

Supplemental information

**Cell-to-cell variability in JAK2/STAT5 pathway
components and cytoplasmic volumes defines
survival threshold in erythroid progenitor cells**

Lorenz Adlung, Paul Stapor, Christian Tönsing, Leonard Schmiester, Luisa E. Schwarzmüller, Lena Postawa, Dantong Wang, Jens Timmer, Ursula Klingmüller, Jan Hasenauer, and Marcel Schilling

1 Supplemental Figures

1.1 Figure S1

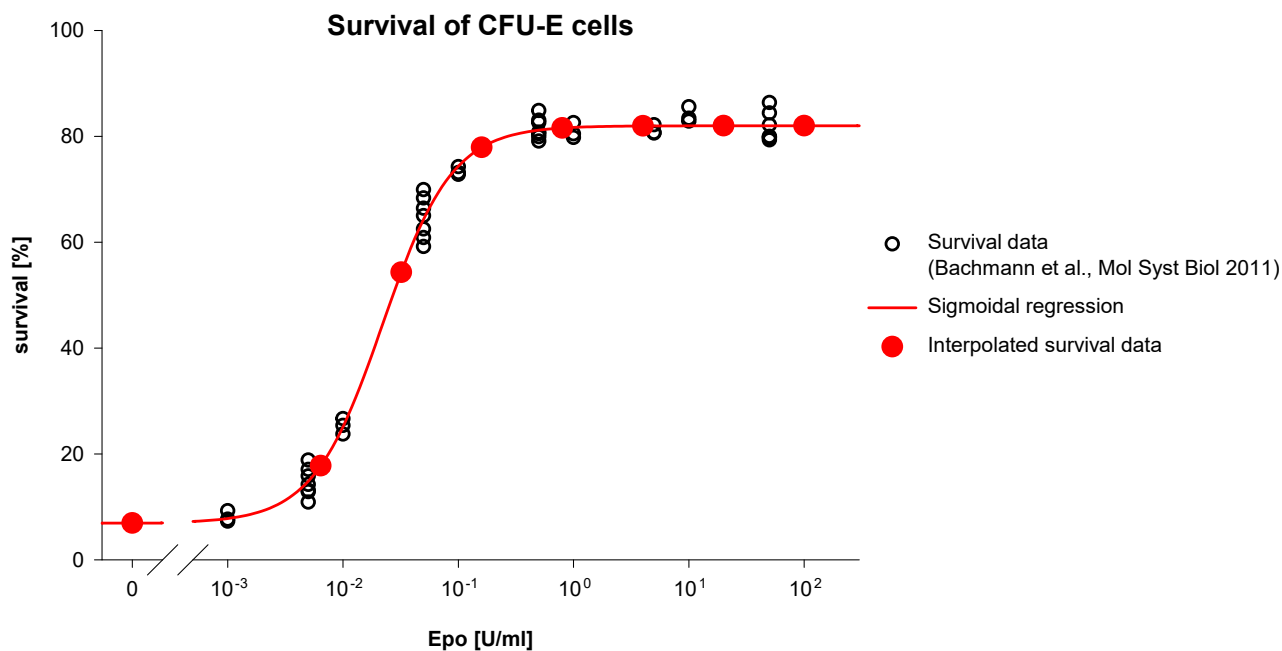


Figure S1 (related to Figure 1): Survival data for CFU-E cells. The percentage of surviving CFU-E cells cultured 24 h in various Epo concentration measured by TUNEL assay as reported in Bachmann et al., Mol Syst Biol 2011 are shown as open black circles. A sigmoidal regression based on a four parameter Hill function was used (solid red line) to interpolate survival data for the Epo concentrations used in Figure 1 (solid red circles).

1.2 Figure S2

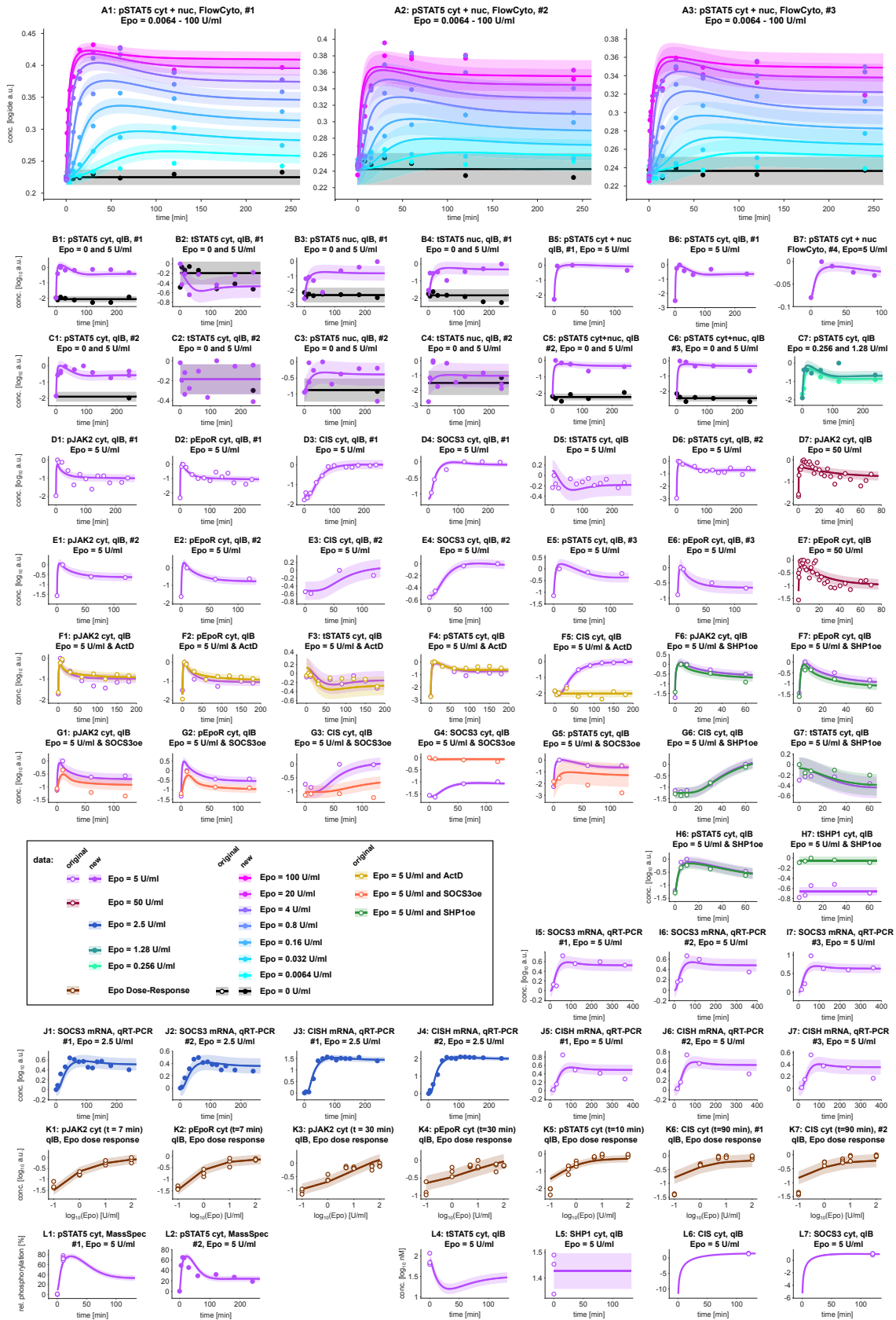


Figure S2 (related to Figure 2): Complete data set and model output of population-average model dynamics. Empty circles indicate original data sets from Bachmann et al., Mol Syst Biol 2011, filled circles is new experimental data. All data shown enter the analyses. Model simulations are shown as solid lines in the color of the corresponding experimental conditions. Shading depicts the error model. Measured entities, experimentation techniques, replicate number (rep #) and experimental conditions as indicated in the subplot titles. cyt: only cytoplasmatic fraction, nuc: only nuclear fraction, cyt + nuc: whole cell lysates. FlowCyto: Flow cytometry (A1-3, B7), qIB: quantitative immunoblotting (B1-H7, L4-L7), qRT-PCR: quantitative Real-Time Transcription PCR (I5-7, J1-7), MassSpec: Mass spectrometry (L1-L2).

1.3 Figure S3

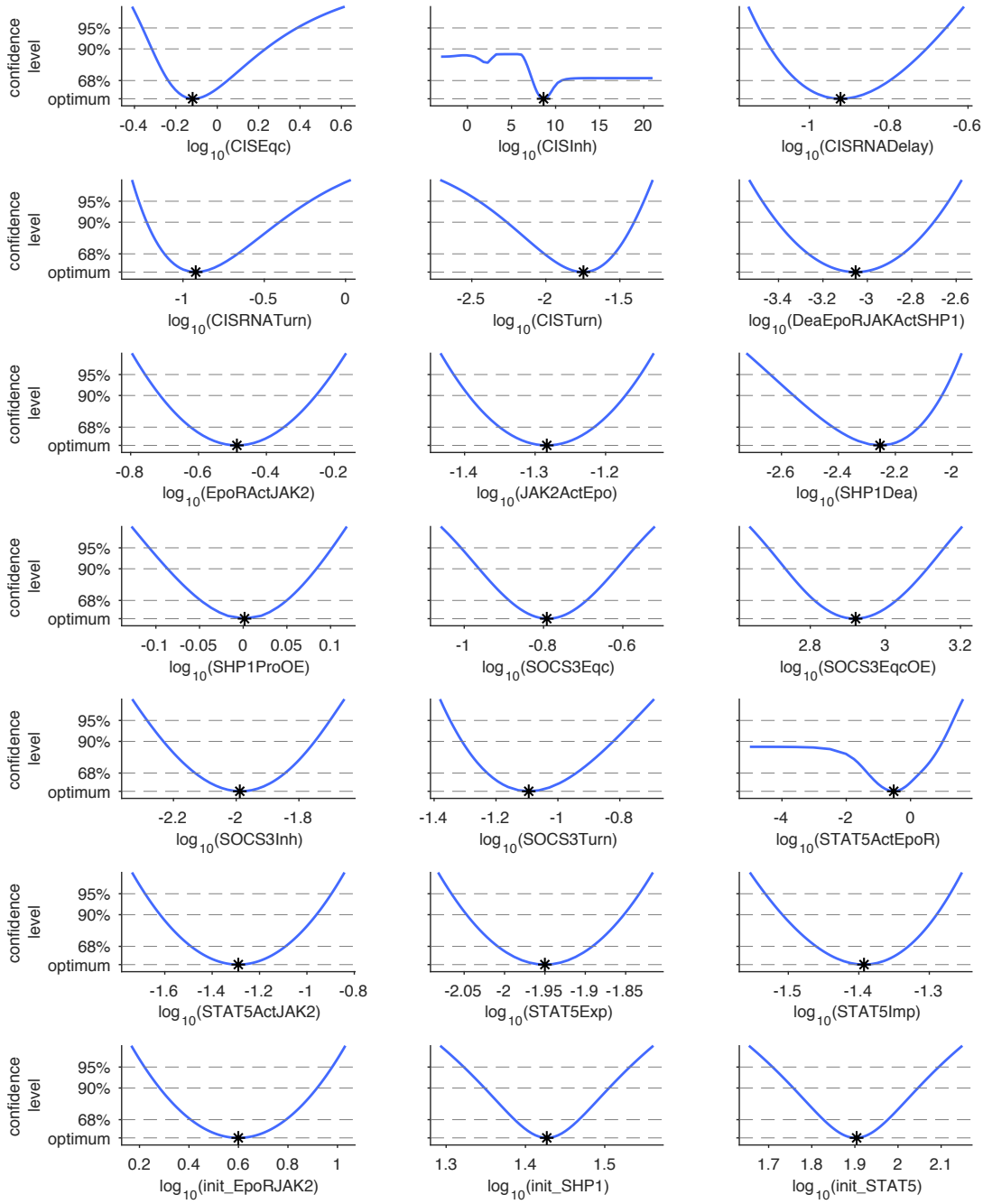


Figure S3 (related to Figure 2): Likelihood profiles of the pathway model with population-level data for all dynamic model parameters (blue lines). Black asterisks indicate the best fit and global optimum. Intersections of the likelihood profiles (blue solid lines) with the statistical thresholds for different confidence levels (gray dashed horizontal lines) indicate the profile likelihood-based confidence intervals. All parameters are identifiable with confidence levels of 68%. Only parameters CISInh and STAT5ActEpoR reveal practical non-identifiabilities for higher confidence levels. These non-identifiabilities could be resolved only by further measurements but they do not weaken the predictive power of the model as a unique optimum of the model parameters is found.

1.4 Figure S4

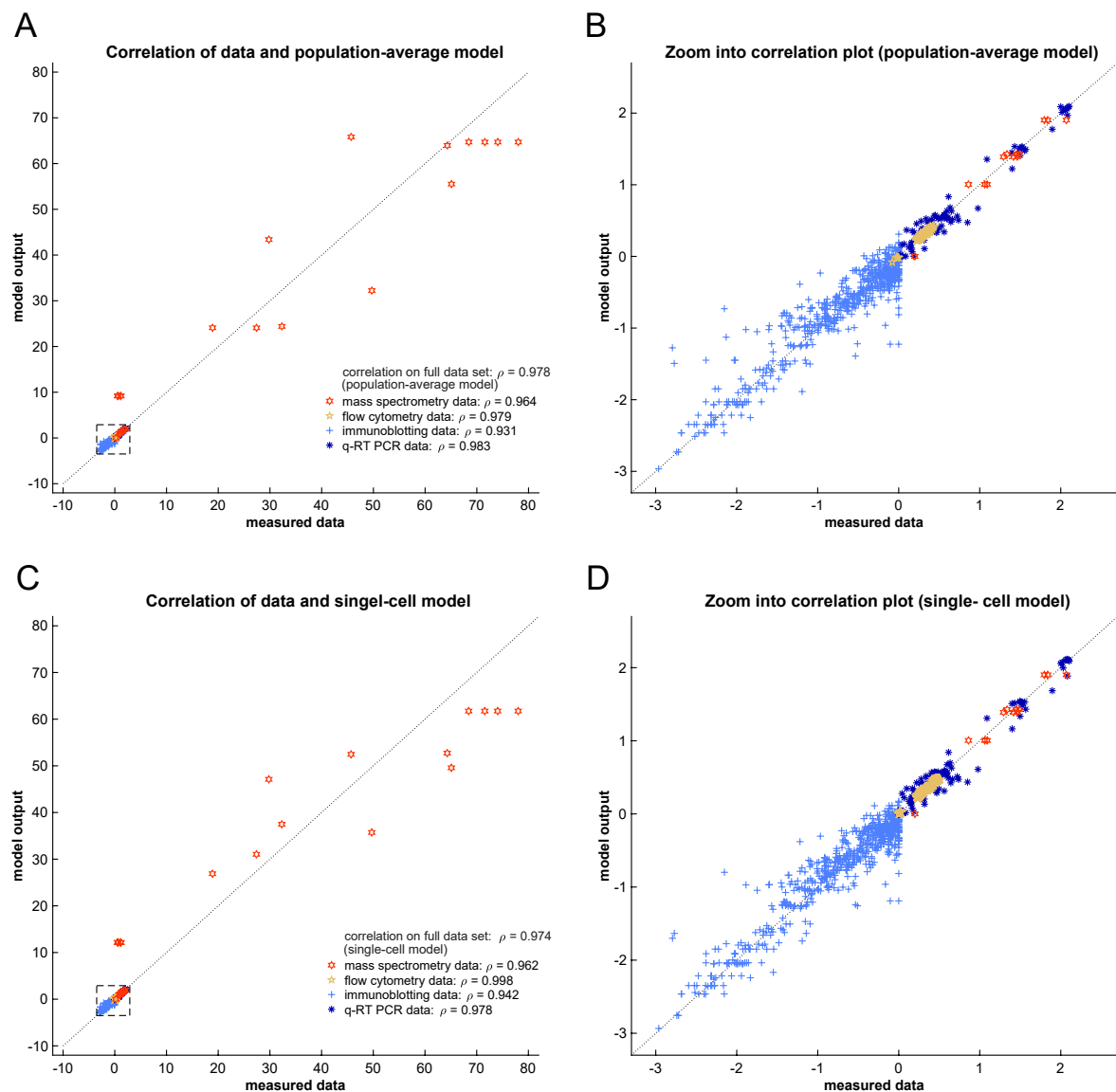
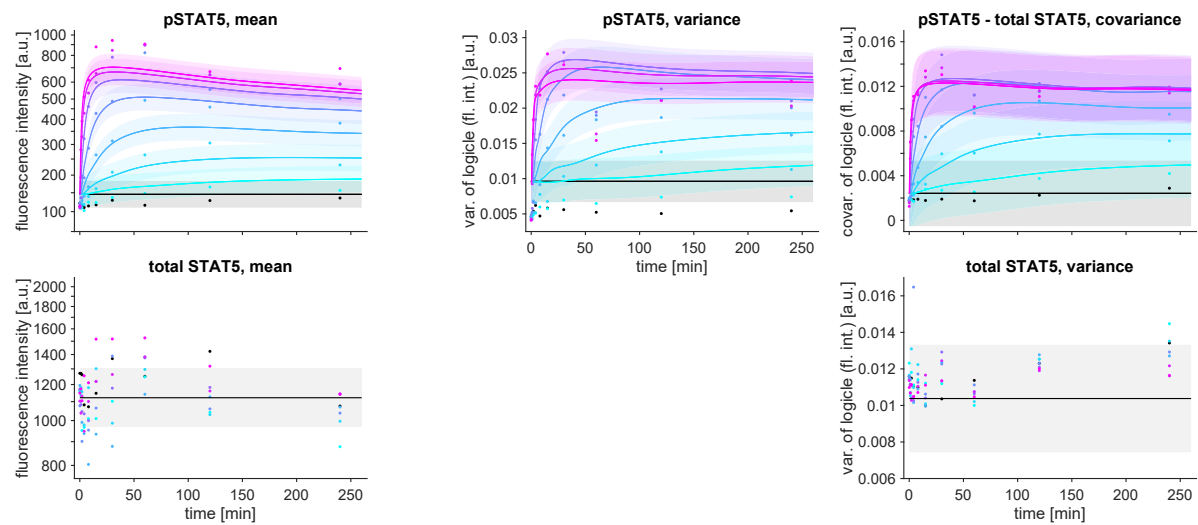


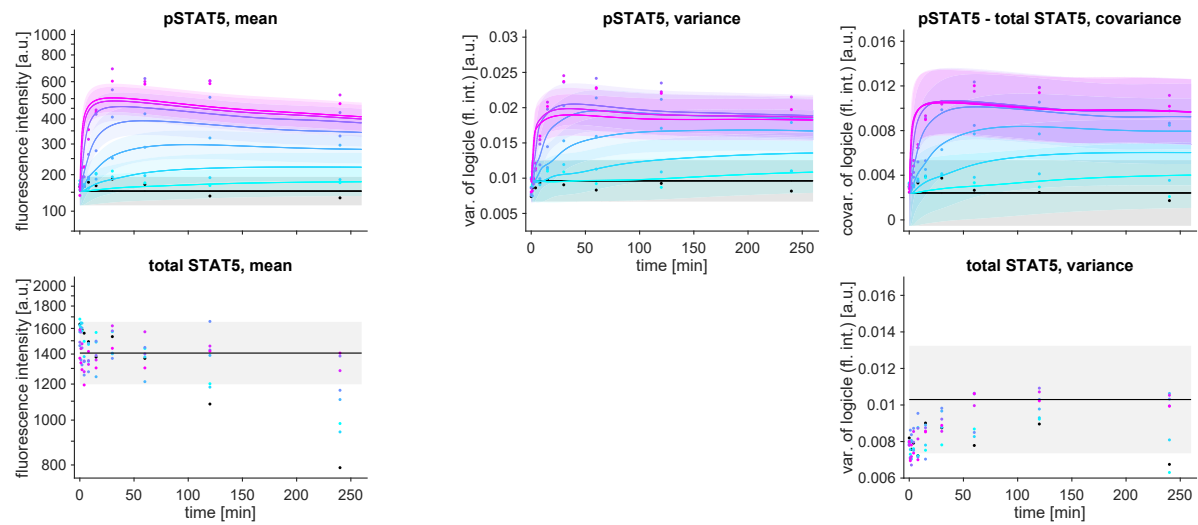
Figure S4 (related to Figure 4): Correlation of experimental data and model output. **(A)** Full data set for population-average model. Red hexagams: mass spectrometry data (non-normalized, linear scale), yellow pentagams: flow cytometry data (non-normalized, logicle scale), light blue plus signs: immunoblotting data (normalized, log-scale), deep blue asterisks: qRT-PCR data (non-normalized, log-scale). **(B)** Zoom into correlation plot for population-average model. **(C)** Full data set for single-cell model. **(D)** Zoom into correlation plot for single-cell model.

1.5 Figure S5

Flow cytometry, time series 1, STAT5 in nucleus and cytoplasm



Flow cytometry, time series 2, STAT5 in nucleus and cytoplasm



Flow cytometry, time series 3, STAT5 in nucleus and cytoplasm

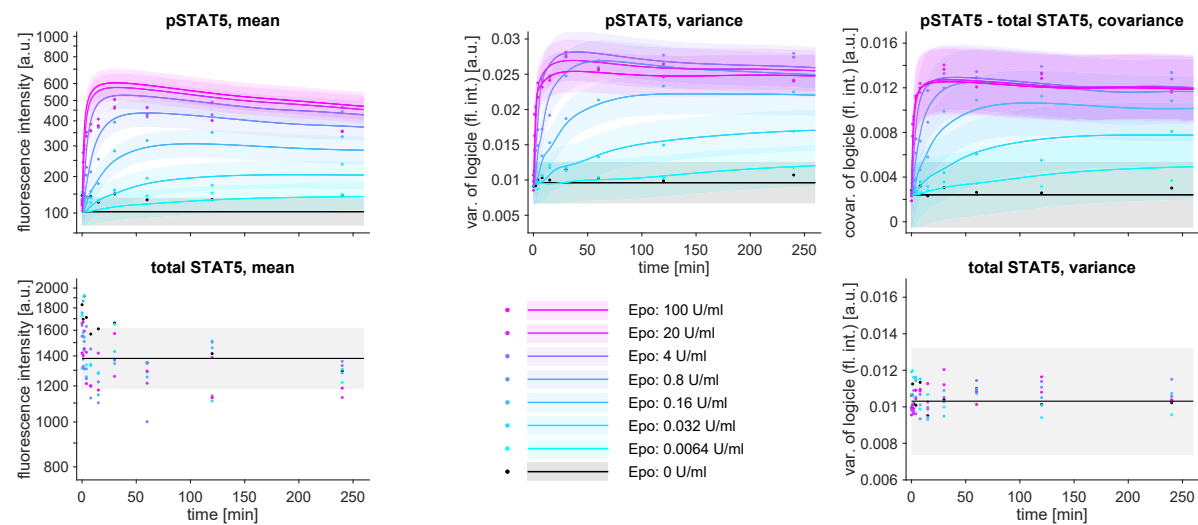


Figure S5 (related to Figure 5): Flow-cytometry data and model fit for the selected single-cell model. Three data sets were measured for eight Epo doses (colors), using flow cytometry and time-resolved experiments. Each experiment shows time-courses of population means (left) of pSTAT5 (upper panel) and total STAT5 (lower panel), and time-course variances and covariance of both species plotted as correlation matrix trajectories (right).

1.6 Figure S6

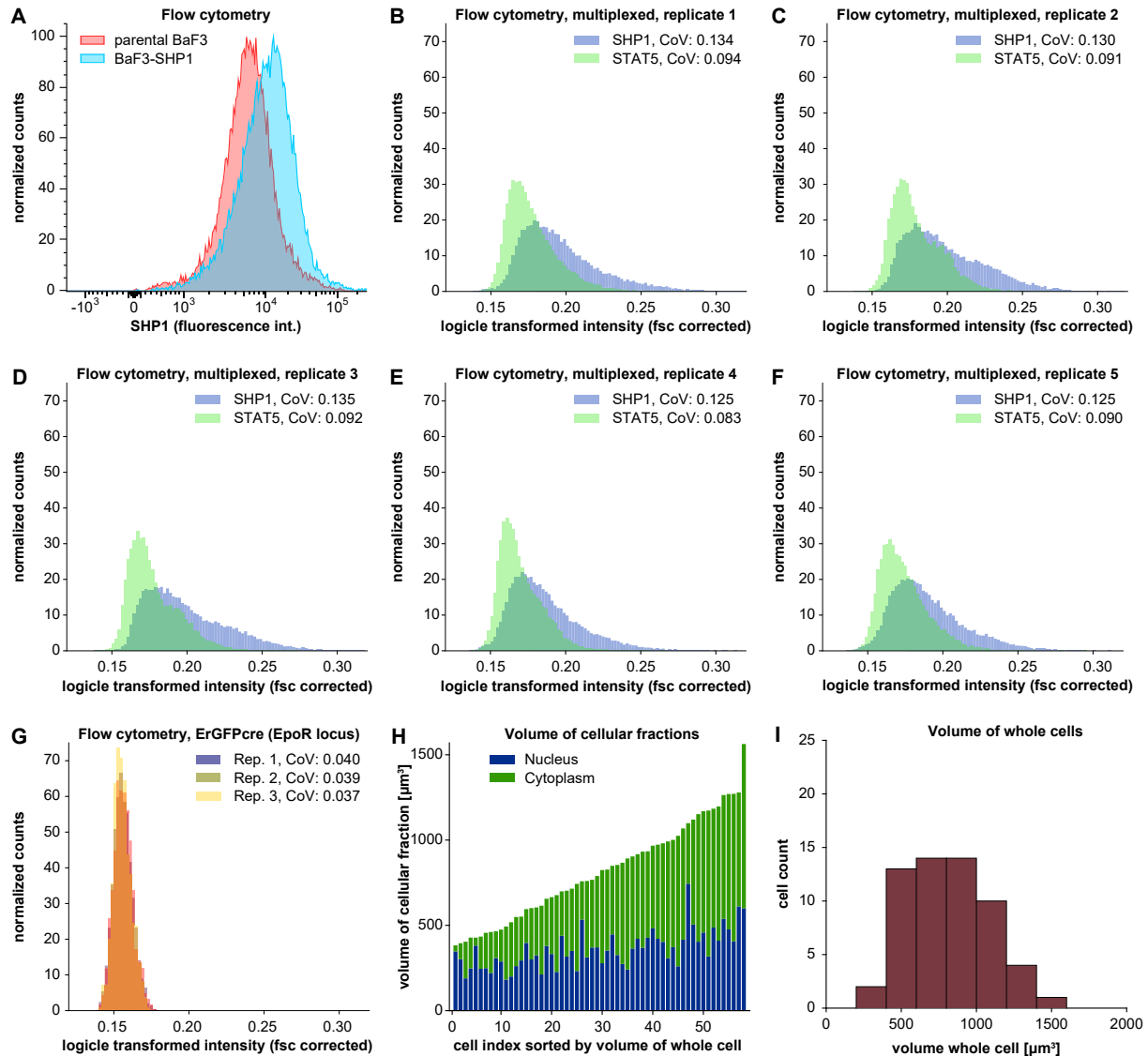


Figure S6 (related to Figure 6): Validation experiments on cell-to-cell variability of protein abundances by flow cytometry and of cytoplasmic and nuclear volumes by confocal microscopy and 3D reconstruction. **(A)** Flow cytometry of SHP1 was performed in parental BaF3 cells and in cells overexpressing SHP1 via a retroviral vector (BaF3-SHP1). **(B-G)** Flow cytometry was performed in unstimulated growth factor-depleted CFU-E cells. Measurement values were transformed to logicle scale, after being normalized with the respective front scatter (fsc) value. The normalization accounts for variable cell sizes and should reflect the actual concentrations of the measured proteins better than the raw measurement values. All measurements were transformed to logicle scale using the same hyperparameters, in order to ensure comparability of biological variability indicated as coefficient of variation (CoV). **(B-F)** Replicate measurements of multiplexed SHP1-STAT5 flow cytometry. **(G)** Three replicate measurements of GFP, using a knock-in mouse model that expresses an GFPcre fusion protein controlled by the endogenous EpoR promoter (ErGFPcre) as a proxy for the actual Epo receptor. **(H)** Analysis of cytoplasmic and nuclear volumes from unstimulated growth factor-depleted CFU-E cells from confocal microscopy and 3D reconstruction. Measured nucleus and cytoplasm volumes from $n = 58$ cells. **(I)** Distribution of whole cell volumes.

1.7 Figure S7

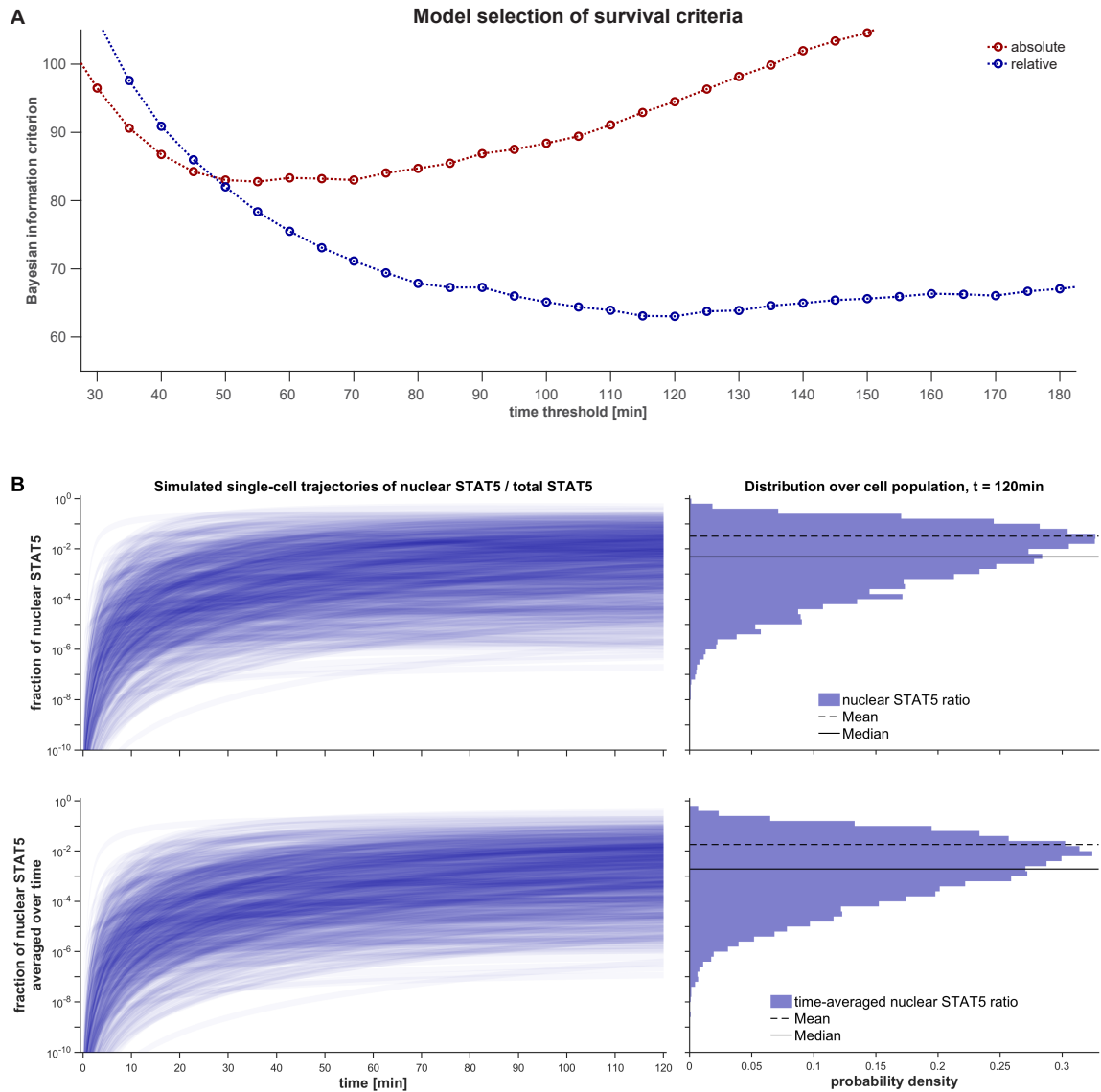


Figure S7 (related to Figure 7): Model simulations to link the amount of pSTAT5 in the nucleus to cell survival. **(A)** BIC results for model selection for different possible models of cell survival. pSTAT5 must be in the nucleus to ensure cell survival. An absolute criterion, i.e., the abundance of STAT5 in the nucleus (red) and a relative criterion, i.e., the percentage of a cell's STAT5 in the nucleus (blue), were compared against each other for different time intervals. Survival thresholds and offset parameters were fitted and the fit quality was compared. The relative criterion for a time interval of 120 minutes provided the best results. **(B)** Trajectories (left panel) for an in-silico population of 10 000 cells at EC_{50} -dose of Epo, which show the fraction of pSTAT5 in the cell's nucleus (upper panel) and the time-average of this fraction from $t = 0$ to the current time (lower panel). The right panel shows the histograms over the population for the two quantities at $t = 120\text{ min}$, which demonstrate the log-normal distribution of the survival signal. The median in the lower right panel represents the survival threshold.