protein concentration vs. pm under the studied range of μ. Red line shows the 3rd order polynomial best fit. #, number of analyzed genes.

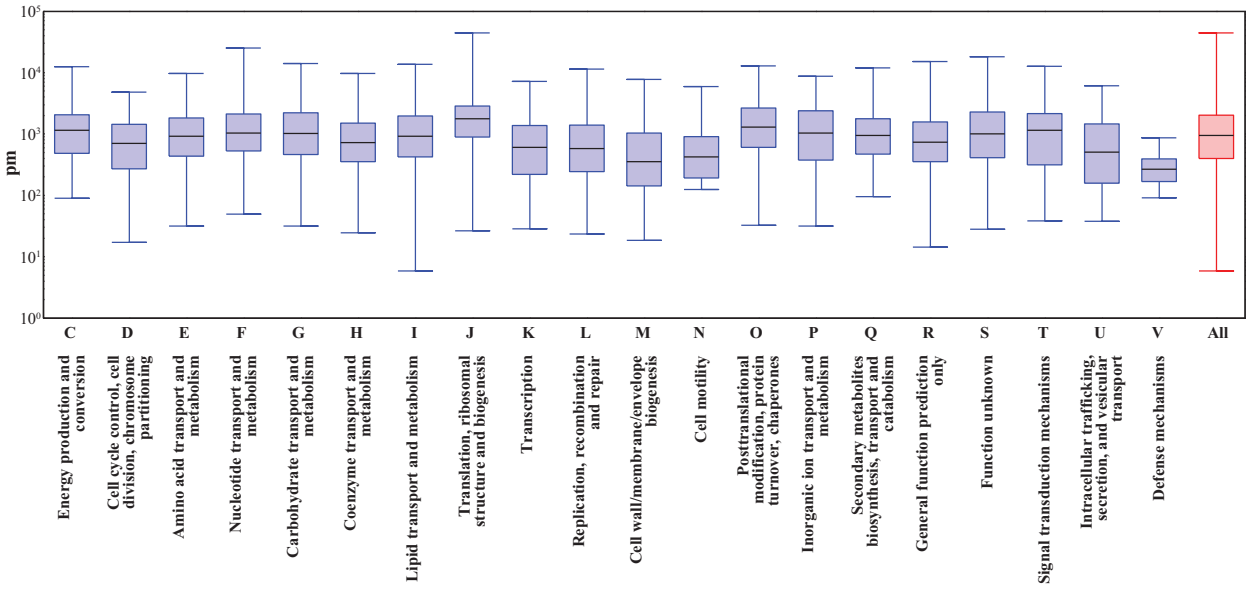


Fig. S4 pm for COG functional classes. pm values over the whole studied range of μ (0.11–0.49 h^{-1}) were taken into account. All quantified mRNAs and proteins were grouped into COG functional classes according to Tatusov *et al.*[1] Horizontal bars represent 25th, 50th (median, black) and 75th percentiles, and whiskers represent the highest and lowest values.

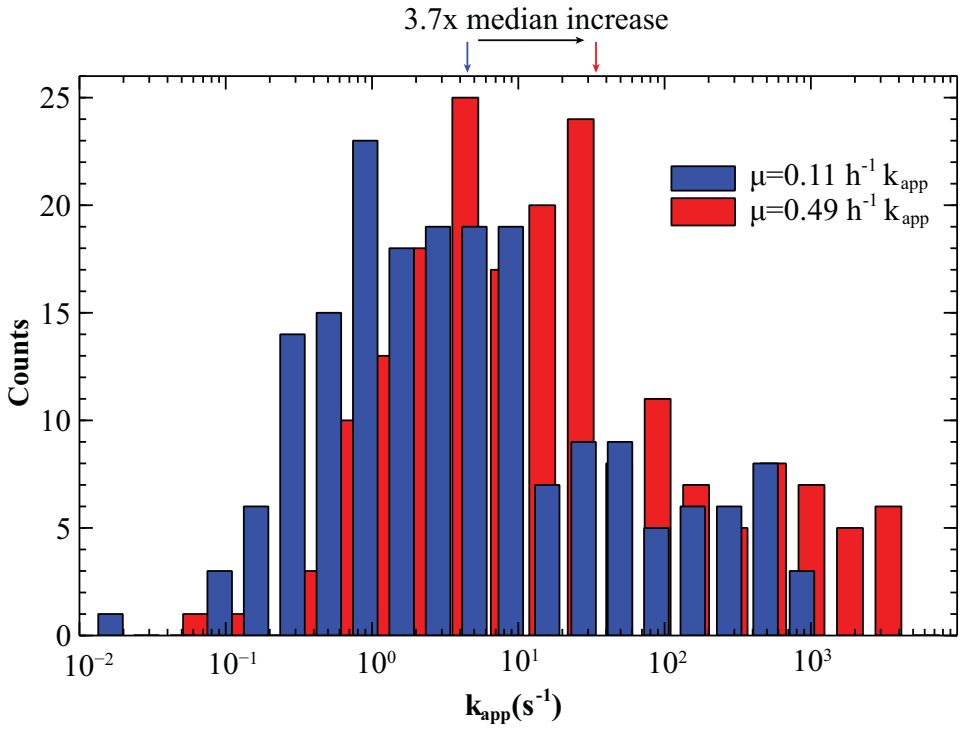


Fig. S5 Dynamic range of 191 central metabolism and biosynthetic enzyme k_{app} values at low and high μ. Blue and red vertical lines with arrowheads above the chart denote the median at μ=0.11 and 0.49 h^{-1}, respectively.

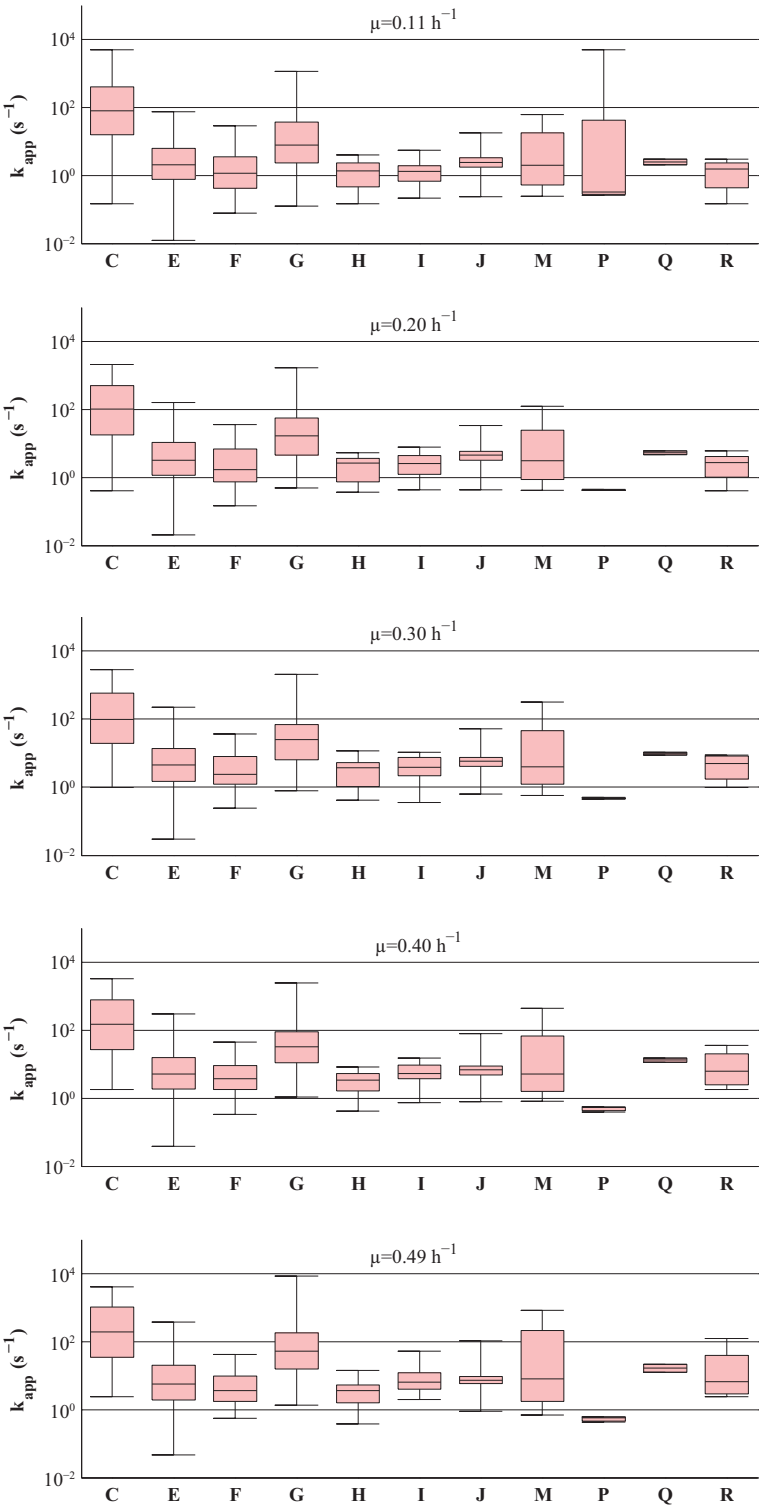


Fig. S6 Distribution of k_{app} values among COG functional classes under the studied range of μ. Refer to Fig. S4 for COG nomenclature. Horizontal bars represent 25th, 50th (median) and 75th percentiles, and whiskers represent the highest and lowest values. Very high k_{app} values could result from the possibility that under *in vivo* conditions actually an iso-enzyme is catalyzing a substantial part of the flux.

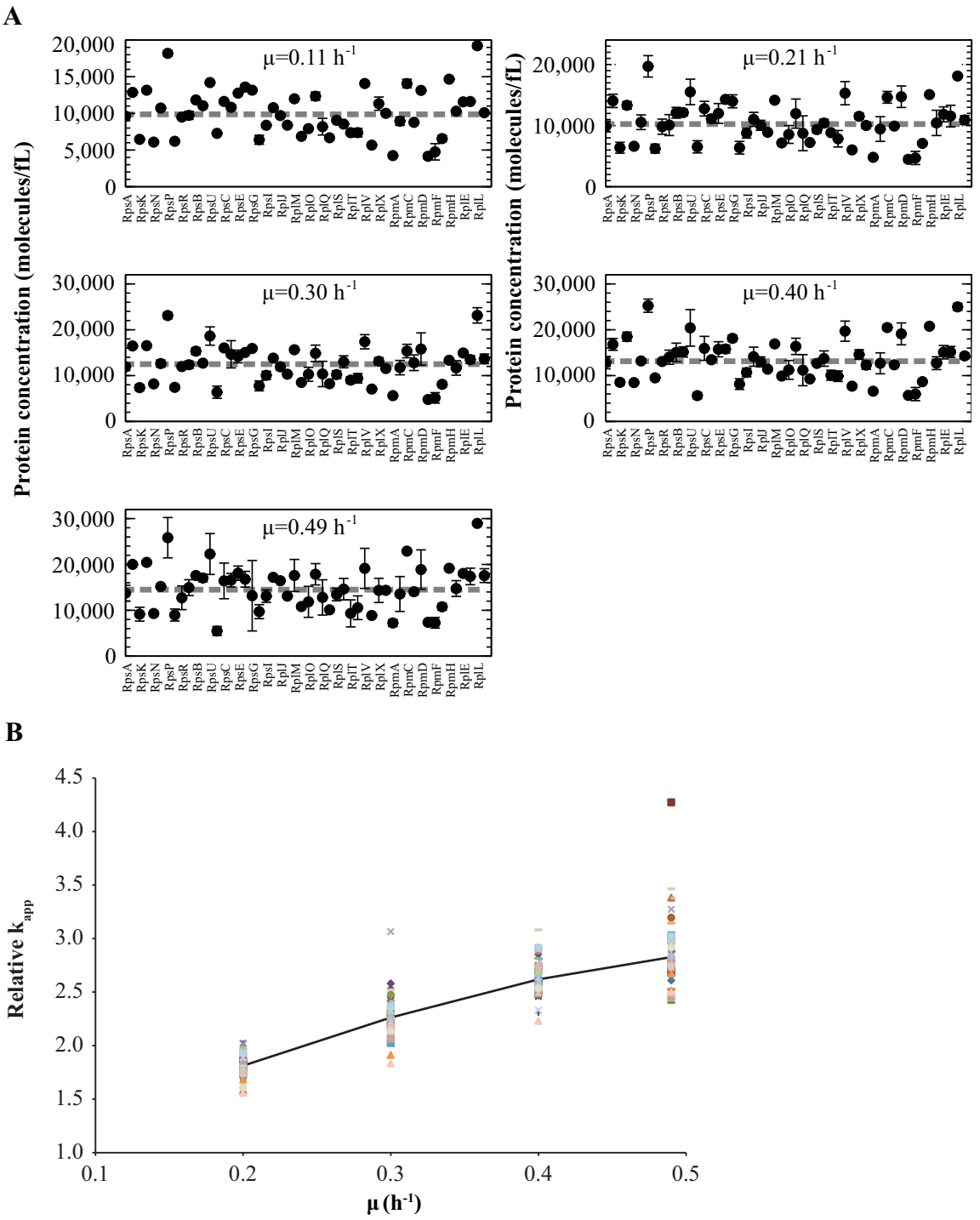


Fig. S7 μ-dependent concentrations and increase of translation rates for the 52 quantified ribosomal proteins. (A) μ-dependent ribosomal protein concentrations. Grey dashed line denotes the median and error bars average absolute deviation between proteome quantification in two independent A-stat experiments. Protein names are shown for the ones with highest concentrations (B) Relative increase of ribosomal protein translation rates expressed through k_{app} values. Relative k_{app} is calculated by comparing k_{app} at each μ to the value at reference μ=0.11 h^{-1}. Black line indicates the median.

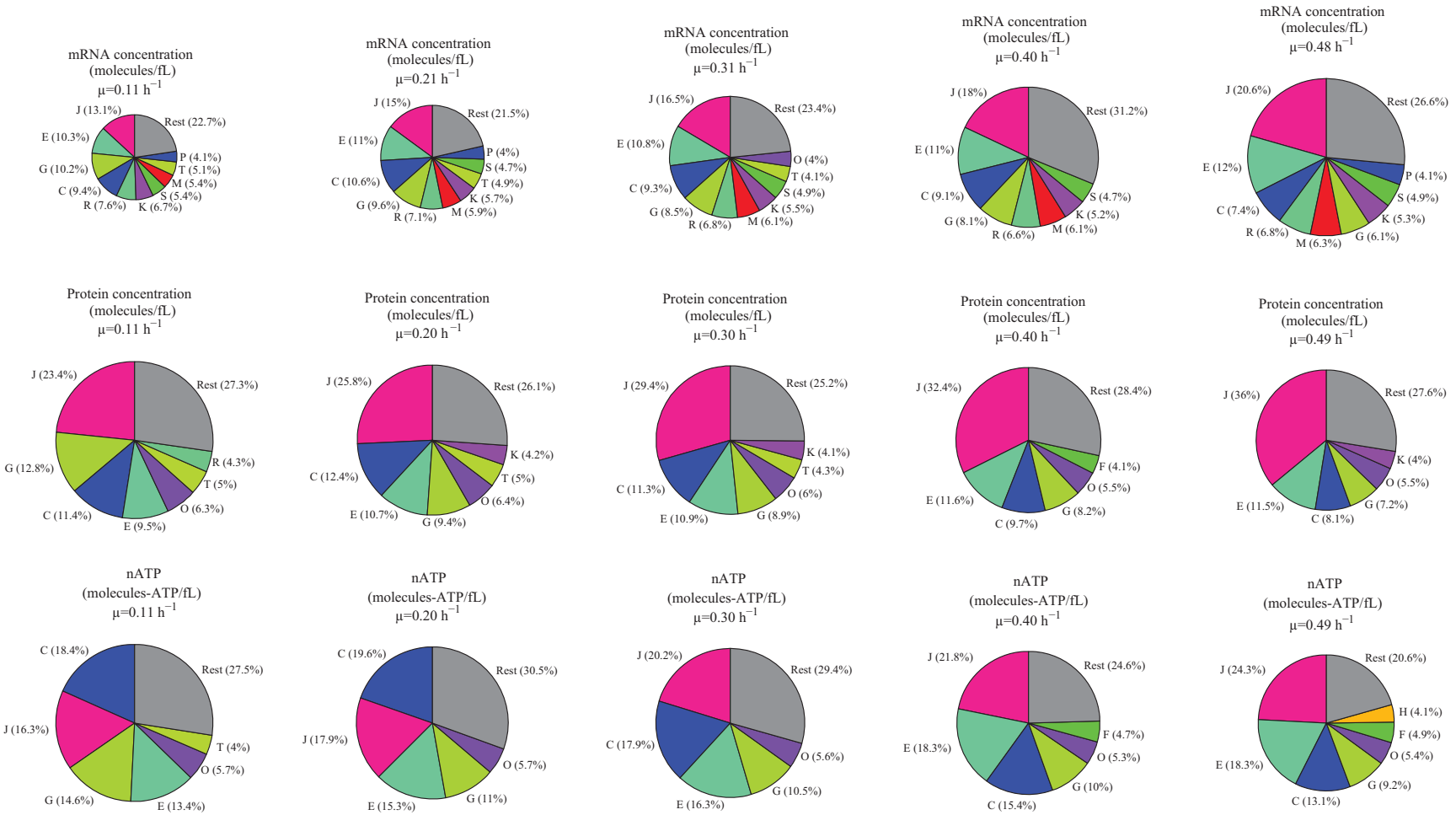
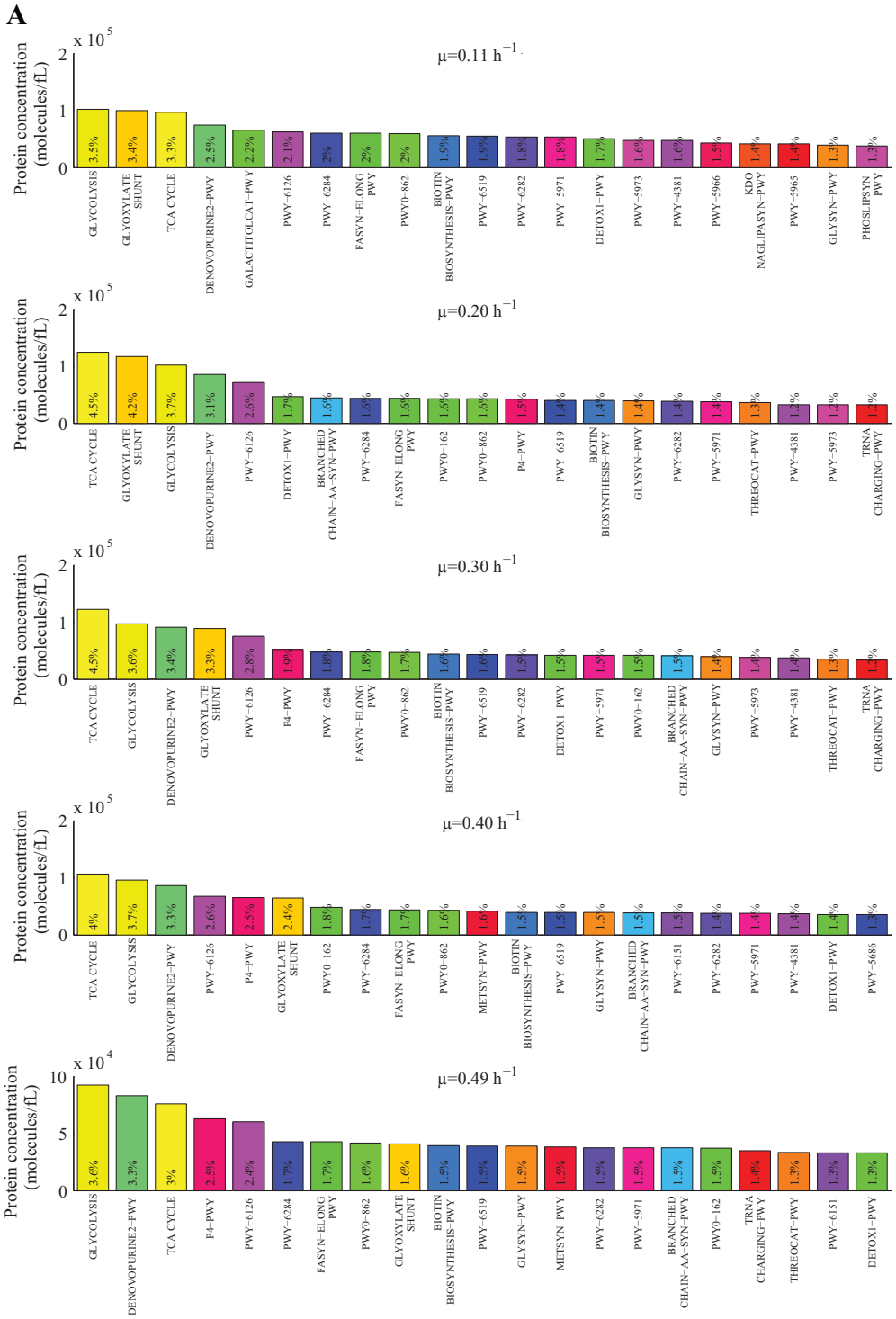


Fig. S8 Distribution of mRNA and protein concentrations and nATP among COG functional classes under the studied range of μ. Refer to Fig. S4 for COG nomenclature and Experimental for nATP calculation. The area of slices and pies within one parameter (mRNA, protein or nATP) are proportional to absolute values. Rest, sum of COGs not shown on the pies independently due to minor proportions from the total pool.



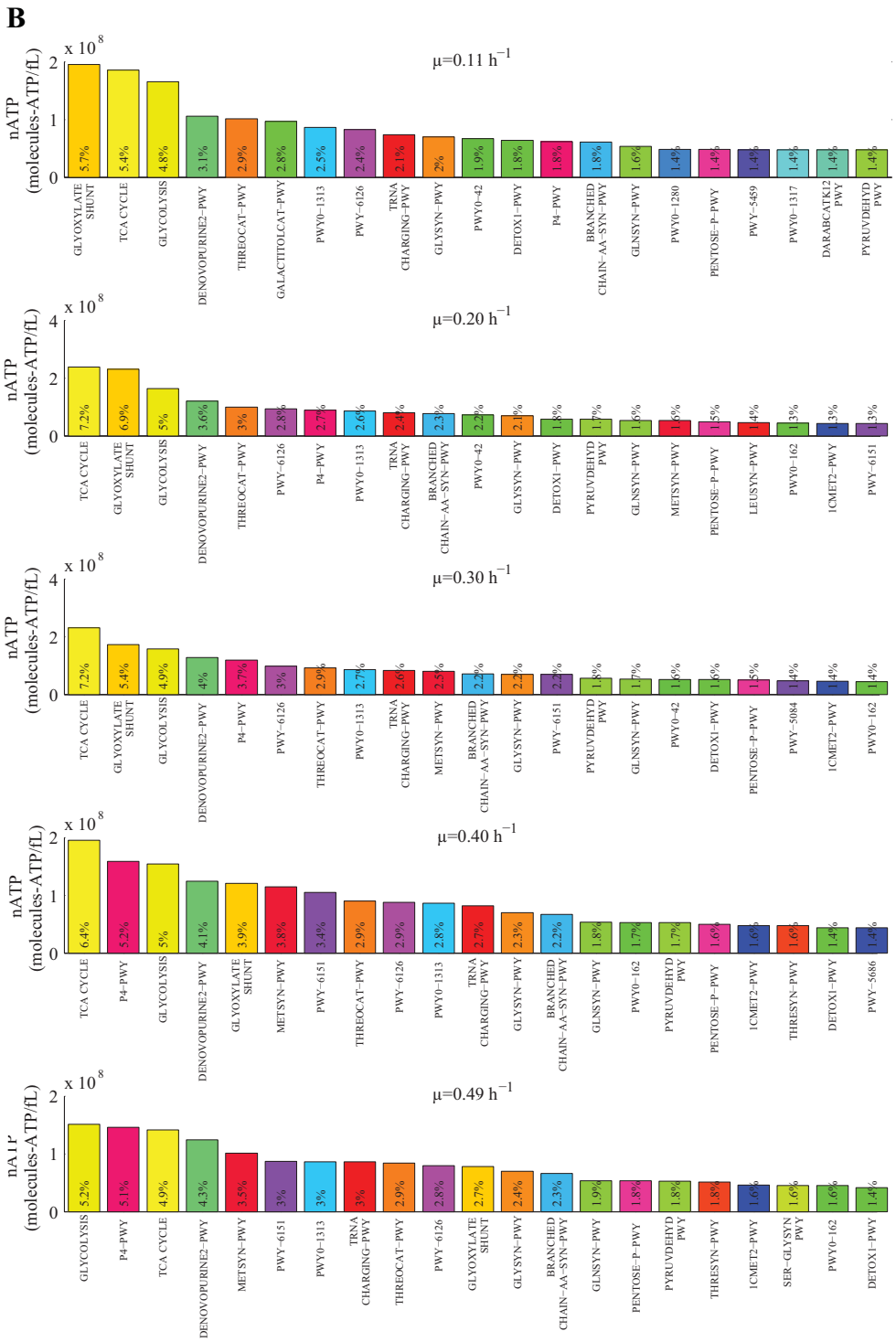
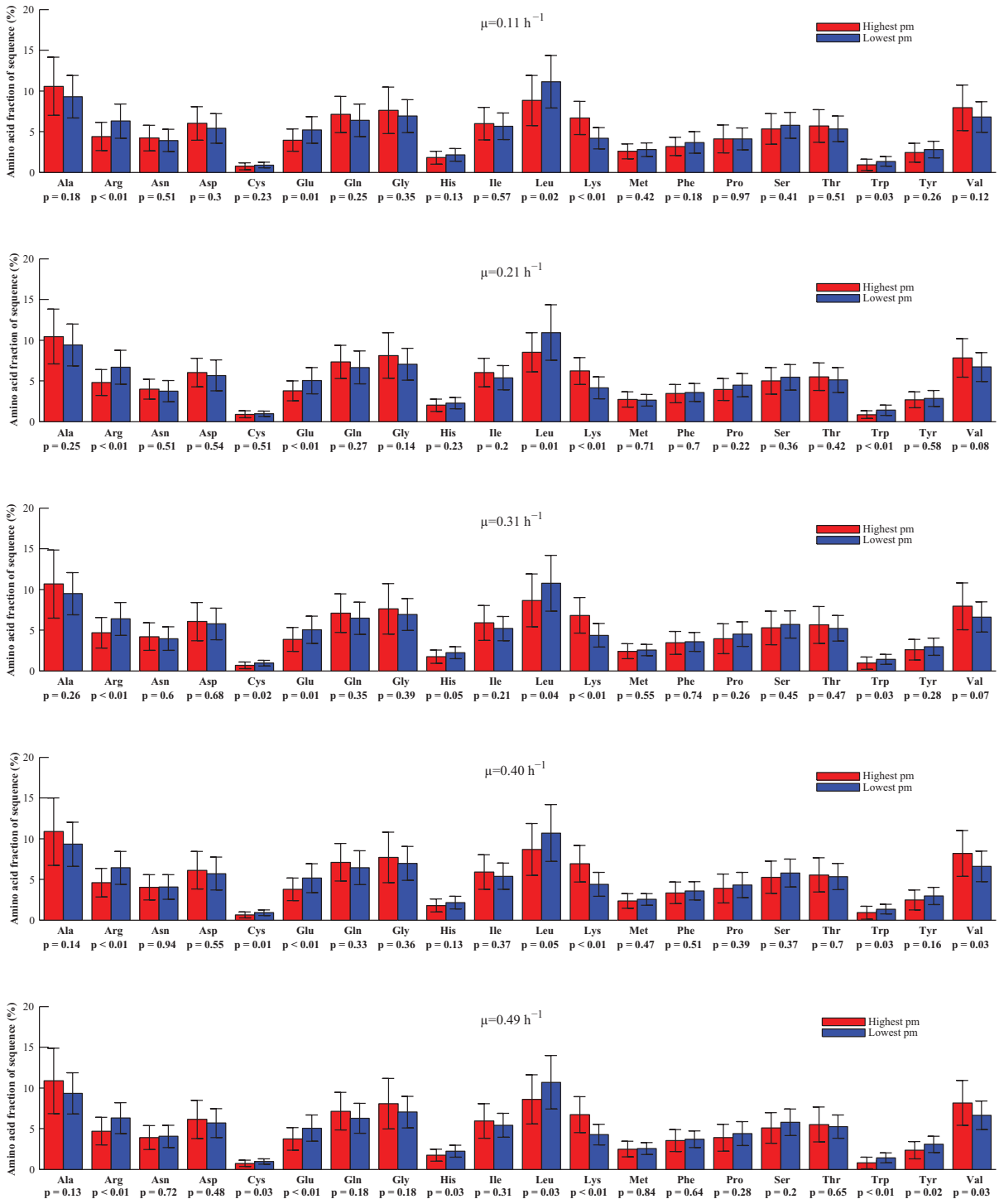


Fig. S9 Pathways with the highest protein concentrations and nATP from the total proteome under the studied range of μ. (A) Pathways with the highest protein concentrations from the total proteome. (B) Pathways with the highest nATP from the total proteome synthesis cost. Fraction of each pathway from the total proteome is shown in percentages. See Table S7 for the genes assigned to pathways according to the EcoCyc database.[2] Refer to Experimental for nATP calculation.

A


μ=0.11 h−1
Highest pm
Lowest pm
Amino acid fraction of sequence (%)
Ala p = 0.18
Arg p < 0.01
Asn p = 0.51
Asp p = 0.3
Cys p = 0.23
Glu p = 0.01
Gln p = 0.25
Gly p = 0.35
His p = 0.13
Ile p = 0.57
Leu p = 0.02
Lys p < 0.01
Met p = 0.42
Phe p = 0.18
Pro p = 0.97
Ser p = 0.41
Thr p = 0.51
Trp p = 0.03
Tyr p = 0.26
Val p = 0.12
μ=0.21 h−1
Ala p = 0.25
Arg p < 0.01
Asn p = 0.51
Asp p = 0.54
Cys p = 0.51
Glu p < 0.01
Gln p = 0.27
Gly p = 0.14
His p = 0.23
Ile p = 0.2
Leu p = 0.01
Lys p < 0.01
Met p = 0.71
Phe p = 0.7
Pro p = 0.22
Ser p = 0.36
Thr p = 0.42
Trp p < 0.01
Tyr p = 0.58
Val p = 0.08
μ=0.31 h−1
Ala p = 0.26
Arg p < 0.01
Asn p = 0.6
Asp p = 0.68
Cys p = 0.02
Glu p = 0.01
Gln p = 0.35
Gly p = 0.39
His p = 0.05
Ile p = 0.21
Leu p = 0.04
Lys p < 0.01
Met p = 0.55
Phe p = 0.74
Pro p = 0.26
Ser p = 0.45
Thr p = 0.47
Trp p = 0.03
Tyr p = 0.28
Val p = 0.07
μ=0.40 h−1
Ala p = 0.14
Arg p < 0.01
Asn p = 0.94
Asp p = 0.55
Cys p = 0.01
Glu p < 0.01
Gln p = 0.33
Gly p = 0.36
His p = 0.13
Ile p = 0.37
Leu p = 0.05
Lys p < 0.01
Met p = 0.47
Phe p = 0.51
Pro p = 0.39
Ser p = 0.37
Thr p = 0.7
Trp p = 0.03
Tyr p = 0.16
Val p = 0.03
μ=0.49 h−1
Ala p = 0.13
Arg p < 0.01
Asn p = 0.72
Asp p = 0.48
Cys p = 0.03
Glu p < 0.01
Gln p = 0.18
Gly p = 0.18
His p = 0.03
Ile p = 0.31
Leu p = 0.03
Lys p < 0.01
Met p = 0.84
Phe p = 0.64
Pro p = 0.28
Ser p = 0.2
Thr p = 0.65
Trp p < 0.01
Tyr p = 0.02
Val p = 0.03

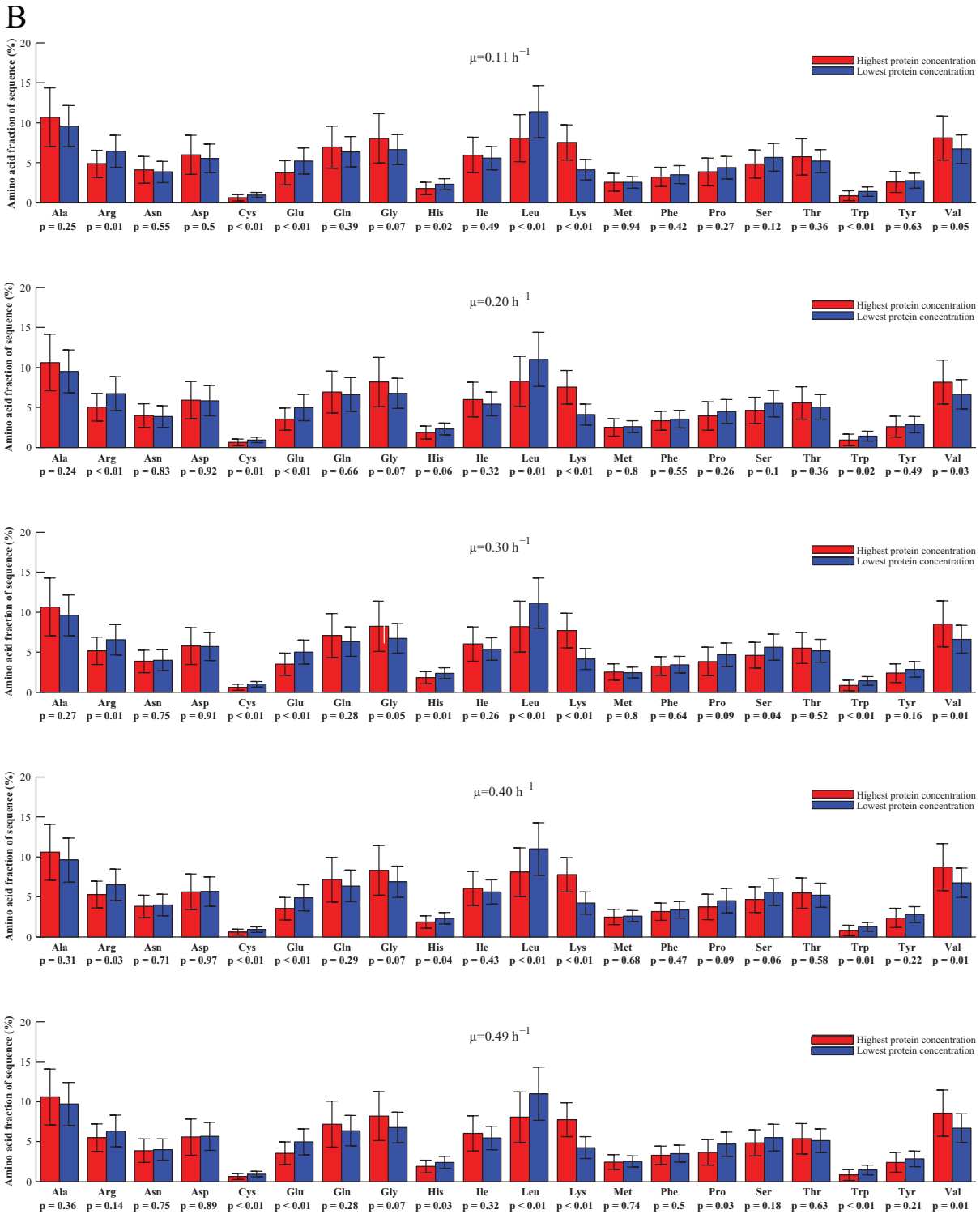


Fig. S10 Amino acid fractions of selected protein sequences under the studied range of μ. (A) Amino acid fractions from sequences of the 100 proteins with highest and lowest pm. (B) Amino acid fractions from sequences of the 100 proteins with highest and lowest concentrations (molecules/fL). Bar heights denote the mean and error bars 95% CI. p-values were calculated between the two compared data sets (highest and lowest) for each amino acid using the t-test.

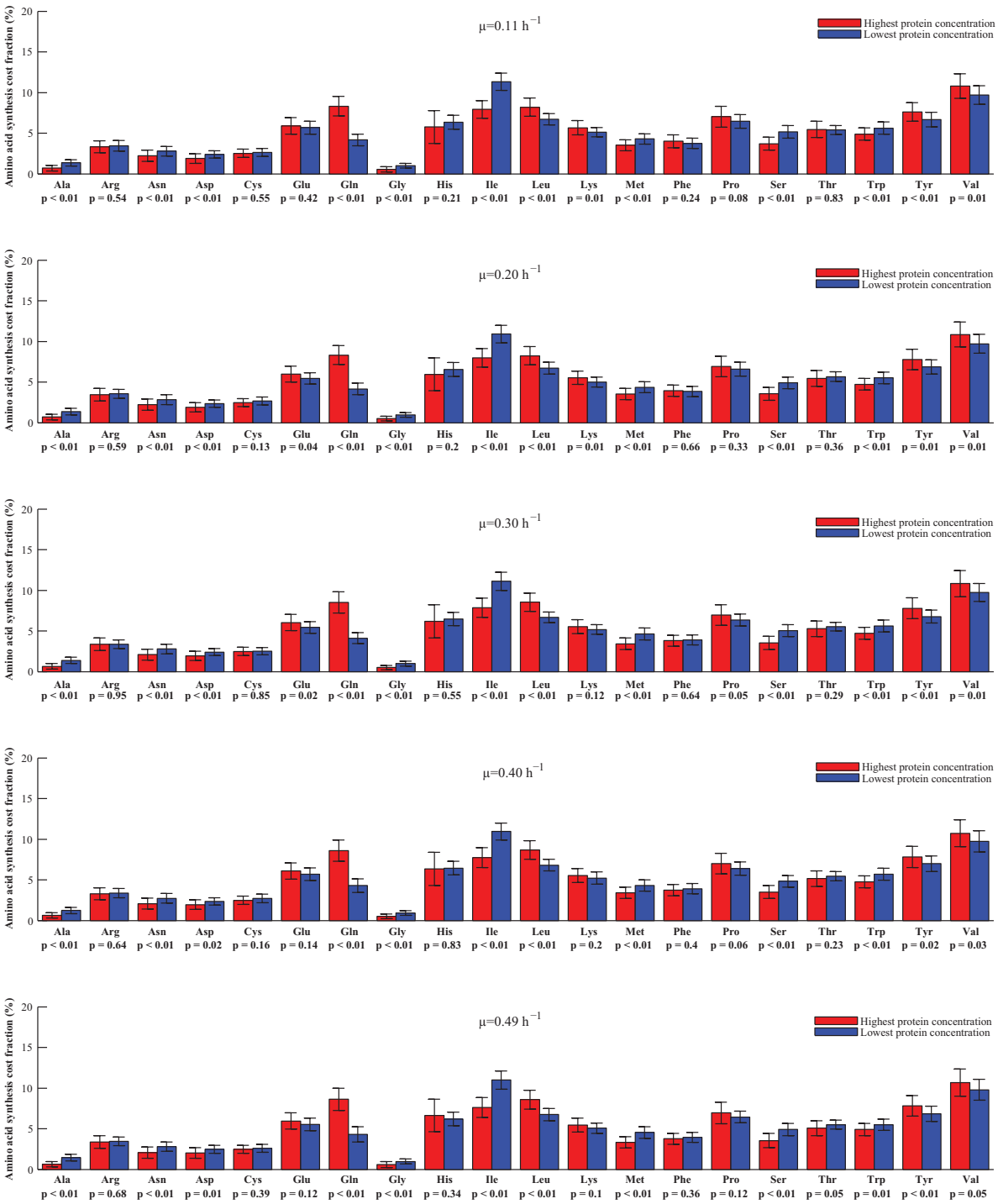


Fig. S11 Amino acid synthesis cost fractions for the 100 proteins with highest and lowest concentrations under the studied range of μ. Amino acid synthesis cost fraction is the fraction of the synthesis cost of a particular amino acid from the total synthesis cost of all the amino acids making up the protein. Amino acid synthesis costs were taken from Akashi and Gojobori.[3] Bar heights denote mean and error bars 95% CI. p-values were calculated between the two compared data sets (highest and lowest) for each amino acid using the t-test. Protein concentration is in molecules/fL.

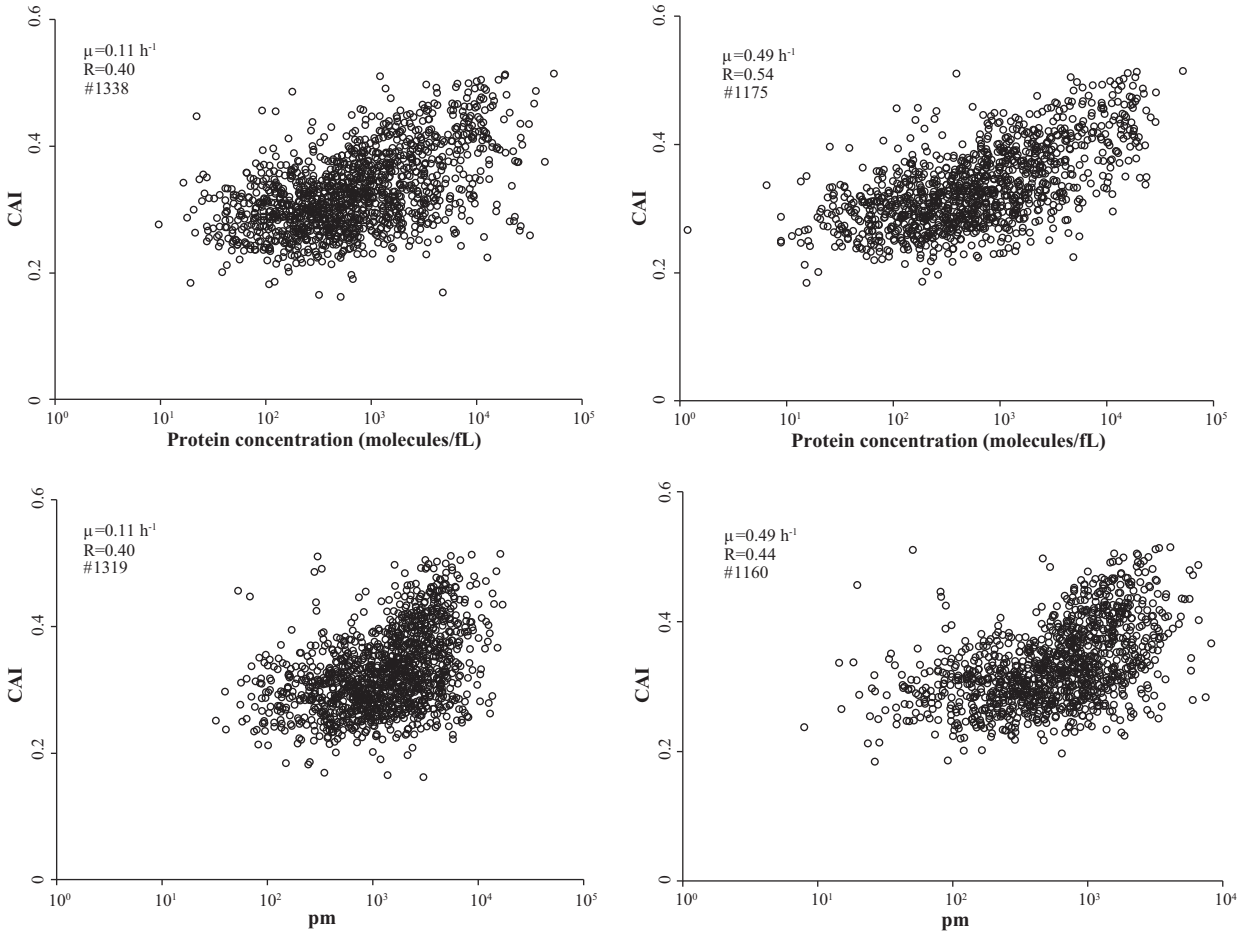
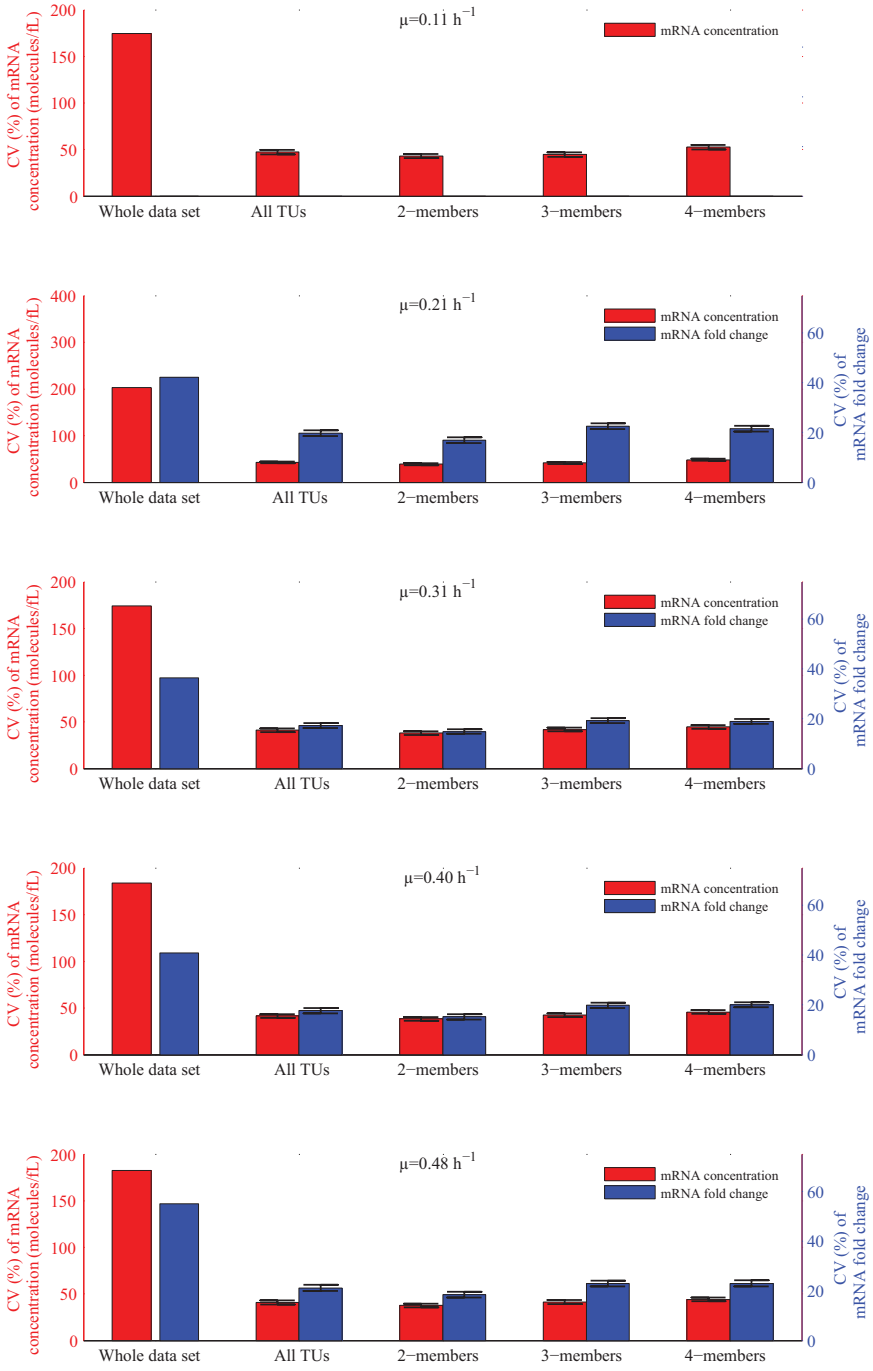


Fig. S12 Correlation of CAI with protein concentration and pm at low and high μ. CAI values for each protein were calculated based on protein abundances (molecules/cell) as follows: proteins which mass accounted for more than 0.5% of the whole proteome mass were chosen for codon usage table calculation with EMBOSS online tool;[4] the acquired codon usage table was used for calculating CAI values using the Seqinr package[5] in the R environment. CAI values for each protein can be seen in Table S2. R, PCC; #, number of analyzed genes.

A



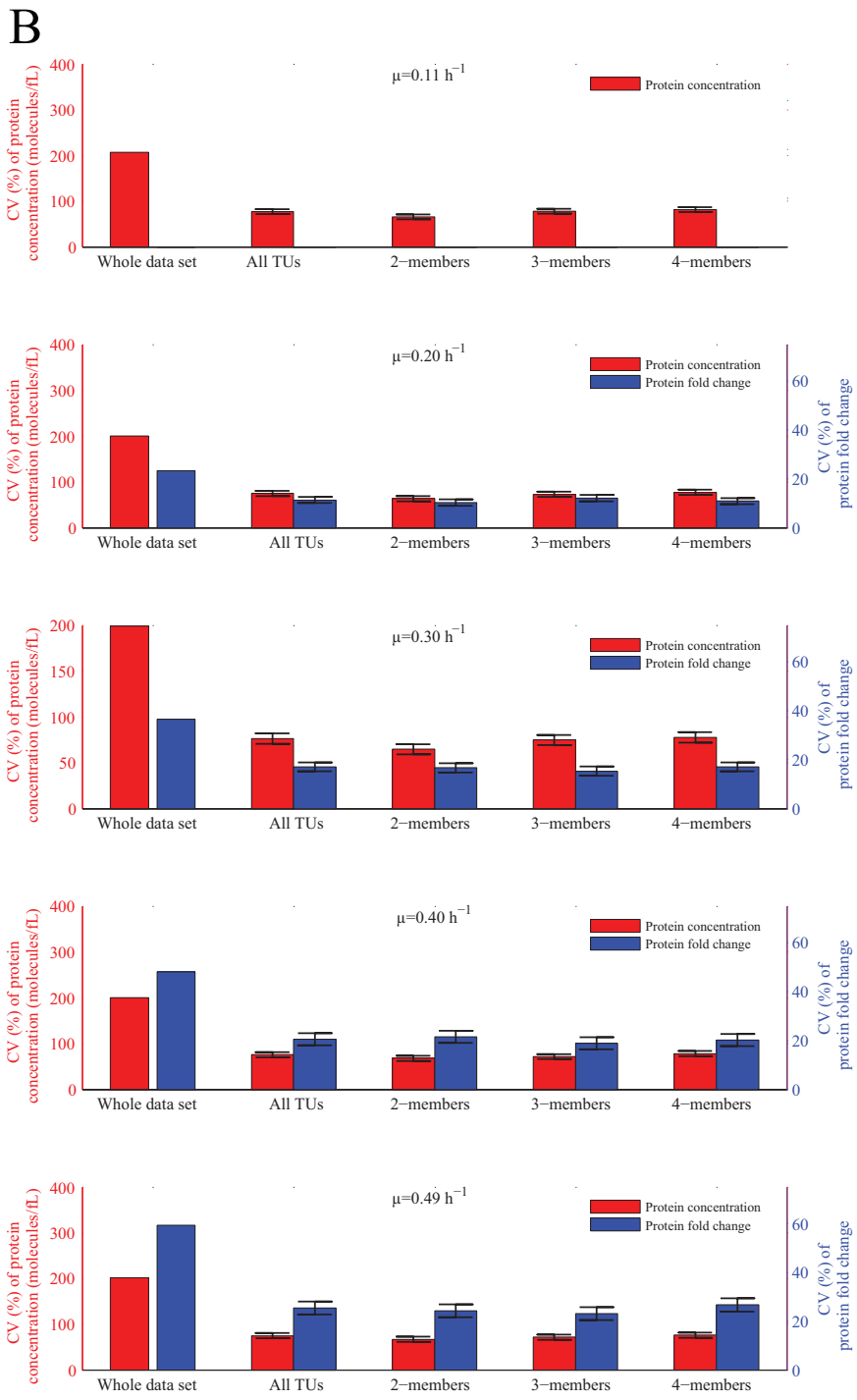


Fig. S13 Variance of mRNA and protein expression within TUs compared to the whole data set under the studied range of μ. (A) mRNA concentration and fold change CV within TUs compared to the whole data set. (B) Protein concentration and fold change CV within TUs compared to the whole data set. Fold change is calculated by comparing mRNA or protein concentration at each μ to the value at reference μ=0.11 h^{-1}. Bar height indicates the mean and error bars 95% CI. All quantified mRNAs and proteins were divided into transcription units according to the EcoCyc database.[2]

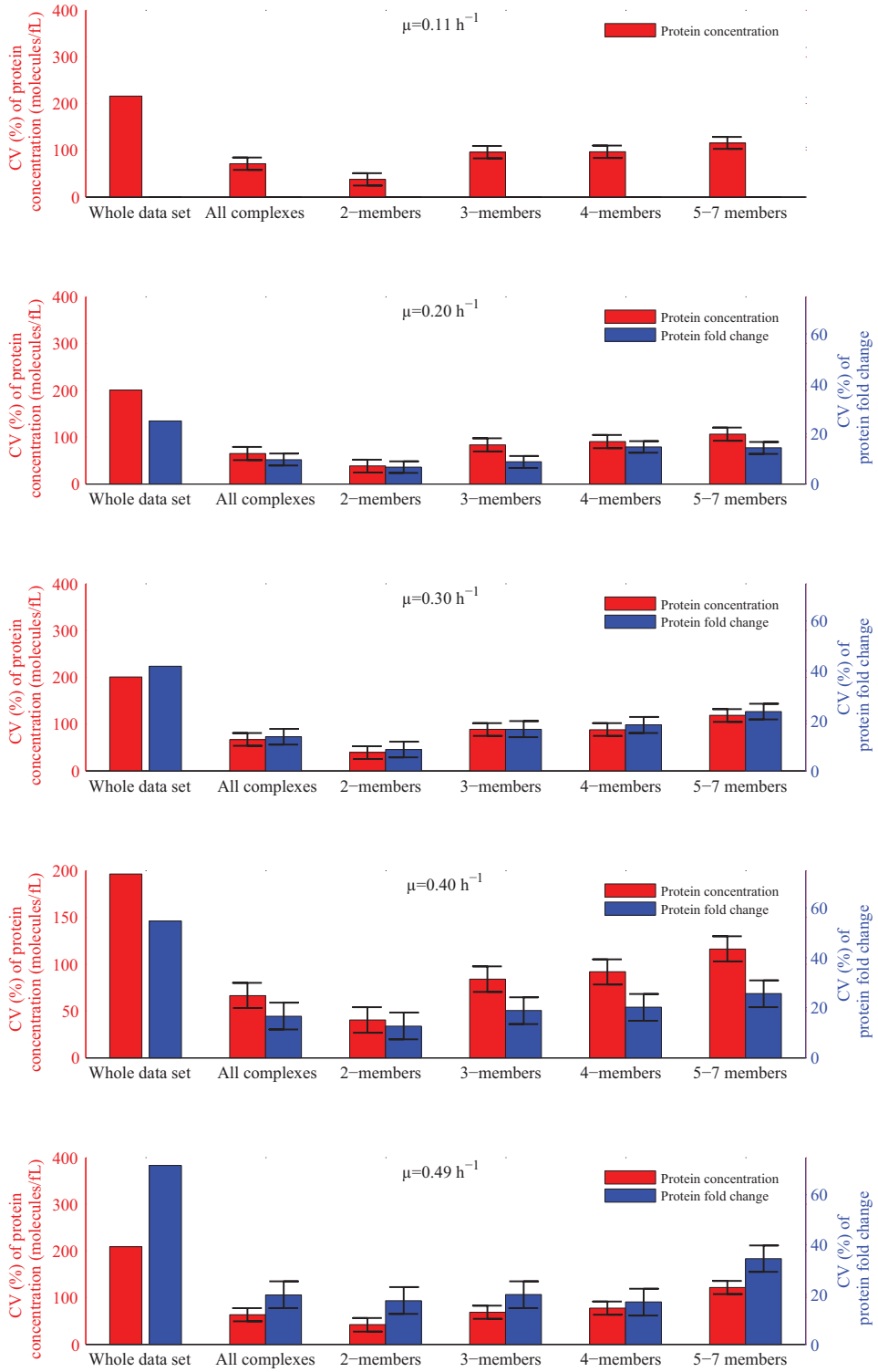


Fig. S14 Variance of protein expression within protein complexes compared to the whole data set under the studied range of μ. Fold change is calculated by comparing protein concentration at each μ to the value at reference μ=0.11 h^{-1}. Bar height indicates the mean and error bars 95% CI. All quantified proteins were divided into functional protein complexes according to the EcoCyc database.[2]

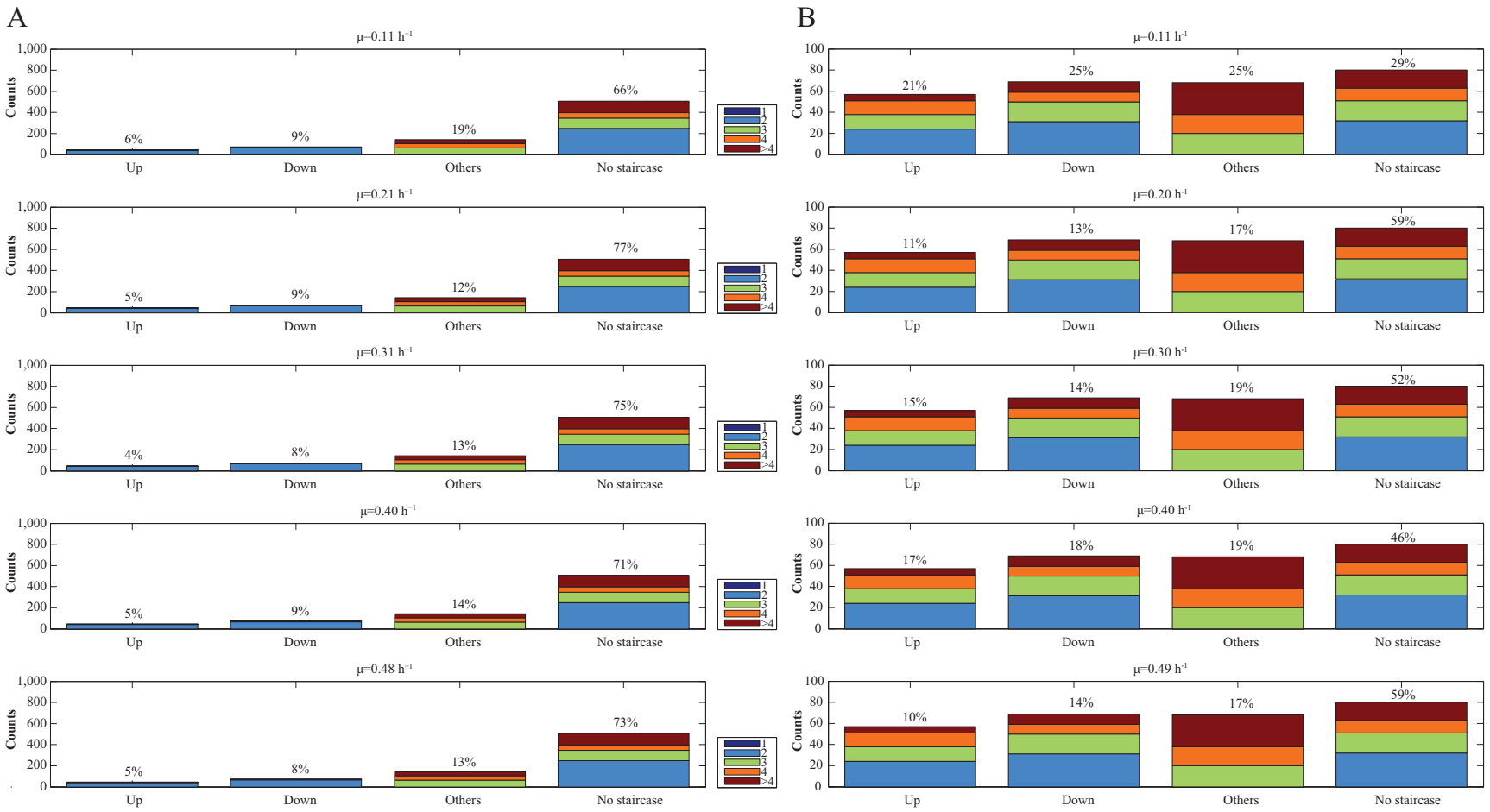


Fig. S15 Distribution of mRNA and protein expression in polycistronic TUs by staircase-behavior type under the studied range of μ. (A) Distribution of mRNA expression in polycistronic TUs by staircase-behavior type. (B) Distribution of protein expression, assigned as translation products of the mRNAs in TUs, in polycistronic TUs by staircase-behavior type. The percentage of TUs with each staircase-behavior type from the total TU amount is shown above the bar. All quantified mRNAs and proteins were divided into transcription units according to the EcoCyc database.[2] Refer to Experimental for details about distribution of TUs among staircase-behavior types. See Fig. S16 for a visual description of staircase-behavior types.

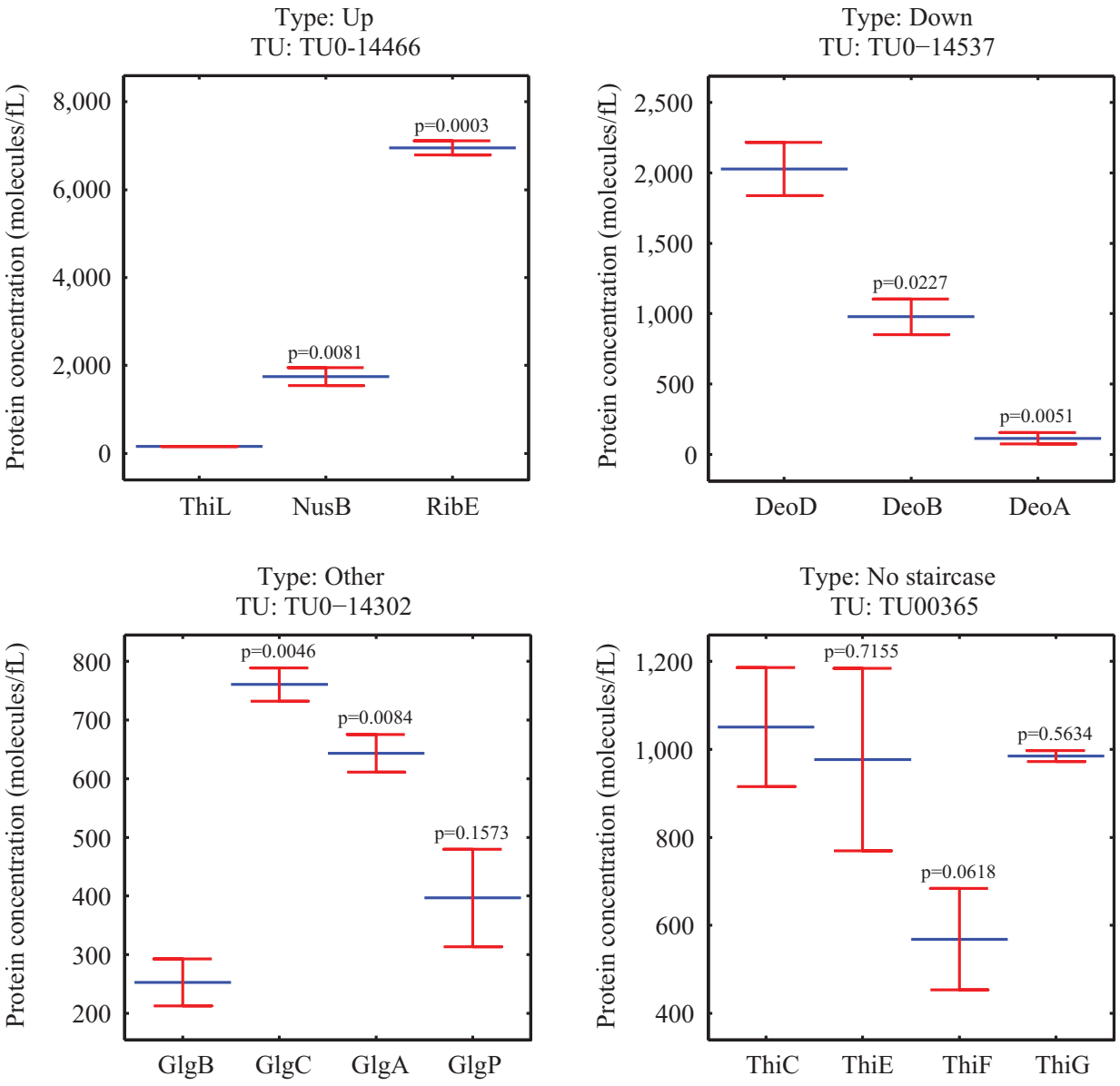


Fig. S16 Examples of the different TU staircase-behavior types determined in this study. Blue lines denote the mean and error bars average absolute deviation. p-values were calculated using the Z-test and show the statistical significance between protein concentrations of the first gene, starting from the direction of transcription, and all other genes in the TU. Refer to Experimental for details about distribution of TUs among staircase-behavior types.

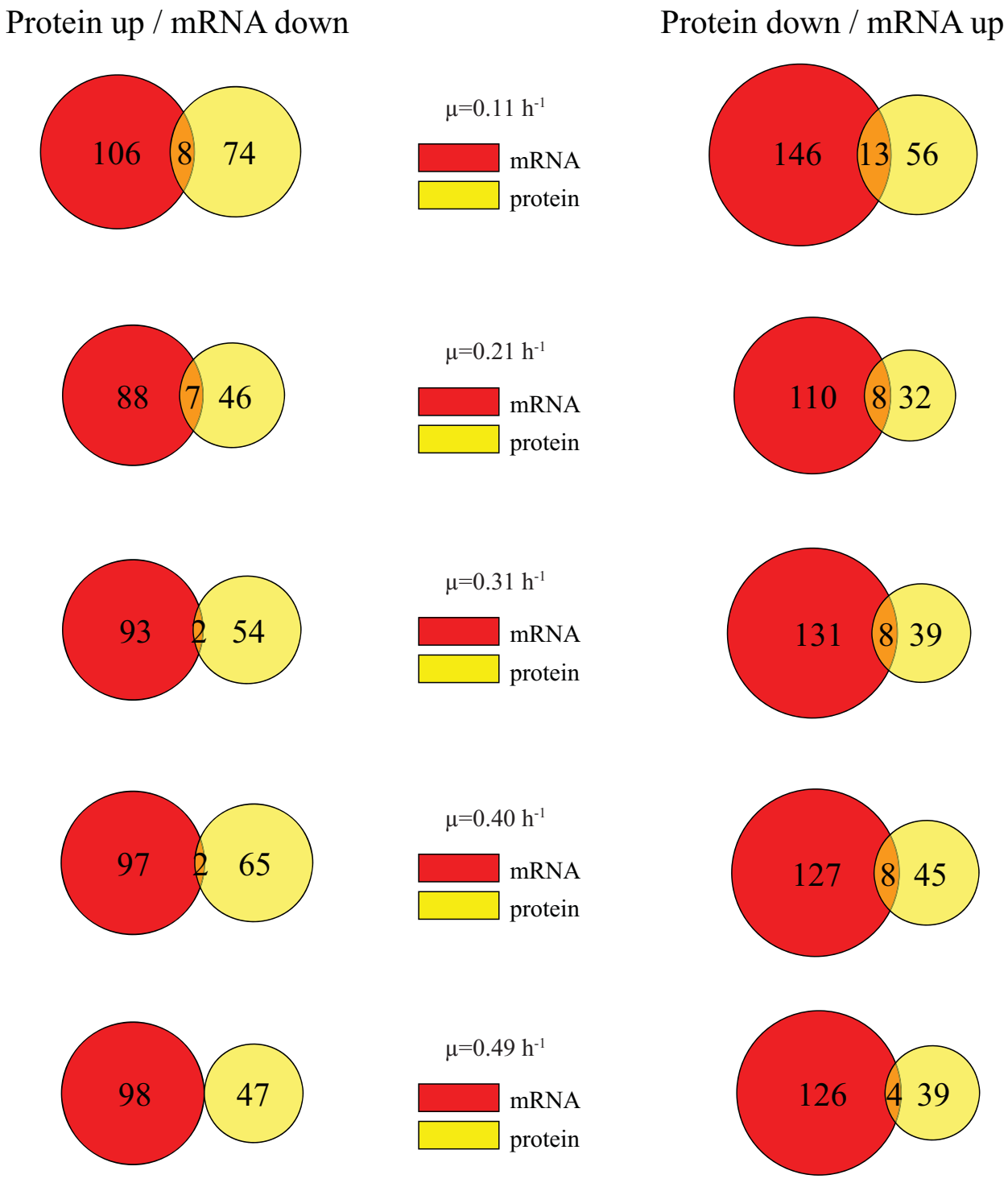


Fig. S17 Venn diagrams for compensatory expression of mRNAs and proteins within polycistronic TUs. Overlapping area indicates TUs which mRNA and protein expression patterns showed opposite staircase-behavior types within the TU.

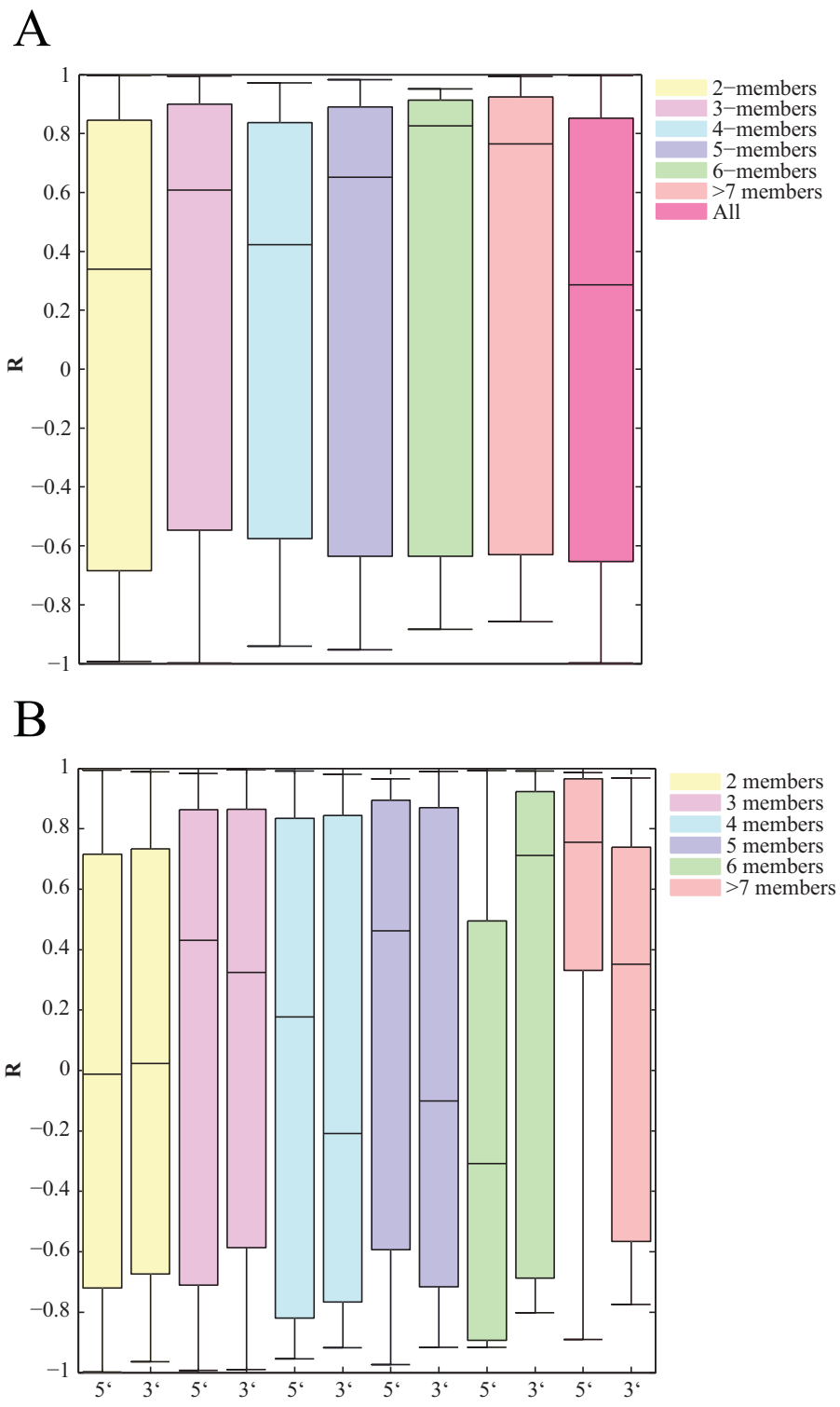


Fig. S18 Correlation of mRNA and protein concentration within polycistronic TUs. (A) Dependence of mRNA and protein concentration correlation on the polycistronic TU length. (B) Dependence of mRNA and protein concentration correlation on gene position in polycistronic TUs. Horizontal bars represent 25th, 50th (median) and 75th percentiles, and whiskers represent the highest and lowest values. R, PCC.

Supplementary References

1 R. L. Tatusov, N. D. Fedorova, J. D. Jackson, A. R. Jacobs, B. Kiryutin, E. V. Koonin, D. M. Krylov, R. Mazumder, S. L. Mekhedov, A. N. Nikolskaya, B. S. Rao, S. Smirnov, A. V. Sverdlov, S. Vasudevan, Y. I. Wolf, J. J. Yin and D. A. Natale, *BMC Bioinformatics,* 2003, **4**, 41.

2 I. M. Keseler, A. Mackie, M. Peralta-Gil, A. Santos-Zavaleta, S. Gama-Castro, C. Bonavides-Martínez, C. Fulcher, A. M. Huerta A. Kothari, M. Krummenacker, M. Latendresse, L. Muñiz-Rascado, Q. Ong, S. Paley, I. Schröder, A. G. Shearer, P. Subhraveti, M. Travers, D. Weerasinghe, V. Weiss, J. Collado-Viles, R. P. Gunsalus, I. Paulsen and P. D. Karp, *Nucleic Acids Res.,* 2013, **41**, D605–12.

3 H. Akashi and T. Gojobori, *Proc. Natl. Acad. Sci. USA,* 2002, **99**, 3695–3700.

4 P. Rice, I. Longden and A. Bleasby, *Trends. Genet.,* 2000, **16**, 2–3.

5 D. Charif, J. Thioulouse, J. R. Lobry and G. Perrière, *Bioinformatics,* 2005, **21**, 545–547.

6 K. Valgepea, K. Adamberg and R. Vilu, *BMC Syst. Biol.,* 2011, **5,** 106.