Supplementary Material (ESI) for Molecular BioSystems

Cytokine-Associated Drug Toxicity in Human Hepatocytes Is Associated with Signaling Network Dysregulation

Benjamin D. Cosgrove, Leonidas G. Alexopoulos, Ta-chun Hang, Bart S. Hendriks, Peter K. Sorger, Linda G. Griffith, and Douglas A. Lauffenburger

Supplementary Material

Supplementary Experimental Methods

Hepatotoxic drugs

Drugs with idiosyncratic and/or inflammation-associated hepatotoxicity and corresponding 'comparison' compounds were classified as described previously.[1] Clarithromycin serves as a less, but still hepatotoxic, 'comparison' compound to telithromycin.[2, 3] Compounds were obtained from Sigma (St. Louis, MO; cimetidine, ranitidine, levofloxacin, buspirone, nefazodone, aspirin, nimesulide, chlorpromazine, nortriptyline, clomipramine, and riluzole), Sequoia Research Products (Pangbourne, UK; clarithromycin and telithromycin), or Pfizer's chemical sample bank (Groton, CT; trovafloxacin). Drugs were dosed at therapeutically appropriate drug exposure levels were defined by average plasma maximum concentration (C_{max}) values observed in humans upon single- or multi-dose administration at commonly recommended therapeutic doses. C_{max} values were previously[1, 4] estimated from a combination of literature searches and available databases. A concentration of 100*C_{max}, encompassing a scaling factor to account for human population pharmacokinetic and toxicodynamic variabilities, was considered as a therapeutically relevant dosing limit for each drug, as previously discussed.[1, 4] 100*C_{max} concentrations correspond to the following molecular concentrations: 1.5 mM cimetidine, 142 µM ranitidine, 1.6 mM levofloxacin, 770 µM trovafloxacin, 0.46 µM buspirone, 86 µM nefazodone, 552 µM aspirin, 2.1 mM nimesulide, 334 µM clarithromycin, 277 µM telithromycin, 111 µM chlorpromazine, 10 µM nortriptyline, 13 µM clomipramine, and 107 µM riluzole.

Multiplexed phosphoprotein assays

Phosphoprotein signaling was quantified using multiplexed bead-based Luminex assays. Cells were plated and treated as described above. Cell lysates were collected at 0 and 20 min and 4, 24, and 48 h following drug and/or cytokine stimulation. At the desired time point, cells were placed on ice and culture medium was removed. Matrigel overlays were partially dissolved by adding ice cold PBS for 15 min at 4°C. PBS was removed and cells were lysed with Phosphoprotein Lysis Buffer (Bio-Rad, Hercules, CA) for 20 min at 4°C. Lysates were collected by scrapping and vigorous pipetting. Lysates were clarified by centrifugation at 16,000*g* for 15 min at 4°C. Clarified lysates were analyzed using a bicinchonicic assay (Pierce, Rockford, IL) to determine the total protein concentration. In each culture plate, a well without cells was maintained, lysed, and analyzed to calculate the protein contribution from the Matrigel overlay alone and estimate the cellular protein concentration in the other wells. Bio-Plex bead-based assays (Bio-Rad) were used to quantify the following 17 phosphoproteins: p-Akt (Ser^{473}), p-CREB (Ser^{133}), p-c-Jun (Ser^{63}), p-GSK-3α/β (Ser^{21}/Ser^{9}), p-IκB-α (Ser^{32}/Ser^{36}), p-IRS-1 (Ser^{636}/Ser^{639}), p-ERK1/2 (Thr^{202}/Tyr^{204}, Thr^{185}/Tyr^{187}), p-Histone H3 (Ser^{10}), p-HSP27 (Ser^{78}), p-JNK (Thr^{183}/Tyr^{185}), p-MEK1 (Ser^{217}/Ser^{221}), p-STAT3 (Ser^{727}), p-STAT6 (Tyr^{641}), p-p38 (Thr^{180}/Tyr^{182}), p-p53 (Ser^{15}), p-p70 S6 kinase (Thr^{421}/Ser^{424}), and p-p90 RSK (Thr^{359}/Ser^{363}). Bio-Plex assays were conducted per manufacturer's recommendations on a Luminex 200 instrument (Luminex) with protein lysates loaded at 10 µg per well in technical duplicate, which were averaged for each biological sample. Multiple positive control treatments were loaded on each assay plate to linearly scale raw fluorescence data to self-consistent relative values. Phosphoprotein assays were validated for high specificity to known activating conditions (Fig. S1).

Kinase inhibitor evaluation and selection

Kinase inhibitors were evaluated for efficacy and toxicity in human hepatocytes (from donor #3) over a range of concentrations at seven 8× serial dilution concentrations from 20 µM (20 µM, 2.5 µM, 0.31 µM, 39 nM, 4.9 nM, 0.61 nM, 76 pM). To evaluate MEK kinase inhibitor efficacy, human hepatocytes were pretreated with inhibitor for 1 h before treatment with 100 ng ml^{-1}

recombinant human TGF-α (R&D Systems, Minneapolis, MN) for 15 min and then were assayed for p-ERK1/2 activation using a bead-based phosphoprotein assay. To evaluate p38 kinase inhibitor efficacy, human hepatocytes were pretreated with inhibitor for 1 h before treatment with 100 ng ml^{-1} recombinant human TNF (R&D Systems) for 15 min and then were assayed for p-HSP27 activation using a bead-based phosphoprotein assay. To evaluate MEK and p38 kinase inhibitor toxicity, inhibitors were added at 1× final concentrations in fresh medium for 48 h, and then medium samples were assayed for LDH release. LDH results were normalized to wells from the same culture plate lysed in 1% Triton X for 10 min, and values were reported as % cell death (with the lysed samples assumed to represent 100% cell death). The following kinase inhibitors/concentrations were selected for their potent signaling inhibition and minimal toxicity and were used to perturb kinase activities in drug- and cytokine co-treatment experiments: 10 μM U0126, 1 μM PD325901, 1 μM PHA-666859, and 1 nM PHA-818637 (see Fig. S8).

Collection and normalization of signal-response data compendia

We collected a cue-signal-response (CSR) data compendia in human hepatocytes from two separate donors. In the initial data compendium (donor #1), human hepatocytes were treated with 66 different combinations of 11 'drug' conditions (six compounds with idiosyncratic hepatotoxicity and/or inflammation-associated hepatic cytotoxicity, four corresponding 'comparison' compounds, and a DMSO control) and six 'cytokine' conditions (no cytokine, IL-1α, LPS, TNF, IL-6, and a mix containing all three cytokines plus LPS). To broadly measure a diverse set of key phosphoprotein activities mechanistically connected to numerous drug- and/or cytokine-induced signaling pathways, we quantitatively assayed the aforementioned 17 phosphoproteins at both early (0 and 20 min) and delayed time-points (4, 24, and 48 h) following drug and/or cytokine stimulation. In this initial data compendium, single biological replicates were used for both phosphoprotein and LDH assays. The total number of individual phosphoprotein signaling measurements in the initial compendium was 4488 (66 conditions × 17 phosphoproteins × 4 time-points × 1 biological replicate).

In the second data compendium (donor #2), human hepatocytes were treated with 14 different combinations of seven 'drug' conditions (two hepatoxicants used in initial compendium, four idiosyncratic hepatoxicants not used in the initial compendium, and a DMSO control) and two 'cytokine' conditions (no cytokine and the 3-cytokine/LPS mix). In this second compendium, quantitative phosphoprotein assays were focused on a reduced set of six highly informative signals (p-MEK1, p-ERK1/2, p-Akt, p-70 S6K, p-p38, p-HSP27). These phosphoproteins were assayed at the same time-points as in the CSR from donor #1, but with some drug/cytokine co-treatment conditions (all those containing the DMSO control and trovafloxacin) also assayed at 1 and 12 h post-stimulation. Biological triplicates were used for both phosphoprotein and LDH assays. The total number of individual phosphoprotein signaling measurements in the second compendium was 1008 (14 conditions × 6 phosphoproteins × 4 time-points × 3 biological replicates).

Phosphoprotein data was fold-change normalized to untreated samples (at 0 min) for each phosphoprotein assay and separately for each hepatocyte donor. LDH release data was fold-change normalized to untreated samples at 48 h post-drug and/or cytokine stimulation separately for each hepatocyte donor. CSR data compendia normalization and fusion was performed using the DataRail toolbox[5] for Matlab (The Mathworks, Inc., Natick, MA). These normalized data are available in *Supplementary Data*.

Metric extraction and scaling

For each phosphoprotein signaling time-course, two time-dependent signaling metrics[6] were extracted: (i) the integral, or area-under-the-curve, for the entire time-course, and (ii) the average of the late time-points (4-48 h), reflecting the steady-state signaling level. These were added to the four time-points (20 min and 4, 24, and 48 h) to yield six signaling metrics for each

assayed phosphoprotein. For each compendia, the signaling metrics from all measured phosphoproteins were then fused into a signaling network data matrix (X). Separately, the toxicity response data were cast into a vector (Y), with both X and Y arrayed across all treatment conditions. In the compendium from donor #1, X was a matrix of 66 rows of treatments and 102 columns (17 phosphoproteins × 6 metrics) of signaling metrics, and Y was a vector of 66 rows of treatments and single column (LDH release measured at 48 h). Before modeling, all columns in the signaling data matrix and response vector were separately mean-centered and scaled to unit-variance to non-dimensionalize different assay measurement dynamic ranges.[7] In modeling test data sets not present in model training, scaling parameters from the training data set were used to scale the test data.

Signal-response data modeling through orthogonal partial-least squares regression

To relate the measured signaling and cell death response data, we assumed a linear relationship between the two data sets, such that:

$$Y = f(X) = X \cdot B,$$

where X is the signaling network data matrix, Y is the cell death response vector, and B is a vector of regression coefficients that reflect how each phosphoprotein signaling metric contributes to cell death. Framed as such, the signaling matrix X is a block of independent variables and the response vector Y is a block of dependent variables. Since the number of signaling metrics (columns of X) exceeds the number of treatment conditions (rows of X), an unique solution to this linear regression problem cannot be identified. Thus, we implemented partial least-squares regression (PLSR) to solve this regression problem. Instead of performing the linear regression in the original multi-dimensional data space, PLSR casts the problem in a principal-component space and regresses principal components-based coefficients associated with independent and dependent variables[8]. The calculation of principal components-based regression coefficients (or, loadings) is biased towards those signaling variables that are most covariant with the response data and to optimize prediction accuracy of the response data in cross-validation.

We implemented PLSR using the NIPALS algorithm in SIMCA-P software (Umetrics, Inc., Kinnelon, NJ) following standard methods[6-12], with some modifications. All models were generated using four principal components under standard optimization criteria.[8] Model calibration was conducted using leave-one-out cross-validation, and model uncertainties were calculated by jack-knifing.[13] Calibrated models were subjected to a principal-component space linear transformation by rotating the projection of the single cell death response variable completely into the first principal-component, thus yielding an orthogonal PLSR (OPLSR) model[14] (see Fig. S2), to allow for simplified interpretation of model loadings and scores. Signaling metric model loadings were calculated using the mean-centered regression coefficients $w_a * c_a$ from the *a*-th OPLSR principal-component.[6] The accuracy of model predictions for both training and test data were assessed using a Pearson correlation coefficient (R)[12] and a model fitness[7, 11] parameter (R^2):

$$R^2 = 1 - \left[\frac{\sum_{i=1}^{n} (\text{Predicted}_i - \text{Observed}_i)^2}{\sum_{i=1}^{n} (\text{Predicted}_i)^2 - \frac{\left(\sum_{i=1}^{n} \text{Predicted}_i \right)^2}{n}} \right],$$

where Predicted_i is the predicted cell death value of the *i*-th treatment condition, Observed_i is the experimentally observed cell death value of the *i*-th treatment condition, and n is the total

number of treatment conditions. This assessment of model fitness postulates a one-to-one equivalence between observed and predicted response values, and is more stringent than a simple correlation assessment that does not penalize for quantitatively inaccurate predictions that are nonetheless qualitatively correlative.[11] An R^2 value of 1 corresponds to a perfect fit between observed and predicted responses. An R^2 value of 0 corresponds a model break-point.[11] Negative R^2 values imply highly inaccurate model predictions.

To interpret the contributions of various signaling pathways to drug- and/or cytokine-induced hepatic cytotoxicity, an initial OPLSR model was trained on the 17-phosphoprotein, 66-condition CSR data compendium from human hepatocyte donor #1, and demonstrated good model fitness (R^2 = 0.92) of cross-validated predictions. All models were regressed against the LDH release data measured at t = 48 h, as models of the LDH release response at earlier time-points were poorly fit (data not shown).

Model reduction

To identify the relative importance of individual phosphoprotein signaling metrics, the information content of each signaling metric was assessed by its variable importance of projection (VIP) score[15]:

$$\mathrm{VIP}_k = \left(\frac{\mathrm{K} \sum_{a=1}^{A} w_{a,k}^2 \, \mathrm{SS}_a}{\sum_{a=1}^{A} \mathrm{SS}_a} \right)^{1/2},$$

where K is the total number of signaling metrics, $w_{a,k}$ is the weight of the *k*-th metric for principal component a, A is the total number of principal components, and SS_a is the sum of squares explained by principal component a . Signaling metrics with a VIP > 1 have significant importance in the model and metrics with a VIP << 1 significantly lack unique information in the model.[7, 15]

To reduce the initial 17-phosphoprotein model, phosphoproteins and all six of their associated signaling metrics were removed from the model step-wise in order of the lowest average VIP score across all six metrics. This model reduction approach yielded a set of 4-to-6-phosphoprotein models (R^2 = 0.87-0.91) that retained the model fitness of the full 17-phosphoprotein model (R^2 = 0.92; Fig. 3D). The robustness of this model reduction approach was examined by testing the ability of a reduced 6-phosphoprotein model, trained on CSR data from human hepatocyte donor #1, to accurately predict signal-response relationships in a 14-condition, 6-phosphoprotein CSR data compendium collected from human hepatocyte donor #2.

Model predictions of kinase inhibitor effects on drug- and cytokine-induced hepatic cytotoxicity

To make *a priori* predictions of kinase inhibitor perturbation of drug- and/or cytokine-induced hepatocellular death responses, a set of 'computationally inhibited' signaling time-courses was generated by reducing the activation levels of the specific phosphoprotein signaling molecules targeted by the kinase inhibitor of interest. These time courses were generated for the treatment conditions of DMSO ± cytokine mix and nortriptyline ± cytokine mix, in the presence or absence of 10 μM U0126, 1 μM PD325901, 1 μM PHA-666859, and 1 nM PHA-818637. To generate the uninhibited time-courses, mean values across donor #1 and #2 (DMSO ± cytokine mix) or from donor #2 only (nortriptyline ± cytokine mix) were used. For a phosphoprotein signals targeted by an inhibitor, the mean observed level at each time-point was reduced by a fraction equivalent to the percent signal reduction observed for that inhibitor in the signaling inhibition studies in donor #3 (Fig. S8), as in [10, 16]. For MEK inhibitors, the phosphoprotein levels of both MEK and ERK at all time-points were reduced by 70% (U0126) and 99% (PD325901),

but all other signaling proteins were not changed. For p38 inhibitors, the phosphoprotein levels of both p38 and HSP27 at all time-points were reduced by 93% (PHA-666859) and 99% (PHA-818637), but all other signaling proteins were not changed. After computationally inhibiting the time-point data, the integral and late average metrics were re-calculated. Predictions of kinase inhibitor effects based on these computationally inhibited signaling metric sets were generated from two different 6-phosphoprotein OPLSR models (featuring either non- or *log*-scaled LDH release response data) trained on a fused CSR data compendium from both donors #1 and #2. Logarithmic scaling of response data can provide more accurate OPLSR model prediction of low response level observations, especially for non-uniformly distributed response data sets, by removing regression bias to high response conditions (see Fig. S9A-B). Prediction accuracy was assessed by comparing to experimental observations collected in human hepatocytes from donor #4 (Fig. S9A-D).

Supplementary Figures

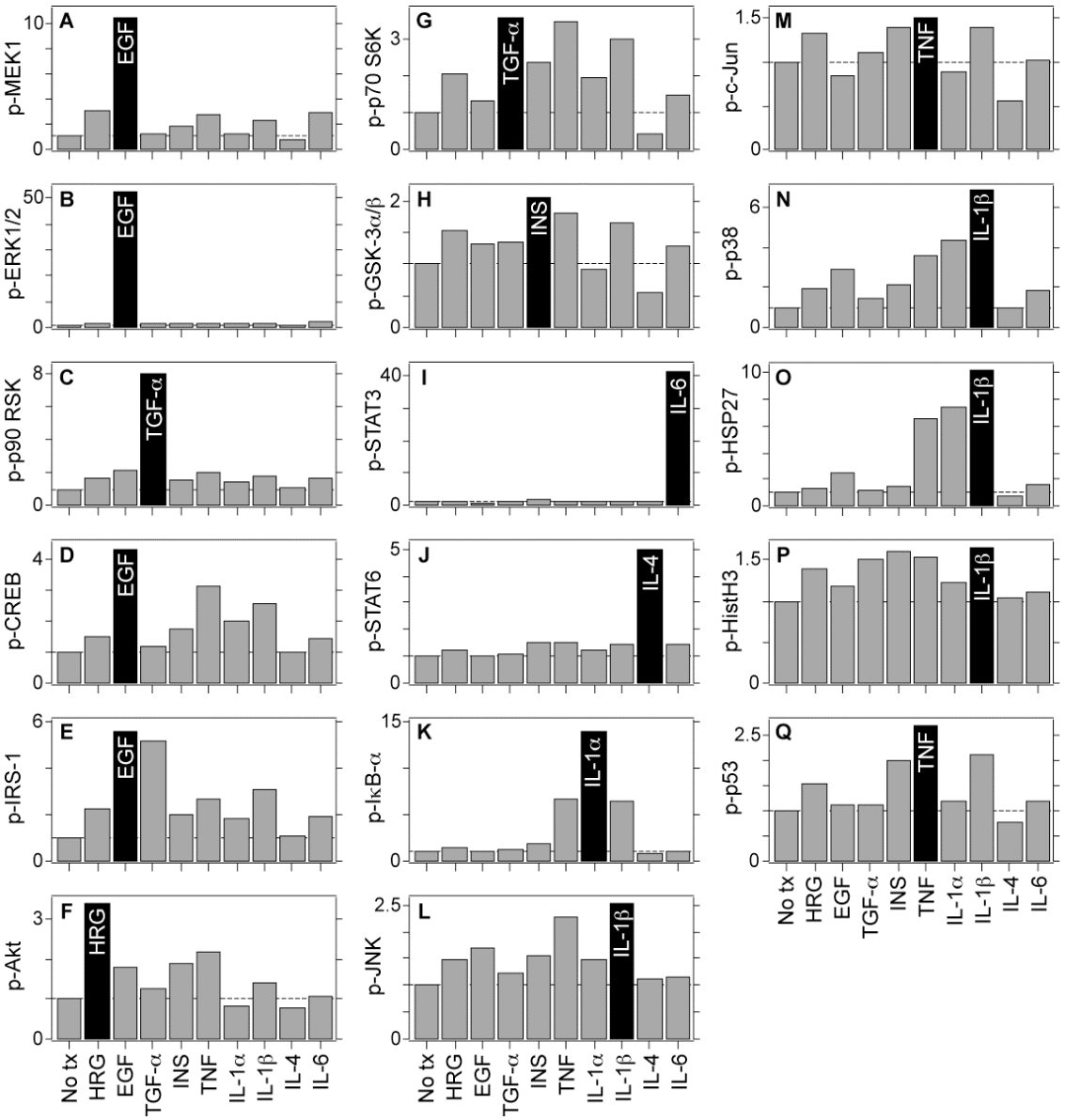

Figure S1. Specificity of multiplexed bead-based phosphoprotein assays. Human hepatocytes were cultured, treated, and lysed as described in the *Experimental* section, except insulin was removed from the culture medium 24 h before stimulation to allow for maximal activation of insulin-related pathways. Cells were stimulated for 20 min with one of the following treatments: no treatment, 100 ng ml^{-1} heregulin (HRG), 100 ng ml^{-1} epidermal growth factor (EGF), 100 ng ml^{-1} transforming growth factor-α (TGF-α), 2 μM insulin (INS), 100 ng ml^{-1} tumor necrosis factor-α (TNF), 20 ng ml^{-1} interleukin-1α (IL-1α), 20 ng ml^{-1} interleukin-1β (IL-1β), 50 ng ml^{-1} interleukin-4 (IL-4), or 100 ng ml^{-1} interleukin-6 (IL-6). (**A**-**Q**) Lysates were analyzed using 17 multiplexed phosphoprotein assays, and assay values are reported as fold-changed normalized to the untreated samples separately for each assay. Dashed lines for the untreated sample values are shown for clarity. Black bars are used for the stimulation leading to the highest signal activation for each phosphoprotein. These data show that the 17 phosphoprotein assays demonstrate high specificity with regards to stimulation condition.

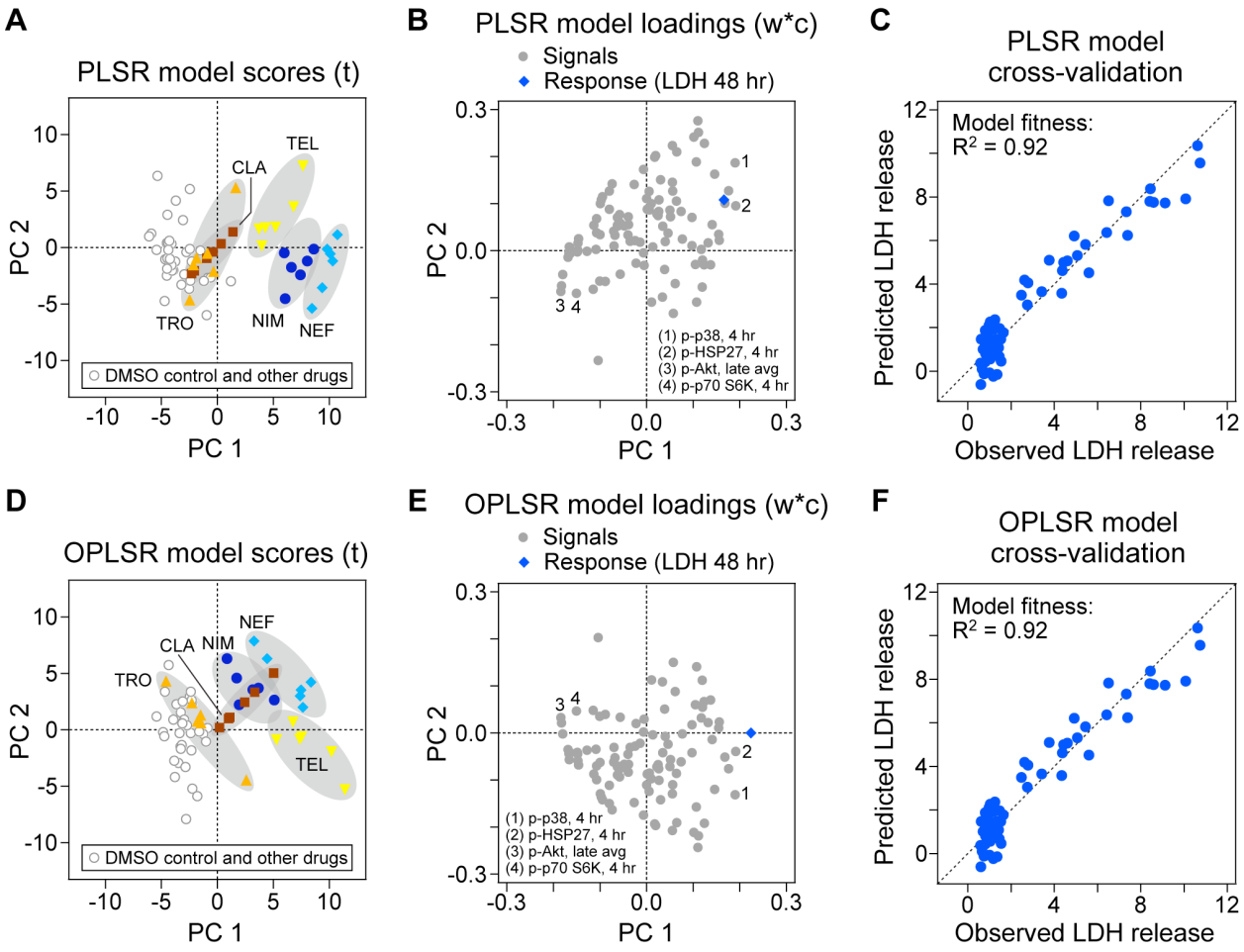

Figure S2. Comparison of partial least-squares regression (PLSR) and orthogonal PLSR models. A PLSR model was generated from the training CSR data compendium using the NIPALS algorithm in SIMCA-P software following standard methods.[6, 8-11] The PLSR model was generated using four principal components under standard optimization criteria[8], and its model scores (**A**), loadings (**B**), and cross-validated predictions (**C**) are plotted. The PLSR model was then subjected to a principal component-space linear transformation by rotating the projection (in 4-dimension principal component-space, with only the first two PC's plotted in (**A**) and (**B**)) of the single cell death response variable completely into the first principal component, with the model scores and loadings of the 102 phosphoprotein signaling metrics similarly rotated. This orthogonal PLSR (OPLSR) model[14] allows for simplified interpretation of model scores (**D**) and loadings (**E**). The OPLSR model demonstrated equivalent cross-validated model predictions and model fitness (**F**) as the original PLSR model (**C**).

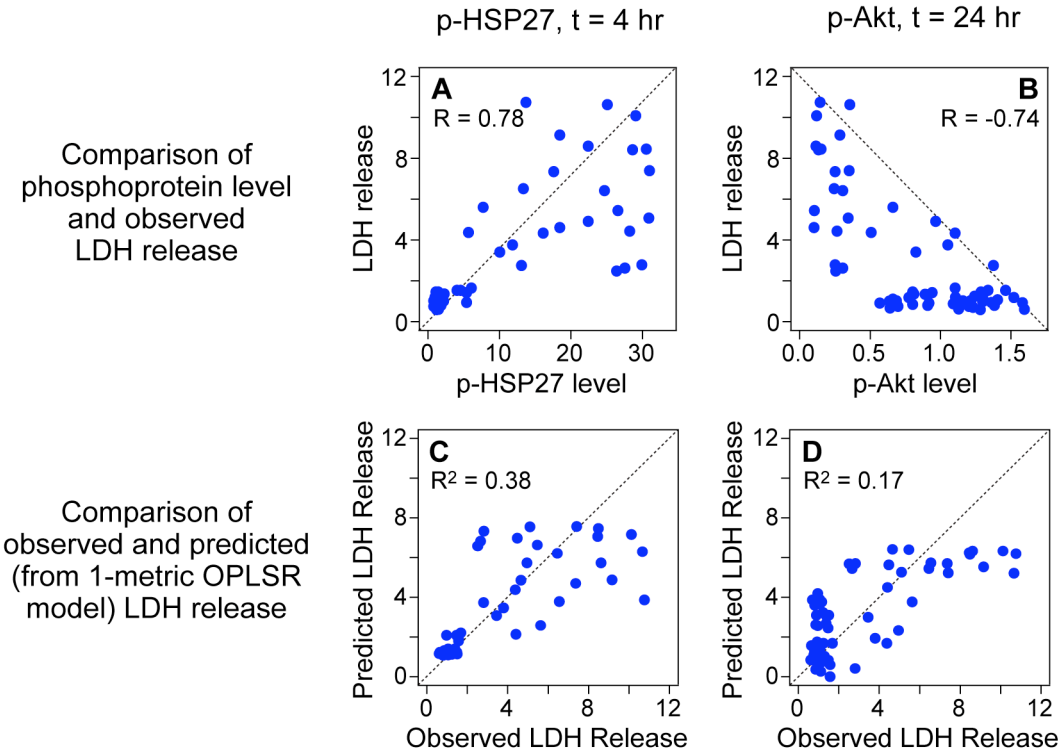

Figure S3. The most correlative signaling metrics are poorly predictive of the observed hepatic cytotoxicity response. (**A**-**B**) In the CSR compendium from donor #1, the single signaling metrics most positively (p-HSP27 at t = 4 h; **A**) and negatively (p-Akt at t = 24 h; **B**) correlated with the observed cell death response (LDH release at t = 48 h) are plotted for all 66 drug-cytokine treatment conditions. Phosphoprotein signal levels are reported as fold-change normalized values to untreated samples at t = 0. LDH release levels are reported as fold-changed normalized values to untreated samples at t = 48 h. One-to-one correlation lines shown for clarity. Pearson correlation coefficients (R) are shown in the insets. (**C**-**D**) OPLSR models trained using these single signaling metrics provide poor model fitness to the observed LDH release data. OPLSR models were trained on either the p-HSP27 at t = 4 h (**C)** or p-Akt at t = 24 h (**D**) phosphoprotein signaling metrics (across all 66 conditions in donor #1). Model cross-validated predictions of LDH release are plotted. Model fitness[7, 11] was assessed (R^2) as described in the supplementary *Experimental* section. Both one-metric OPLSR models show poor predictive accuracy ($R^2 < 0.4$). A one-to-one correlation line demonstrating perfect model fitness ($R^2 = 1$) is shown for clarity.

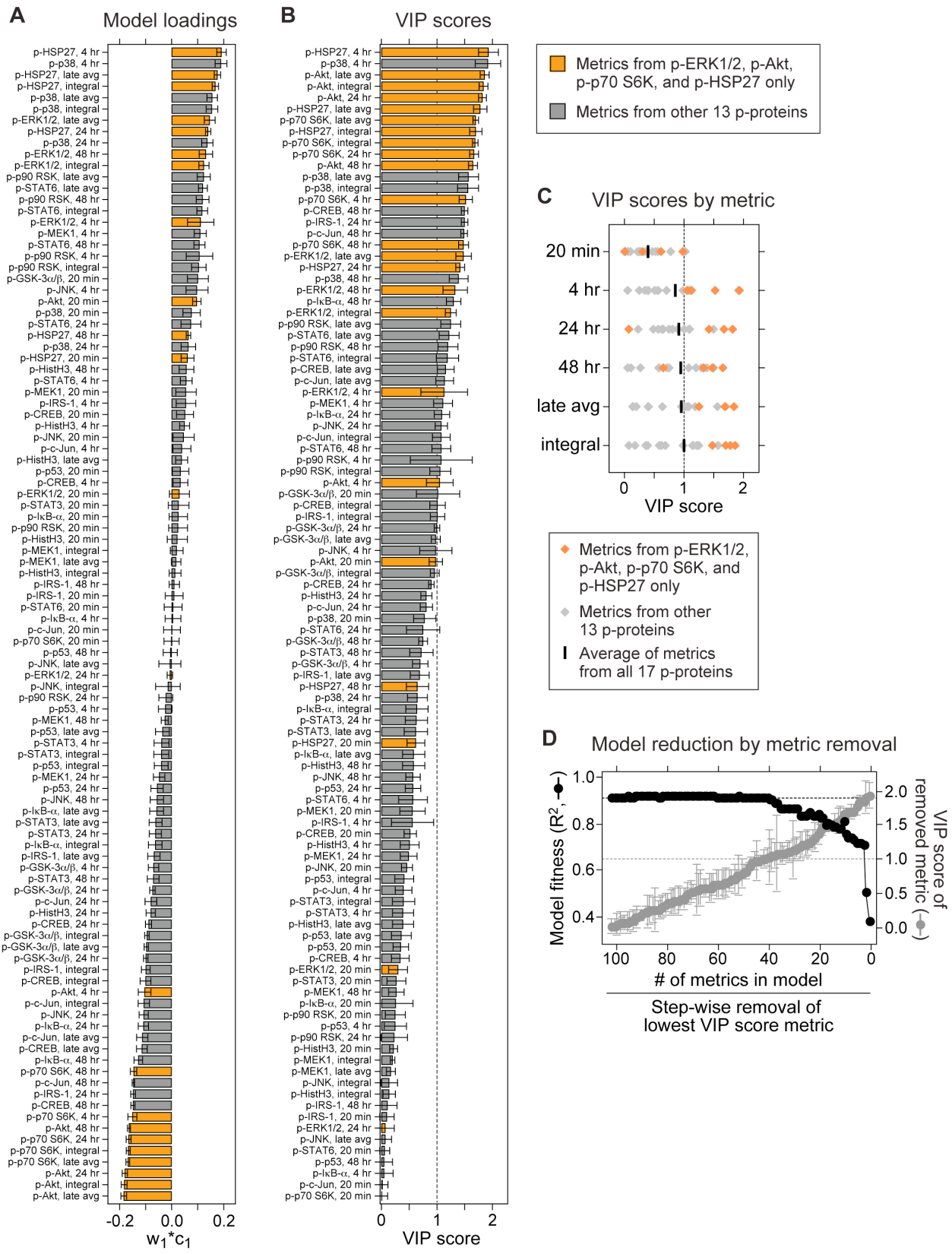
A
Model loadings
p-HSP27, 4 hr
p-p38, 4 hr
p-HSP27, late avg
p-HSP27, integral
p-p38, late avg
p-p38, integral
p-ERK1/2, late avg
p-HSP27, 24 hr
p-p38, 24 hr
p-ERK1/2, 48 hr
p-ERK1/2, integral
p-p90 RSK, late avg
p-STAT6, late avg
p-p90 RSK, 48 hr
p-STAT6, integral
p-ERK1/2, 4 hr
p-MEK1, 4 hr
p-STAT6, 48 hr
p-p90 RSK, 4 hr
p-p90 RSK, integral
p-GSK-3α/β, 20 min
p-JNK, 4 hr
p-Akt, 20 min
p-p38, 20 min
p-STAT6, 24 hr
p-HSP27, 48 hr
p-p38, 24 hr
p-HSP27, 20 min
p-HistH3, 48 hr
p-STAT6, 4 hr
p-MEK1, 20 min
p-IRS-1, 4 hr
p-CREB, 20 min
p-HistH3, 4 hr
p-JNK, 20 min
p-c-Jun, 4 hr
p-HistH3, late avg
p-p53, 20 min
p-CREB, 4 hr
p-ERK1/2, 20 min
p-STAT3, 20 min
p-IκB-α, 20 min
p-p90 RSK, 20 min
p-HistH3, 20 min
p-MEK1, integral
p-MEK1, late avg
p-HistH3, integral
p-IRS-1, 48 hr
p-IRS-1, 20 min
p-STAT6, 20 min
p-IκB-α, 4 hr
p-c-Jun, 20 min
p-p70 S6K, 20 min
p-p53, 48 hr
p-JNK, late avg
p-ERK1/2, 24 hr
p-JNK, integral
p-p90 RSK, 24 hr
p-p53, 4 hr
p-MEK1, 48 hr
p-p53, late avg
p-STAT3, 4 hr
p-STAT3, integral
p-p53, integral
p-MEK1, 24 hr
p-p53, 24 hr
p-JNK, 48 hr
p-IκB-α, late avg
p-STAT3, late avg
p-STAT3, 24 hr
p-IκB-α, integral
p-IRS-1, late avg
p-GSK-3α/β, 4 hr
p-STAT3, 48 hr
p-GSK-3α/β, 24 hr
p-c-Jun, 24 hr
p-HistH3, 24 hr
p-CREB, 24 hr
p-GSK-3α/β, integral
p-GSK-3α/β, late avg
p-GSK-3α/β, 24 hr
p-IRS-1, integral
p-CREB, integral
p-Akt, 4 hr
p-c-Jun, integral
p-JNK, 24 hr
p-IκB-α, 24 hr
p-c-Jun, late avg
p-CREB, late avg
p-IκB-α, 48 hr
p-p70 S6K, 48 hr
p-c-Jun, 48 hr
p-IRS-1, 24 hr
p-CREB, 48 hr
p-p70 S6K, 4 hr
p-Akt, 48 hr
p-p70 S6K, 24 hr
p-p70 S6K, integral
p-p70 S6K, late avg
p-Akt, 24 hr
p-Akt, integral
p-Akt, late avg
-0.2
0.0
0.2
w1*c1
B
VIP scores
p-HSP27, 4 hr
p-p38, 4 hr
p-Akt, late avg
p-Akt, integral
p-Akt, 24 hr
p-HSP27, late avg
p-p70 S6K, late avg
p-HSP27, integral
p-p70 S6K, integral
p-p70 S6K, 24 hr
p-Akt, 48 hr
p-p38, late avg
p-p38, integral
p-p70 S6K, 4 hr
p-CREB, 48 hr
p-IRS-1, 24 hr
p-c-Jun, 48 hr
p-p70 S6K, 48 hr
p-ERK1/2, late avg
p-HSP27, 24 hr
p-p38, 48 hr
p-ERK1/2, 48 hr
p-IκB-α, 48 hr
p-ERK1/2, integral
p-p90 RSK, late avg
p-STAT6, late avg
p-p90 RSK, 48 hr
p-STAT6, integral
p-CREB, late avg
p-c-Jun, late avg
p-ERK1/2, 4 hr
p-MEK1, 4 hr
p-IκB-α, 24 hr
p-JNK, 24 hr
p-c-Jun, integral
p-STAT6, 48 hr
p-p90 RSK, 4 hr
p-p90 RSK, integral
p-Akt, 4 hr
p-GSK-3α/β, 20 min
p-CREB, integral
p-IRS-1, integral
p-GSK-3α/β, 24 hr
p-GSK-3α/β, late avg
p-JNK, 4 hr
p-Akt, 20 min
p-GSK-3α/β, integral
p-CREB, 24 hr
p-HistH3, 24 hr
p-c-Jun, 24 hr
p-p38, 20 min
p-STAT6, 24 hr
p-GSK-3α/β, 48 hr
p-STAT3, 48 hr
p-GSK-3α/β, 4 hr
p-IRS-1, late avg
p-HSP27, 48 hr
p-p38, 24 hr
p-IκB-α, integral
p-STAT3, 24 hr
p-STAT3, late avg
p-HSP27, 20 min
p-IκB-α, late avg
p-HistH3, 48 hr
p-JNK, 48 hr
p-p53, 24 hr
p-STAT6, 4 hr
p-MEK1, 20 min
p-IRS-1, 4 hr
p-CREB, 20 min
p-HistH3, 4 hr
p-MEK1, 24 hr
p-JNK, 20 min
p-p53, integral
p-c-Jun, 4 hr
p-STAT3, integral
p-STAT3, 4 hr
p-HistH3, late avg
p-p53, late avg
p-p53, 20 min
p-CREB, 4 hr
p-ERK1/2, 20 min
p-STAT3, 20 min
p-MEK1, 48 hr
p-IκB-α, 20 min
p-p90 RSK, 20 min
p-p53, 4 hr
p-p90 RSK, 24 hr
p-HistH3, 20 min
p-MEK1, integral
p-MEK1, late avg
p-JNK, integral
p-HistH3, integral
p-IRS-1, 48 hr
p-IRS-1, 20 min
p-ERK1/2, 24 hr
p-JNK, late avg
p-STAT6, 20 min
p-p53, 48 hr
p-IκB-α, 4 hr
p-c-Jun, 20 min
p-p70 S6K, 20 min
0
1
2
VIP score
Metrics from p-ERK1/2, p-Akt, p-p70 S6K, and p-HSP27 only
Metrics from other 13 p-proteins
C
VIP scores by metric
20 min
4 hr
24 hr
48 hr
late avg
integral
0
1
2
VIP score
Metrics from p-ERK1/2, p-Akt, p-p70 S6K, and p-HSP27 only
Metrics from other 13 p-proteins
Average of metrics from all 17 p-proteins
D
Model reduction by metric removal
Model fitness (R2, ●)
1.0
0.8
0.6
0.4
VIP score of removed metric (–)
2.0
1.5
1.0
0.5
0.0
100
80
60
40
20
0
# of metrics in model
Step-wise removal of lowest VIP score metric

Figure S4. (Previous page) OPLSR model loadings and VIP scores. (**A**-**B**) Model loadings (**A**) and VIP scores (**B**) are plotted for all 102 phosphoprotein signaling metrics, with metrics from p-ERK1/2, p-Akt, p-p70 S6K, and p-HSP27 noted, sorted by loading and VIP score values, respectively. Model loadings (**A**) and VIP scores (**A**,**D**) are presented as the mean values ± cross-validation standard error, calculated by jack-knifing.[13] (**C**) VIP scores grouped by signaling metric, with metrics from p-ERK1/2, p-Akt, p-p70 S6K, and p-HSP27 noted. (**D**) Model fitness sensitivity to the removal of individual phosphoprotein signaling metrics. Model complexity was reduced by step-wise removal of the signaling metric with the lowest VIP score. Model fitness[7, 11] (R^2, black/left axis) and the VIP score of the removed metric (grey/right axis) are plotted for each reduced model. In (**B**-**D**), a line (grey color in **D**) indicating the threshold value of 1 for informative VIP scores[7, 15] is shown for clarity. In (**D**), a line (black) indicating the model fitness (R^2 = 0.92) of the full 102-metric OPLSR is shown for clarity.

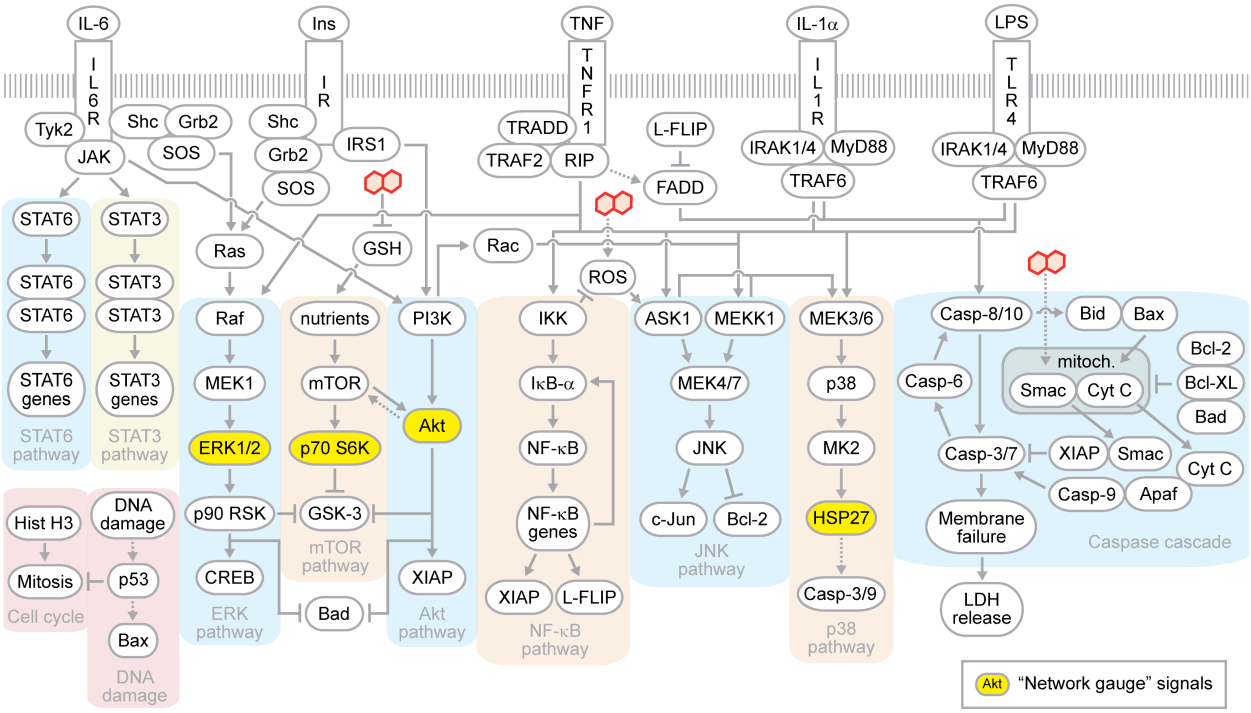

Figure S5. A schematic of the drug- and cytokine-induced hepatocellular death signaling network (see Fig. 1) highlighting the four 'network gauge' phosphoprotein signals (p-ERK1/2, p-p70 S6K, p-Akt, and p-HSP27) identified by VIP score-based reduction of the full 17-phosphoprotein OPLSR model (see Fig. 3D).

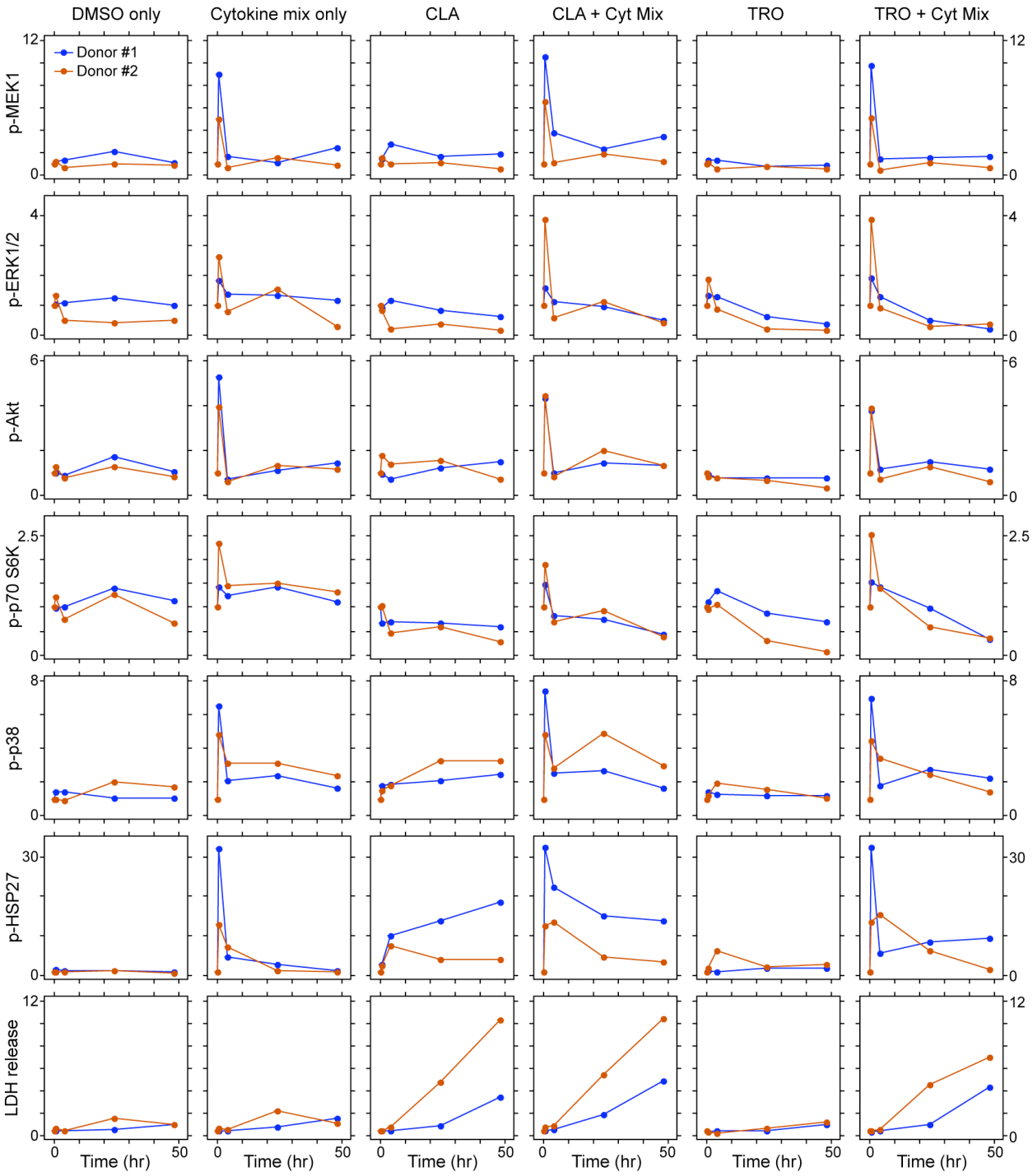

Figure S6. Comparison of phosphoprotein and LDH release data in human hepatocyte donors #1 and #2. Phosphoprotein and LDH release data were collected and fold-changed normalized (separately for each donor/assay) from cultured human hepatocytes from donors #1 and #2 as described in Fig. 2 and Fig. 4. Mean values from assays (p-MEK1, p-ERK1/2, p-Akt, p-p70 S6K, p-38, p-HSP27, and LDH release) and treatments (DMSO only, the three-cytokine/LPS mix only, 334 μM clarithromycin [CLA], CLA + cytokine mix, 770 μM trovafloxacin [TRO], TRO + cytokine mix) conducted in both donors are presented, with full data sets shown in Fig. 2 (donor #1, n = 1 biological replicate) and Fig. 4 (donor #2, n = 3 biological replicates).

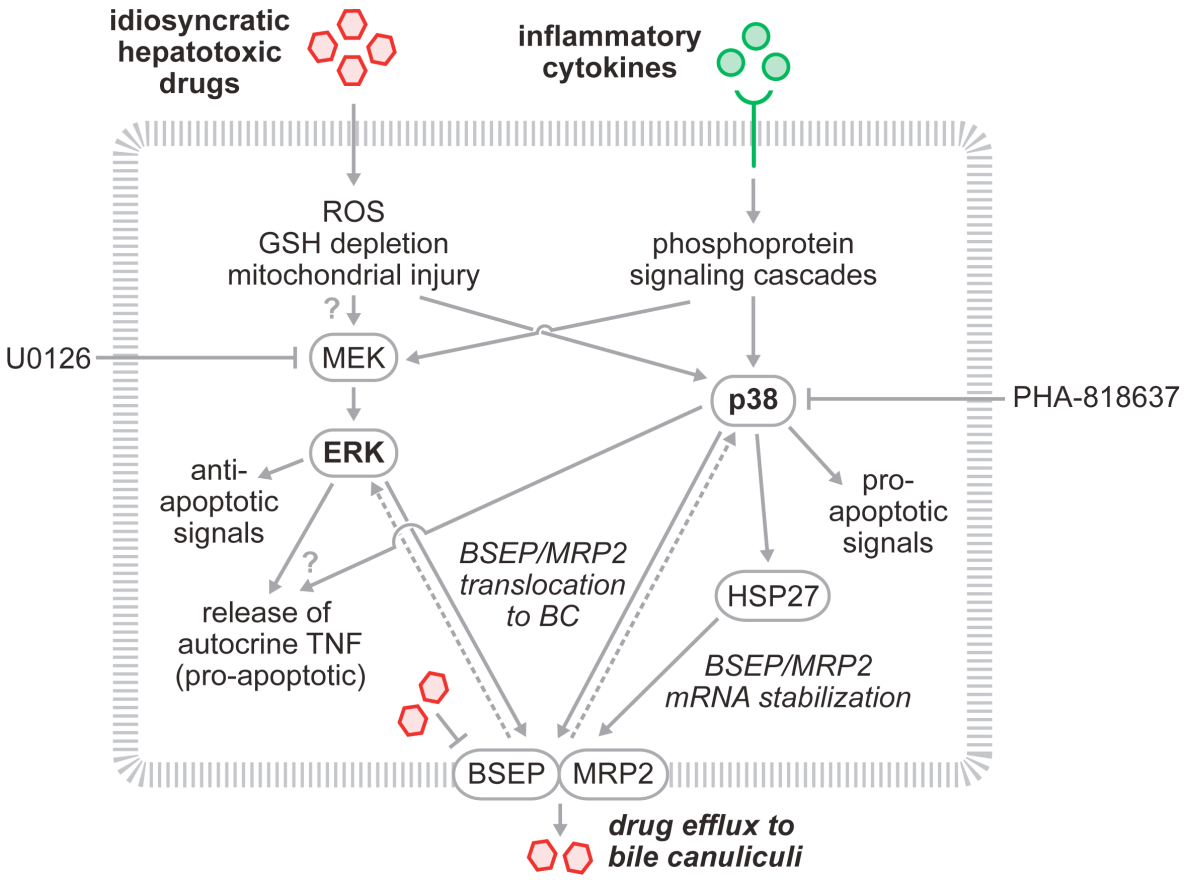

Figure S7. ERK and p38 pathways at the nexus of drug/cytokine-induced signaling and drug efflux transporter regulation. Interpretations from the OPLSR model of hepatic cytotoxicity suggest that the ERK and p38 pathways are activated by drug- and/or cytokine treatments and positively regulate the resulting cell death response. Whereas the ERK pathway is generally considered pro-survival through its activation of anti-apoptotic effectors such as Bad[17, 18], the p38 pathway is generally considered pro-apoptotic due to its transcription regulation of effector caspases[19]. Both pathways have been implicated in the regulation of the translocation of hepatocyte drug efflux transporters, including the bile salt export pump (BSEP) and the conjugate export pump (MRP2), to the bile canaliculi (BC). Consequently, inhibitors of the kinase activities of MEK[20] and p38[21, 22] decrease drug efflux transporter translocation and activity, leading to cholestasis upon prolonged administration. Possibly due to their inhibition of drug efflux transporter activities or, in the case of ERK, due to their perturbation of apoptosis regulatory mechanisms, MEK[23, 24] and p38[25, 26] inhibitors elicit liver toxicity some cellular and animals models and clinical investigations. Further, some idiosyncratic hepatotoxicants (*e.g.* nefazodone[27]) inhibit BSEP and/or MRP2 activities. This inhibition itself can induce transient activation of ERK and p38 signaling and consequently stimulate additional transporter protein translocation to the BC to enable recovery of drug efflux capacity. Additionally, activated ERK[28] and p38[29] can both positively regulate the metalloproteinase TACE and its proteolytic release of TNF.[30] TNF is a crucial pro-death cytokine in inflammation-induced drug hepatotoxicity *in vivo*[31, 32], where its TACE-regulated release by neutrophils is p38-dependent[29], and in our *in vitro* drug/cytokine co-treatment system.[1] Thus, the experimental perturbation of the ERK and p38 pathways must be made with careful consideration to their complex, and possibly counter-acting, functions in response to drug and/or cytokine stimuli.

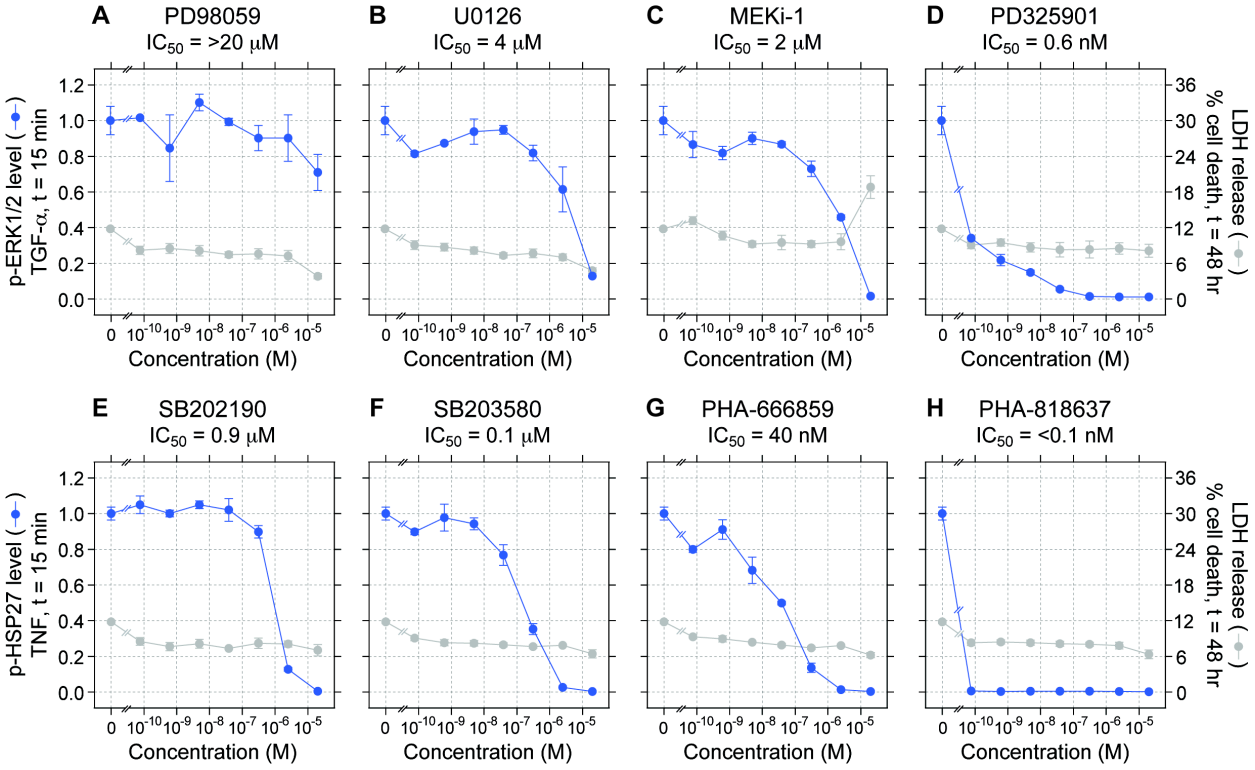

Figure S8. Selection of MEK and p38 kinase inhibitors based on potent signaling inhibition efficacy and minimal toxicity in human hepatocytes. MEK (**A**-**D**) and p38 (**E**-**H**) kinase inhibitors were evaluated for efficacy and toxicity in human hepatocytes (from donor #3) over a range of concentrations at seven 8× serial dilution concentrations from 20 μM (20 μM, 2.5 μM, 0.31 μM, 39 nM, 4.9 nM, 0.61 nM, 76 pM). Data are presented as the mean ± s.e.m. of three biological replicates. (**A**-**H**) To evaluate MEK and p38 kinase inhibitor toxicity, inhibitors were added at 1× final concentrations in fresh medium for 48 h, then medium samples were assayed for LDH release. LDH results were normalized to wells from the same culture plate lysed in 1% Triton X for 10 min, and values were reported as % cell death, with the lysed samples assumed to represent 100% cell death. (Note that only 20 μM MEKi-1 elicited significant toxicity.) (**A**-**D**) To evaluate MEK kinase inhibitor efficacy, human hepatocytes were pretreated with inhibitor for 1 h before treatment with 100 ng ml^{-1} TGF-α for 15 min and then were assayed for p-ERK1/2 activation. (**E**-**H**) To evaluate p38 kinase inhibitor efficacy, human hepatocytes were pretreated with inhibitor for 1 h before treatment with 100 ng ml^{-1} TNF for 15 min and then were assayed for p-HSP27 activation. (**A**-**H**) Signaling inhibition IC_{50} values were manually estimated from the signaling down-regulation curves by identifying the inhibitor concentrations that elicited half-maximal phosphoprotein activation. The following kinase inhibitors/ concentrations were selected for their potent signaling inhibition and minimal toxicity and were used to perturb kinase activities in drug- and cytokine co-treatment experiments (see Fig. 6 and Fig. S9): 10 μM U0126, 1 μM PD325901, 1 μM PHA-666859, and 1 nM PHA-818637. At these concentrations, these inhibitors yield the following reductions in phosphoprotein signaling: U0126, 70% reduction, and PD325901, 99%, for p-MEK1/p-ERK1/2 inhibition; and PHA-666859, 93%, and PHA-818637, 99%, for p-p38/p-HSP27 inhibition.

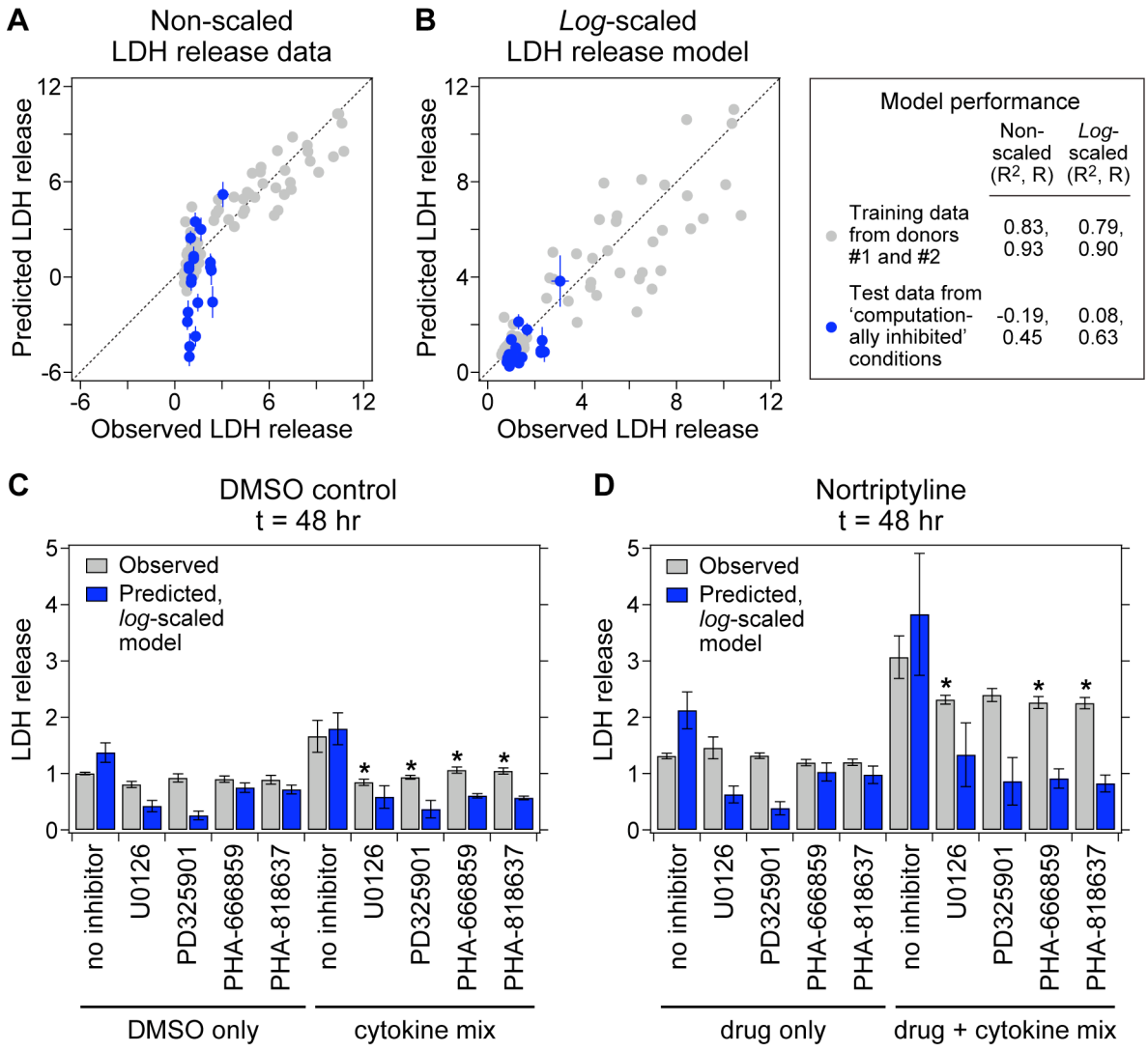

Figure S9. Six-phosphoprotein OPLSR model makes qualitatively accurate predictions of the effects of MEK and p38 inhibitors in perturbing drug- and cytokine-induced hepatic cytotoxicity. To make *a priori* predictions of a kinase inhibitor perturbation to a drug- and/or cytokine-induced hepatocellular death response, 'computationally inhibited' signaling time-courses (based on uninhibited signaling data from donors #1 and/or #2, limited to the six phosphoproteins assayed in samples from both donors) were generated by reducing the activation time-courses of the specific phosphoprotein signaling molecules targeted by a kinase inhibitor by an amount based on the inhibitor's experimentally measurement signaling inhibition (in samples from donor #3; see Fig. S8). Signaling proteins from pathways not targeted by the kinase inhibitor were left unchanged. For example, p-MEK1 and p-ERK1/2 activation time-courses were decreased by 70% for the MEK inhibitor U0126, but the p-Akt, p-p70 S6K, p-p38, and p-HSP27 were unchanged. These computationally inhibited time courses were generated for the treatment conditions of DMSO ± cytokine mix (20 ng ml^{-1} IL-1α, 10 μg ml^{-1} LPS, 100 ng ml^{-1} TNF, and 20 ng ml^{-1} IL-6) and nortriptyline ± cytokine mix, in the presence or absence of 10 μM U0126, 1 μM PD325901, 1 μM PHA-666859, and 1 nM PHA-818637. After computationally inhibiting the time-course, the integral and late average metrics were re-calculated. The resultant computationally inhibited signaling metric data were used to predict cell death responses from OPLSR models trained on a fused (uninhibited) signaling data compendium from donors #1 and

#2. Model fitness of the predicted inhibitor responses was assessed by comparing to experimental measurements collected in human hepatocytes from donor #4. Human hepatocytes (from donor #4) were treated with DMSO control ± cytokine mix (**C**) or 10 μM nortriptyline ± cytokine mix (**D**), in the absence or presence of a kinase inhibitor. To inhibit MEK kinase activity, cells were pretreated with 10 μM U0126 or 1 μM PD325901 1 h before drug/cytokine stimulation. To inhibit p38 kinase activity, cells were pretreated with 1 μM PHA-666859 or 1 nM PHA-818637 1 h before drug/cytokine stimulation. After 24 or 48 h, conditioned medium samples were assayed for LDH release. Two different 6-phosphoprotein OPLSR models were trained on the fused CSR data compendium from both donors #1 and #2. (**A**) One model was trained using non-scaled cell death response data. Although this OPLSR model demonstrated good model fitness[7, 11] for the training data (R^2 = 0.83), it poorly predicted the inhibitor test data (R^2 = -0.19) and led to significantly under-predicted cell death responses. (**B**-**D**) A second OPLSR model was generated from the fused training compendium by regressing *log*-scaled cell death response data. The *log*-scaled model demonstrated reasonable model fitness for the training data (R^2 = 0.79) but only qualitatively accurate predictions of the test computationally inhibited data (R^2 = 0.08). Using a less stringent Pearson correlation metric, the *log*-scaled model yielded predictions of the training (R = 0.90) and test inhibition (R = 0.63) data sets that were well-correlated with the observed responses. In (**A**) to (**D**), test experimental data are presented as mean ± s.e.m. of eight biological replicates. In (**A**) and (**B**), model-predicted responses are presented as the mean prediction ± cross-validation standard error, calculated by jack-knifing[13], and experimental and prediction uncertainties are not shown for the training data for simplicity. In (**A**) and (**B**), a one-to-one correlation line demonstrating perfect model fitness (R^2 = 1) is shown for clarity. In (**C**) and (**D**), for experimental data, differences between uninhibited and kinase inhibitor pretreatments are labeled as significant (*) if $P < 0.05$ by a Student's *t* test.

Table S1. Human hepatocyte donor information.

Donor ID	Experiment	Data in	Culture start date	Lot**	Donor information and history*			
					Gender, age	Obesity†	Smoking, drug, alcohol use	Disease, viral infection
1	CSR training data set	Fig. 2	05/20/07	Hu4000	Female, 4 yrs old	Not obese, BMI unknown	None known	None known††
2	CSR test data set	Fig. 4	10/18/08	Hu0921	Male, 38 yrs old	Not obese, BMI 25	None known	None known
3	MEK and p38 inhibitor dosing study	Fig. S8	10/17/08	Hu0920	Female, 52 yrs old	Not obese, BMI 21	None known	None known
4	Kinase inhibitor drug-cytokine toxicity study	Fig. 6 Fig. S9	11/13/08	Hu0935	Female, 54 yrs old	Not obese, BMI 20	None known	None known

* All donor information provided by CellzDirect, Inc.
** Lot number assigned by CellzDirect, Inc.
† BMI = body mass index. A BMI of 30 or greater is generally considered obese.
†† No known history of or exposure to Hepatitis B, Hepatitis C, cirrhosis, biliary disease or HIV. Not serology-tested.

Supplementary References

1. B. D. Cosgrove, B. M. King, M. A. Hasan, L. G. Alexopoulos, P. A. Farazi, B. S. Hendriks, L. G. Griffith, P. K. Sorger, B. Tidor, J. J. Xu and D. A. Lauffenburger, *Toxicol Appl Pharmacol*, 2009, **237**, 317-330.

2. T. S. Peters, *Toxicol Pathol*, 2005, **33**, 146-154.

3. K. D. Clay, J. S. Hanson, S. D. Pope, R. W. Rissmiller, P. P. Purdum, 3rd and P. M. Banks, *Ann Intern Med*, 2006, **144**, 415-420.

4. J. J. Xu, P. V. Henstock, M. C. Dunn, A. R. Smith, J. R. Chabot and D. de Graaf, *Toxicol. Sci.*, 2008, **105**, 97-105.

5. J. Saez-Rodriguez, A. Goldsipe, J. Muhlich, L. G. Alexopoulos, B. Millard, D. A. Lauffenburger and P. K. Sorger, *Bioinformatics*, 2008, **24**, 840-847.

6. K. A. Janes, J. G. Albeck, S. Gaudet, P. K. Sorger, D. A. Lauffenburger and M. B. Yaffe, *Science*, 2005, **310**, 1646-1653.

7. S. Gaudet, K. A. Janes, J. G. Albeck, E. A. Pace, D. A. Lauffenburger and P. K. Sorger, *Mol Cell Proteomics*, 2005, **4**, 1569-1590.

8. P. Geladi and B. Kowalski, *Anal. Chim. Acta*, 1986, **185**, 1-17.

9. N. Kumar, R. Afeyan, H. D. Kim and D. A. Lauffenburger, *Mol Pharmacol*, 2008, **73**, 1668-1678.

10. K. Miller-Jensen, K. A. Janes, J. S. Brugge and D. A. Lauffenburger, *Nature*, 2007, **448**, 604-608.

11. K. A. Janes, H. C. Reinhardt and M. B. Yaffe, *Cell*, 2008, **135**, 343-354.

12. N. Kumar, A. Wolf-Yadlin, F. M. White and D. A. Lauffenburger, *PLoS Comput Biol*, 2007, **3**, e4.

13. B. Efron and R. J. Tibshirani, *An introduction to the bootstrap*, 1993, Chapman and Hall, London.

14. M. Bylesjö, M. Rantalainen, O. Cloarec, J. K. Nicholson, E. Holmes and J. Trygg, *J Chemometrics*, 2006, **20**, 341-351.

15. S. Wold, *Chemometr Intell Lab*, 1994, **23**, 149-161.

16. M. L. Kemp, L. Wille, C. L. Lewis, L. B. Nicholson and D. A. Lauffenburger, *J Immunol*, 2007, **178**, 4984-4992.

17. J. A. McCubrey, M. Milella, A. Tafuri, A. M. Martelli, P. Lunghi, A. Bonati, M. Cervello, J. T. Lee and L. S. Steelman, *Curr Opin Investig Drugs*, 2008, **9**, 614-630.

18. B. A. Ballif and J. Blenis, *Cell Growth Differ*, 2001, **12**, 397-408.

19. K. Ono and J. Han, *Cell Signal*, 2000, **12**, 1-13.

20. F. Schliess, A. K. Kurz, S. vom Dahl and D. Haussinger, *Gastroenterology*, 1997, **113**, 1306-1314.

21. R. Kubitz, G. Sutfels, T. Kuhlkamp, R. Kolling and D. Haussinger, *Gastroenterology*, 2004, **126**, 541-553.

22. A. K. Kurz, D. Graf, M. Schmitt, S. Vom Dahl and D. Haussinger, *Gastroenterology*, 2001, **121**, 407-419.

23. S. C. Wentz, H. Wu, M. T. Yip-Schneider, M. Hennig, P. J. Klein, J. Sebolt-Leopold and C. M. Schmidt, *J Gastrointest Surg*, 2008, **12**, 30-37.

24. A. P. Brown, T. C. Carlson, C. M. Loi and M. J. Graziano, *Cancer Chemother Pharmacol*, 2007, **59**, 671-679.

25. J. J. Xu, B. S. Hendriks, J. Zhao and D. de Graaf, *FEBS Lett*, 2008, **582**, 1276-1282.

26. D. M. Dambach, *Curr Top Med Chem*, 2005, **5**, 929-939.

27. S. E. Kostrubsky, S. C. Strom, A. S. Kalgutkar, S. Kulkarni, J. Atherton, R. Mireles, B. Feng, R. Kubik, J. Hanson, E. Urda and A. E. Mutlib, *Toxicol Sci*, 2006, **90**, 451-459.

28. M. Chokki, H. Eguchi, I. Hamamura, H. Mitsuhashi and T. Kamimura, *Febs J*, 2005, **272**, 6387-6399.

29. X. Deng, J. Lu, L. D. Lehman-McKeeman, E. Malle, D. L. Crandall, P. E. Ganey and R. A. Roth, *J Pharmacol Exp Ther*, 2008, **326**, 144-152.

30. G. T. Le and G. Abbenante, *Curr Med Chem*, 2005, **12**, 2963-2977.

31. P. J. Shaw, M. J. Hopfensperger, P. E. Ganey and R. A. Roth, *Toxicol Sci*, 2007, **100**, 259-266.

32. F. F. Tukov, J. P. Luyendyk, P. E. Ganey and R. A. Roth, *Toxicol Sci*, 2007, **100**, 267-280.