Supporting Information

Kravchenko-Balasha et al. 10.1073/pnas.1404462111

SI Text

Supporting Theory Surprisal Analysis

Theory. We apply the procedure of maximal entropy to describe a two-cell system under constraints.

Assumptions and applications:

1. The two-cell system is in a state of maximal entropy subject to constraints at every distance range. Our goal is to find the (distance dependent) constraints that prevent the entropy from reaching the global maximum. The global maximum of the entropy is the steady state, a state without any constraints on the biological networks. The constraints are detected and quantified by identifying how the protein expression levels respond to those constraints. For each measured protein, the extent of participation of the protein in the biological process responding to constraints is defined.
2. The experimentally determined mean values, X_i, of the functional proteins as a function of distance are used to identify the constraints.
3. Constraints prevent the entropy from reaching the global maximum, and thereby generate deviations from the steady state. We seek to define the protein concentrations X_i^o at the steady state—the state at the global maximum of the entropy—and compare it to the experimental protein levels.
4. We seek to define distance ranges with minimal deviations from the steady state; these are the most stable states of the system and thus are most probable.
5. To determine the distance ranges with minimal constraints we use surprisal analysis as presented in Eq. **1** in the main text.

Here we give a very brief summary how this equation was developed. For more details, see refs. 1–5. First, we relate protein concentrations to the chemical potential using the fundamental physical chemical relationships (6):

$$\begin{aligned} \underbrace{\mu_i}_{\text{free energy in a solution}} &= \underbrace{\mu_i^{ss}}_{\text{standard free energy}} + kT\ln X_i \\ \underbrace{\mu_i^o}_{\text{free energy at the steady state}} &= \underbrace{\mu_i^{ss}}_{\text{standard free energy}} + kT\ln X_i^o \end{aligned} \qquad [S1]$$

By solving the system of the above equations we obtain

$$\ln X_i = \ln X_i^o + \left(\mu_i - \mu_i^o\right)/kT. \qquad [S2]$$

From the experimental data X_i, we seek to obtain X_i^o, which is the functional protein expression expected at the steady state, as well as $(\mu_i - \mu_i^o)/kT$, which is the deviation from the steady state. Note from the equation that this deviation represents a change in a free energy of the system due to the constraint(s). As written above, we assume that at every distance range the biological system is in a state of maximal entropy subject to constraints. In addition to the calculation of the expected protein distributions at the maximal entropy X_i^o and deviations from X_i^o as a function of distance, we want to find biological networks responding to the constraints, and to calculate a weight λ_α for every existing constraint α.

To determine the change in the chemical potentials, we recall that they are Lagrange multipliers (6, 7) and so we use the technique of undetermined multipliers. For this technique, each multiplier is associated with a constraint on the system. A way to perform such a search (5) is to find the maximum entropy of the system subject to constraints:

$$£ = \text{entropy} - \sum_\alpha \lambda_\alpha \,\text{constraint}_\alpha \qquad [S3]$$

The sum over α is over all of the constraints. The numbers λ_α (weights of the constraints) are the Lagrange undetermined multipliers. Seeking the unconstrained maximum of £ is the solution to our problem of maximization of the entropy subject to constraints. To develop a working equation, the expression for the entropy of mixture of species in the cell was used (5). Constraints are presented as $\langle G_\alpha\rangle = \sum_i G_{i\alpha}X_i$. This equation means that the number X_i of protein molecules i is limited by the quantity $\langle G_\alpha\rangle$, which represents a biological network responding to the constraints α. Thus, α is the label of the constraint. $G_{i\alpha}$ is the extent of participation of protein i in the biological process responding to the constraint α (5).

The maximum of the function (Eq. **S3**) leads to the exponential form

$$X_i(r) = X_i^o \exp\left(-\sum_{\alpha=1} G_{i\alpha}\lambda_\alpha(r)\right). \qquad [S4]$$

The values $G_{i\alpha}$ represent the extent of participation of a protein i in a constraint α. The exponent of the above equation, $-\sum_{\alpha=1} G_{i\alpha}\lambda_\alpha(r)$, describes deviation from the steady state. Thereby we have our key result $\sum_{\alpha=1} G_{i\alpha}\lambda_\alpha(r) = (\mu_i - \mu_i^o)/kT$. By taking a natural logarithm in Eq. **S4** we obtain the working equation (Eq. **1** in the main text):

$$\ln X_i(r) = \ln X_i^o(r) - \sum_{\alpha=1} G_{i\alpha}\lambda_\alpha(r) \qquad [S5]$$

To make the steady-state term mimic the look of a deviation term, it is convenient to write $\ln X_i^o(r) = G_{i0}\lambda_0(r)$ (5). In this way, Eq. **1** can be written as $\ln X_i(r) = \sum_{\alpha=0} G_{i\alpha}\lambda_\alpha(r)$.

Because the cells in the chamber are not free to move, we impose the constraints separately at each range of cell–cell distance r. Thereby there is a set of Lagrange multipliers for each range of cell–cell distances. A given constraint α can have a large weight (λ_α) at some specific range of intercellular separation distances, but can have effectively a negligible weight at other distances (see Fig. S2*A* for more details). We look for those distance ranges that have very similar values of $X_i(r)$ and $X_i^o(r)$, meaning that $\sum_{\alpha=1} G_{i\alpha}\lambda_\alpha(r) \approx 0$.

Calculation of the parameters λ_α and $G_{i\alpha}$. We want to fit the sum of the terms as shown on the right-hand side of Eq. **S5** (Eq. **1** in the main text) to the logarithm of the measured expression level of protein i at the distance r, and this has to be repeated for every distance. We use the singular value decomposition (SVD) (ref. 5 and more below) method for diagonalizing a matrix that includes mean values of functional protein distributions from two separated cells (Table S2, rows) at a certain distance range (Table S2, columns). Table S2 contains the mean values of six functional protein distributions at six distance ranges. Using this data we determine two sets of parameters: the Lagrange multipliers λ_α for all α's at each distance, and the time-independent vector $G_{i\alpha}$ for each of the i proteins.

SVD Analysis. For the SVD we generate a matrix **Y** from the entries of Table S2 using the natural log of the average functional protein

levels (5), as indicated in the main text. Using SVD, we construct two square (and symmetric) matrices; the first one is $\mathbf{Y}^T\mathbf{Y}$. To determine constraints we use $\mathbf{Y}^T\mathbf{Y}\ \boldsymbol{\lambda}_\alpha = \omega_\alpha^2\ \boldsymbol{\lambda}_\alpha, \alpha = 0, 1, 2, .., \text{A-1}$. $\boldsymbol{\lambda}_\alpha$ is an eigenvector and ω_α^2 is an eigenvalue of the matrix $\mathbf{Y}^T\mathbf{Y}$. The components of the vector $\boldsymbol{\lambda}_\alpha$ provide the weights of the constraint α at each of the distance ranges r. The maximal number of nonzero eigenvalues $\boldsymbol{\lambda}_\alpha$ is A, where A is the smaller of the dimensions of the matrix $\mathbf{Y}$. The mathematically exact statement is that A is the rank of the matrix $\mathbf{Y}$. Here A, the maximal number of possible constraints, is equal to the number of binned distances [six for U87 EGF receptor variant III (EGFRvIII) cells]. However, as discussed in the main text, only one or two of those constraints have nonzero eigenvalues and thus play significant roles.

To determine the $G_{\alpha i}$, we generate a second matrix: $\mathbf{Y}\mathbf{Y}^T$ (5). This conjugated matrix has the same value of eigenvalues ω_α^2: $\mathbf{Y}\mathbf{Y}^T\mathbf{G}_\alpha = \omega_\alpha^2\ \mathbf{G}_\alpha, \alpha = 0, 1, 2, .., \text{A-1}$. $\mathbf{G}_\alpha$ is a column vector of six components, each component corresponding to a measured protein.

Output of the Surprisal Analysis. As shown in Table S3, $\alpha = 0$ has the biggest weight. This is the steady state. Constraints 3–5 are within the noise value as indicated by their low values. As written above, we use experimental protein levels $X_i(r)$ to calculate the steady-state expression levels and deviations from the steady state using $\ln X_i(r) = \sum_{\alpha=0} G_{i\alpha}\lambda_\alpha(r)$. From Table S3 and Fig. 2*B* one can see that the steady-state levels yield the largest contribution to the observed expression level of each protein at each distance range. The constraint $\alpha = 1$, which is the most dominant constraint, changes sign ~80–100 μm (Table S3). Thus, between 80 and 100 μm, the two-cell system hardly deviates from the steady state according to the constraint $\alpha = 1$. $\lambda_\alpha(r)$ values for U87PTEN cells are provided in Table S4. For the input we used a natural logarithm of the protein expression levels as a function of distance.

***DPE* Function.** We define *DPE* function (deviation in free energy per protein molecule), which is calculated using the equation below (Eq. **2** in the main text):

$$DPE(r) = \left\{ \sum_i X_i(r)\left[\ln\left(X_i(r)/X_i^o(r)\right)\right] - \sum_i \left(X_i(r) - X_i^o(r)\right) \right\} \Big/ \left\{ \sum_i X_i(r) \right\}$$

DPE is calculated using the measured levels (copy numbers), $X_i(r)$, of the proteins in the EGFR network and calculated protein levels at the steady state $X_i^o(r)$. When $X_i(r) = X_i^o(r)$, the value of DPE(r) at that distance range will vanish, meaning no deviation from the steady state at that r. The numerator in Eq. **2** is divided by $\sum_i X_i(r)$ (sum of all protein molecules at a given distance range r) to obtain a deviation from the steady state or a change in a free energy per protein molecule. At all values of r where the *DPE* is not zero it is mathematically guaranteed to be positive.

To obtain accurate values for the distance range where the $\lambda_1(r)$ function equals zero for the U87EGFRvIII, a polynomial curve was fitted to the $\lambda_1(r)$ function with the following restrictions: all of the original points have had to overlap with the curve as well as the region 80–100 μm ($R^2 = 1$). The same considerations were applied to $\lambda_2(r)$ and U87PTEN cells. $\lambda_\alpha(r)$ values were obtained from the curve fit for the original distance ranges and for the point close to the zero.

For molecular species in solution, the free energy is calculated as follows (6):

$$\underbrace{\mu_i}_{\text{free energy in a solution}} = \underbrace{\mu_i^{ss}}_{\text{standard free energy}} + kT\ln X_i$$

$$\underbrace{\mu_i^o}_{\text{free energy at the steady state}} = \underbrace{\mu_i^{ss}}_{\text{standard free energy}} + kT\ln X_i^o$$

These equations form the base of the Eq. **1**: $\ln X_i(r) = \ln X_i^o(r) - \sum_{\alpha=1} G_{i\alpha}\lambda_\alpha(r)$ in the main text. To calculate Eq. **1** we use $\ln X_i$ instead of $kT\ln X_i$, obtaining a deviation from the steady state $\sum_{\alpha=1} G_{i\alpha}\lambda_\alpha(r) = (\mu_i - \mu_i^o)/kT$. Therefore, the *DPE* function, which expresses the deviation from the steady state (1, 8): $\sum_i X_i(r)[\ln(X_i(r)/X_i^o(r))]$ is dimensionless (because it is divided by the unit of thermal energy kT).

Determination of the Errors for the Calculated Parameters Obtained from Surprisal Analysis. SVD analysis was performed three times: once for the mean values of the functional proteins, measured as a function of distance; once for the mean values of the protein distribution − SEM for each measured protein; and once for the mean of the distribution + SEM. Thus, three values for each parameter: $\lambda_\alpha(r)$ and $G_{i\alpha}$ were obtained. Results of surprisal analysis presented in the article are the mean values of the corresponded parameters ± SD (9, 10).

Cell–Cell Distance Distribution Measurements and Calculations of $u(r)$ Function. Two initial starting conditions were used for the bulk culture experiments. Either 33,000 (low) or 66,000 cells per well [density comparable to the single-cell barcode chip (SCBC) experiments] were seeded into six-well plates. At 6 h later, the medium was replaced with fresh medium and the cells were incubated at 37 °C for 24, 48, and 72 h. At each time point, microscopic images were analyzed using custom algorithms, detecting the centers of the cells and calculating distances between the cells, from cell center to cell center. All of the pairs having a distance up to 200 μm were binned into the histograms to calculate a probability to find a pair of the cell at a certain distance. The obtained probability was divided by a random probability of cell–cell distance distributions computed using Monte Carlo simulation. Random simulations used the same cell density as detected experimentally for each time point. Using the equation relating radial distribution function (RDF) to the two-cell potential of interaction $u(r)$, these two parameters were calculated: $RDF = P_{\text{exp}}/P_{random} = \exp(-u(r)/kT)$, where P is the probability to find two cells at a particular distance and $u(r)$ indicates potential of interaction of two-cell interaction (7).

SI Methods

Cell Culture. U87EGFRvIII and U87PTEN were constructed as reported (11) and were kindly provided by Paul Mischel's laboratory (University of California, San Diego, La Jolla, CA). The cells were cultured in DMEM (ATCC) medium supplemented with 10% (vol/vol) FBS in a 5% (vol/vol) CO_2 incubator at 37 °C.

Antibody Microarray (12). The microarrays are initially patterned as DNA arrays, because DNA has the physical and chemical stability to withstand the various processing steps of microfluidics fabrication. The DNA itself is patterned onto a polylysine-coated glass slide by two sequential microchannel-based flow patterning steps. The two microchips for those two steps are based on polydimethylsiloxane (PDMS) elastomer sealed with a glass slide, so that 20 microchannels are formed winding from one end to the other end of the glass slide. The first flow-patterning step generates 20-μm-wide 50-μm pitch lines of three unique DNA oligomers, whereas the second DNA patterning step is carried out at right angles to the first. We design the ssDNA sequences for the first and second patterning in such a way that the inter-

section of the two sets of lines will remain a unique ssDNA for assembly location for complementary ssDNA–antibody conjugates. In the current study, we chose a 3 × 3 array, which has the capacity of multiplexed measurement of up to eight different proteins. This addressable nine-element array has been repeated ~19,000 times with various orientations across the whole glass slide, which will be mated with a SCBC elastomer layer. The DNA microarrays can be stored in a desiccator at room temperature for at least a month without detectable degradation.

We validated the patterned DNA array using 100-nM cy5-conjugated cDNA oligomers in 1% BSA in PBS. After incubation for 1 h at 37 °C, the slide was rinsed with 3% BSA in PBS twice and deionized water and dried under flowing N2 before signal readout by a GenePix scanner (Molecular Devices, LLC). The validation procedure provides a check on the cross-reactivity between the anchor, bridging, and terminal ssDNA oligomers. In addition, the fluorescence intensity per unit area can be compared against standard DNA spotting approaches as a means of gauging surface coverage. Finally, the validation procedure provides an assessment of the fidelity of the microfluidics flow-patterning steps.

The DNA array was converted into antibody array using complementary ssDNA–antibody conjugates. The procedure of conjugation has been described previously (12). Calibration and cross-reactivity characterizations of the barcode assay have been performed before application to single cells. We used recombinant pEGFR, pERK, pAkt, pS6k, VEGF, and IL-6 proteins (R&D Systems, Inc.) at various concentrations to calibrate the fluorescence intensity of a barcode assay with molecule number in a cell chamber with a volume of 0.15 nL. Antibody cross-reactivity assays were carried out using spotted arrays and identical biomolecular reagents.

Cell Sample Preparation for Single-Cell Assays. U87 EGFRvIII cells and U87PTEN cells were harvested for experiment after achieving 80% confluency. A total of 1 μM 5-chloromethylfluorescein diacetate (CMFDA) in PBS was used to stain live cells and visualize cells in microchambers for counting.

On-Chip Assay and Quantification (12). The PDMS layer of SCBC chip is mated onto a barcode glass slide. The microchambers are aligned with the microarray so that each microchamber is ensured of overlapping a complete barcode. Then the microchambers and microchannels were blocked with 3% BSA in PBS for 1 h before filling a mixture of 10 μg/mL antibody–DNA conjugates. After incubation for 1 h and removal of unlinked conjugates, the SCBC was filled with a cell sample and then mounted onto a clamp. With appropriate pressure exerted by the clamp, all microchambers were closed where random number of cells resided in each microchamber. The whole structure was incubated for 6 h at 37 °C in a 5% CO_2 incubator. The clamp was then partially released to draw into the microchambers lysis buffer (20 mM Tris·HCl, 150 mM NaCl, 1 mM Na2EDTA, 1 mM EGTA, 1% Triton, 2.5 mM sodium pyrophosphate, 1 mM β-glycerophosphate, 1 mM Na3VO4, and 1 μg/mL leupeptin; Cell Signaling) supplemented with protease/phosphatase inhibitor mixture (Cell Signaling). The SCBC was incubated at room temperature for 2 h to allow proteins captured by the microarray. After labeling the microchannels using 20 μM Alexa 514 succinimidyl ester (Molecular Probes, Inc.), biotinylated detection antibody mixtures in 1:18 dilution prepared according to product instruction were injected into the microchannels. Unbound antibodies were washed off, and 1 μM cy5 streptavidin and 100 nM H-cy3 in PBS with 3% BSA were used to complete sandwich–ELISA procedure and label barcode arrays. Finally, the glass slide containing the developed arrays was detached and cleaned before scanning by a GenePix Scanner. Fluorescence intensity of each barcode across the whole slide was digitalized and exported to a custom algorithm written in MatLab (MathWorks, Inc.). One array spot [detected using the green reference channel (532 nm)] provides an alignment marker to permit protein identification by spatial location on each of the other array spots [fluorescence detection using red data channel (635 nm)]. The fluorescence intensity of 174,000 barcode spots across the whole slide was exacted and analyzed by a MatLab (MathWorks, Inc.) program developed in our laboratory.

The intensity can be further converted into protein copy number using a calibration curve.

Cell Proliferation Assay. A total of 33,000 or 66,000 cells per well were seeded into six-well plates. At 6 h later, the medium was replaced with fresh medium and the cells were incubated at 37 °C for 24, 48, and 72 h. At each time point indicated, microscopic images were taken and number of cells, normalized by the images' area, was computed using MatLab.

1. Kravchenko-Balasha N, et al. (2012) On a fundamental structure of gene networks in living cells. *Proc Natl Acad Sci USA* 109(12):4702–4707.
2. Laidler K (1996) A glossary of terms used in chemical kinetics, including reaction dynamics (IUPAC Recommendations 1996). *Pure Appl Chem* 68(1):149–192.
3. Levine RD (2005) *Molecular Reaction Dynamics* (Cambridge Univ Press, Cambridge, UK).
4. Levine RD, Bernstein RB (1974) Energy disposal and energy consumption in elementary chemical reactions. Information theoretic approach. *Acc Chem Res* 7(12):393–400.
5. Remacle F, Kravchenko-Balasha N, Levitzki A, Levine RD (2010) Information-theoretic analysis of phenotype changes in early stages of carcinogenesis. *Proc Natl Acad Sci USA* 107(22):10324–10329.
6. Mayer JE, Mayer MG (1966) *Statistical Mechanics* (Wiley, New York).
7. McQuarrie DA (2000) *Statistical Mechanics* (University Science Books, Sausalito, CA).
8. Gross A, Li CM, Remacle F, Levine RD (2013) Free energy rhythms in *Saccharomyces cerevisiae*: A dynamic perspective with implications for ribosomal biogenesis. *Biochemistry* 52(9):1641–1648.
9. Alhassid Y, Levine RD (1980) Experimental and inherent uncertainties in the information theoretic approach. *Chem Phys Lett* 73(1):16–20.
10. Gross A, Levine RD (2013) Surprisal analysis of transcripts expression levels in the presence of noise: A reliable determination of the onset of a tumor phenotype. *PLoS ONE* 8(4):e61554.
11. Wang MY, et al. (2006) Mammalian target of rapamycin inhibition promotes response to epidermal growth factor receptor kinase inhibitors in PTEN-deficient and PTEN-intact glioblastoma cells. *Cancer Res* 66(16):7864–7869.
12. Wang J, et al. (2012) Quantitating cell–cell interaction functions with applications to glioblastoma multiforme cancer cells. *Nano Lett* 12(12):6101–6106.

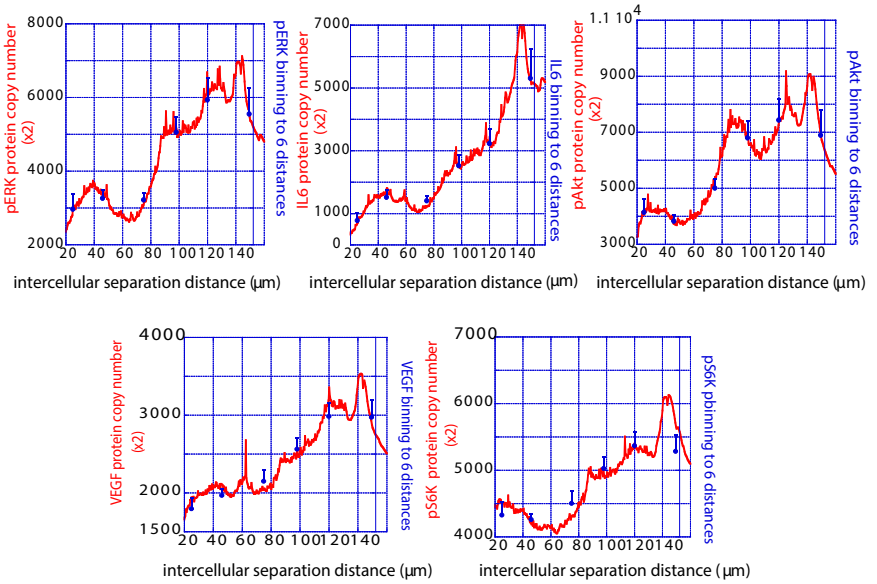

Fig. S1. Functional protein expression levels as a function of distance in U87EGFRvIII cells. Data representing functional protein expression levels as a function of distance from 500 cell pairs (×2) of U87EGFRvIII cells (smooth red curve). The data were binned into six different distance ranges (blue dots). Blue dots represent mean values of protein copy numbers at every distance range. Values are mean ± SEM.

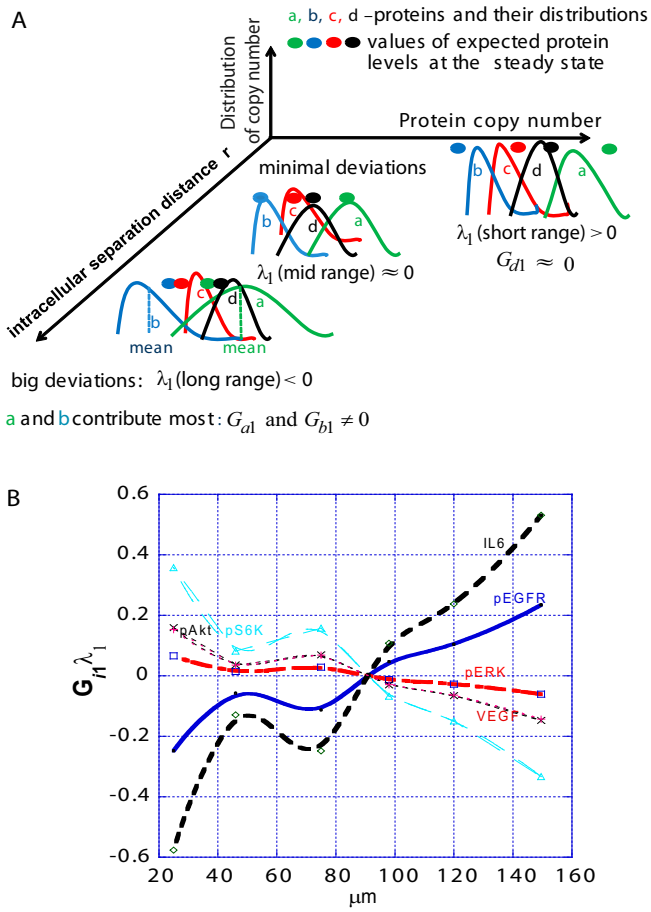

Fig. S2. Surprisal analysis of protein expression levels as a function of distance in U87EGFRvIII cells. (*A*) Illustration of surprisal analysis for a system with four measured proteins at three distance ranges. The steady-state term and one constraint ($\alpha = 1$) are illustrated. Surprisal analysis identifies protein expression levels at the steady state (the colored dots) for proteins a–d. Proteins a and b contribute most to the $\alpha = 1$ constraint (G_{a1} and G_{b1} are relatively big) as indicated by large differences between the measured mean protein levels and the expected steady-state levels. Protein d barely deviates at all distance ranges. The constraint $\alpha = 1$ has its largest weights [$\lambda_1(r) \neq 0$] at the long and short distance ranges, but is essentially zero at the midranges, implying that this midrange distance corresponds to the steady state of the cells. (*B*) $G_{i1}\lambda_1(r)$ products as calculated for U87EGFRvIII cells. $G_{i1}\lambda_1(r)$ represents the extent of deviation of each protein X_i from the steady state in U87EGFRvIII cells as a function of distance. $G_{i1}\lambda_1(r)$ for all measured proteins are plotted.

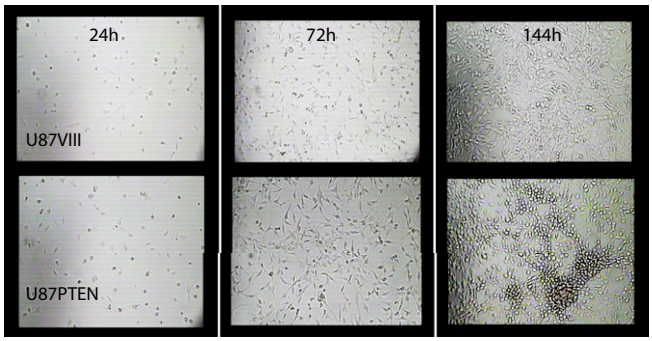

Fig. S3. Formation of 3D cell structures in U87EGFRvIII and U87PTEN cells. U87EGFRvIII cells and U87PTEN cells were seeded at a density of 66,000 cells per well in six-well plates. Three-dimensional dense structures were observed in U87PTEN cells but not in U87EGFRvIII cells after 144 h in culture.

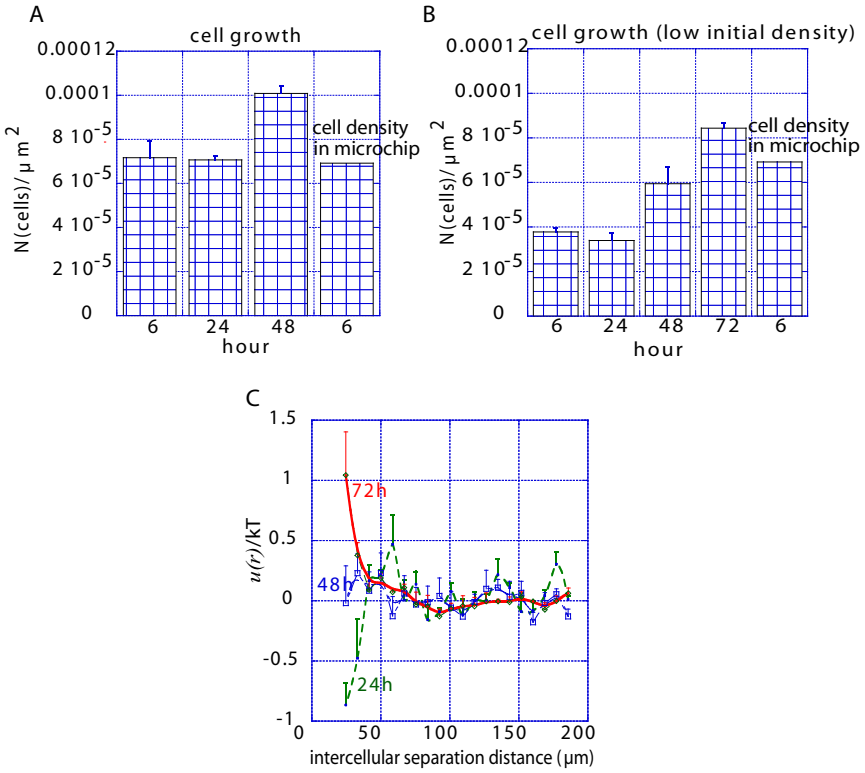

Fig. S4. Cell density analysis and calculations of $u(r)$ functions for U87EGFRvIII cells using bulk experiments. U87EGFRvIII cells were seeded at the normal density, 66,000 cells per well (*A*) and low density, 33,000 cells per well (*B*) in six-well plates. Plots represent time course analysis of the cell density. Number of cells was divided by the area. Values are mean values ± SD (at least $n = 4$). (*C*) Time-course analysis of the $u(r)$ functions calculated for U87EGFRvIII cells seeded at the low density. After 72 h, the $u(r)$ function attained the same shape (red) as the function calculated for U87EGFRvIII cells seeded at normal densities and cultured 24 h (Fig. 3).

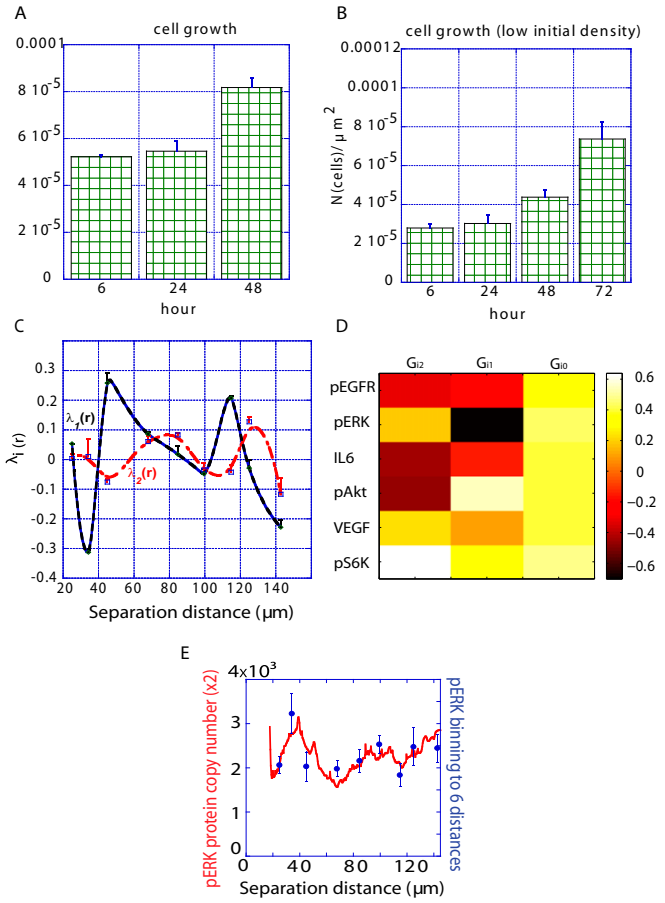

Fig. S5. Cell density analysis using bulk experiments (*A* and *B*) and results of surprisal analysis using two-cell SCBC experiments for U87PTEN cells (*C* and *D*). U87PTEN cells were seeded at a density of 66,000 cells per well (*A*) and 33,000 cells per well (*B*) in six-well plates. Plots represent time-course analysis of the cell density as described in Fig. S1. Number of cells was divided by the area. Values are mean values ± SD (at least $n = 4$). (*C*) Surprisal analysis was performed using mean values of phosphorylation/secretion as a function of distance range (mean values were calculated from $16 \leq n \leq 70$ cell pairs). $\lambda_\alpha(r)$ represents extent of deviation of the EGFR network from the steady state in U87PTEN cells as a function of distance. $\lambda_\alpha(r)$ for the constraints $\alpha = 1,2$ are plotted. The error bars for $\lambda_\alpha(r)$ were calculated by propagating errors associated with the mean values of measured proteins as a function of distance and as indicated in *Surprisal Analysis*. (*D*) Extent of participation in the biological processes ($G_{i\alpha}$), related to the constraints $\alpha = 0,1,2$ was computed for every detected protein. (*E*) SCBC data showing the measured level of phospho-ERK as a function of cell separation distance from 310 pairs of U87PTEN cells (smooth red curve). The data were binned into nine distance ranges for every measured protein (blue dots). The $16 \leq N \leq 67$ cell pairs were analyzed for every distance range.

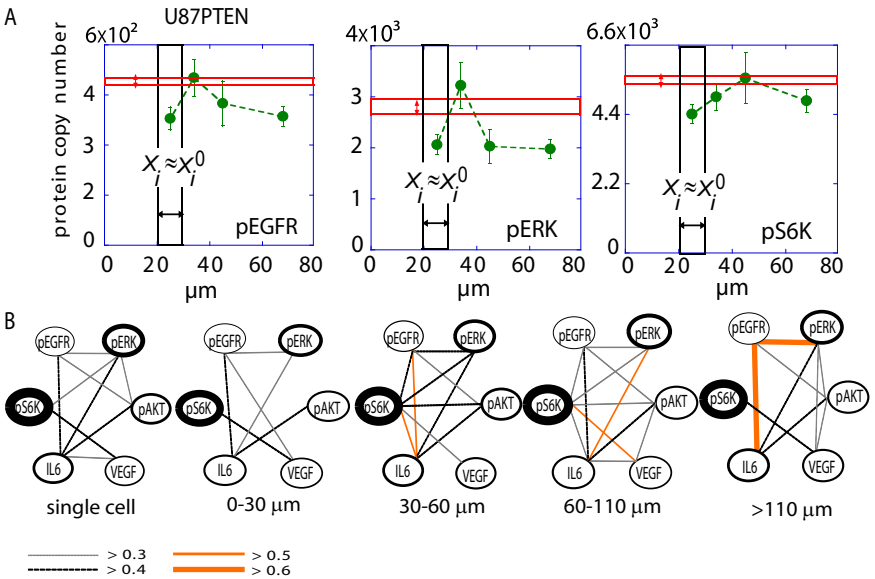

Fig. S6. Comparisons of noninteraction U87PTEN single cells with interacting cell pairs. (*A*) Mean values of protein levels as measured for U87PTEN single cells ($n = 550$) are compared against measurement of those same proteins (blue and green dots) for cell pairs at cell-separation distance ranges close to the steady state (20–30 μm for U87PTEN cells) or deviating significantly (30–50 μm). The red box represents the measured range of copy numbers (multiplied by 2, ± SEM) for single cells. The black box is that range for cell pairs at the steady state. (*B*) Protein–protein coordination maps were generated using U87PTEN two-cell and single-cell data. To compute the coordination maps the data were binned similarly to Fig. 5 with one change: the 0- to 60-μm distance range was divided further in two subranges: 0–30 μm, the distance range close to the steady state, and 30–60 μm, the most deviating distance range. All indicated correlations pass a *P* test ($P < 0.05$). The thickness of the lines encircling the protein names reflects the relative abundance of those proteins. The thickness and color of the edges reflect the extent of the protein–protein coordination, as provided in the key below the networks.

Table S1. List of reagents for GBM single/two-cell proteomic assay

Reagent	Source and sequences
Antibody name	
Mouse antihuman phospho-EGFR	R&D Y1068
Biotin-labeled goat antihuman EGFR	R&D BAF231
Antihuman phospho-ERK kit	R&D DYC1825
Anti –IL-6 kit	R&D DYC206
Antihuman phospho-mTOR kit	R&D DYC1665
Antihuman phospho-Akt1 kit	R&D DYC2289
Anti-VEGF kit	R&D DYC293B
Antihuman p70 s6 kinase	R&D DYC896
Oligonucleotide name	
2X A-3′-polyA	ATCCTGGAGCTAAGTCCGTAAAAAAAAAAAAAAAAAAAATCCTGGAGCTAAGTCCGTAAAAAAAAAAAAAAAAAAAA
2X B-3′-polyA	GCCTCATTGAATCATGCCTAAAAAAAAAAAAAAAAAAAAGCCTCATTGAATCATGCCTAAAAAAAAAAAAAAAAAAAA
2X C-3′-polyA	GCACTCGTCTACTATCGCTAAAAAAAAAAAAAAAAAAAAGCACTCGTCTACTATCGCTAAAAAAAAAAAAAAAAAAAA
A′-D-D	ATGGTCGAGATGTCAGAGTAAAAAAAAAAATGGTCGAGATGTCAGAGTAAAAAAAAAATACGGACTTAGCTCCAGGAT
D3′	NH_2-AAAAAAAAAAAAATACTCTGACATCTCGACCAT
B′-E-E	ATGTGAAGTGGCAGTATCTAAAAAAAAAAATGTGAAGTGGCAGTATCTAAAAAAAAAAATAGGCATGATTCAATGAGGC
E3′	NH_2-AAAAAAAAAAAAATAGATACTGCCACTTCACAT
C′-F-F	ATCAGGTAAGGTTCACGGTAAAAAAAAAAATCAGGTAAGGTTCACGGTAAAAAAAAAAATAGCGATAGTAGACGAGTGC
F3′	NH_2-AAAAAAAAAAAAATACCGTGAACCTTACCTGAT
A′-G-G	GAGTAGCCTTCCCGAGCATTAAAAAAAAAAGAGTAGCCTTCCCGAGCATTAAAAAAAAAATACGGACTTAGCTCCAGGAT
G′	NH_2-AAAAAAAAAAAAAAAATGCTCGGGAAGGCTACTC
B′-H-H	ATTGACCAAACTGCGGTGCGAAAAAAAAAAATTGACCAAACTGCGGTGCGAAAAAAAAAATAGGCATGATTCAATGAGGC
H′	NH_2-AAAAAAAAAAAAACGCACCGCAGTTTGGTCAAT
C′-I-I	TGCCCTATTGTTGCGTCGGAAAAAAAAAATGCCCTATTGTTGCGTCGGAAAAAAAAAATAGCGATAGTAGACGAGTGC
I′	NH_2-AAAAAAAAAAAAATCCGACGCAACAATAGGGCA
A′-J-J	TCTTCTAGTTGTCGAGCAGGAAAAAAAAAATCTTCTAGTTGTCGAGCAGGAAAAAAAAAATACGGACTTAGCTCCAGGAT
J′	NH_2-AAAAAAAAAAAAACCTGCTCGACAACTAGAAGA
B′-K-K	TAATCTAATTCTGGTCGCGGAAAAAAAAAATAATCTAATTCTGGTCGCGGAAAAAAAAAATAGGCATGATTCAATGAGGC
K′	NH_2-AAAAAAAAAAAAACCGCGACCAGAATTAGATTA
C′-L-L	GTGATTAAGTCTGCTTCGGCAAAAAAAAAAGTGATTAAGTCTGCTTCGGCAAAAAAAAAATAGCGATAGTAGACGAGTGC
L′	NH_2-AAAAAAAAAAAAAGCCGAAGCAGACTTAATCAC
Streptavidin cy 5, cell lysis buffer	Cell Signaling
FBS	Sigma-Aldrich
Protease/phosphatase inhibitor mixture	Cell Signaling
Alexa 514 succinimidyl ester	Invitrogen
CMFDA	Invitrogen

Table S2. The input data for SVD analysis [$X_i(r)$ values from U87EGFRvIII cells]

	Distance, μm					
Protein	25	46	75	98	120	149.5
pEGFR	698.1	837.5	830.1	1,620.5	2,070.0	1,981.7
pERK	2,960.9	3,258.8	3,211.3	5,051.3	5,931.0	5,544.1
IL-6	773.4	1,514.7	1,398.3	2,517.8	3,222.8	5,304.8
pAkt	4,131.4	3,824.3	5,003.8	6,796.2	7,428.8	6,891.8
VEGF	1,793.0	1,965.0	2,148.6	2,560.2	2,980.7	2,971.0
pS6K	4,324.3	4,259.3	4,501.1	5,024.9	5,365.1	5,280.5

Table S3. $\lambda_\alpha(r)$ values for U87EGFRvIII cells

Constraint α	λ_α, 25 μm	λ_α, 46 μm	λ_α, 75 μm	λ_α, 98 μm	λ_α, 120 μm	λ_α, 149 μm
0	18.6	19.0	19.1	20.0	20.4	20.5
1	0.7	0.2	0.3	−0.1	−0.3	−0.7
2	0.1	−0.2	−0.1	0.1	0.2	−0.1
3	0.0	0.1	−0.1	0.0	0.0	0.0
4	0.0	0.0	0.0	0.0	0.0	0.0
5	0.0	0.0	0.0	0.0	0.0	0.0

Table S4. $\lambda_\alpha(r)$ values for U87PTEN cells

Constraint α	λ_α, 125 μm	λ_α, 34 μm	λ_α, 45 μm	λ_α, 68 μm	λ_α, 85 μm	λ_α, 99 μm	λ_α, 119 μm	λ_α, 125 μm	λ_α, 143 μm
0	17.78	18.20	18.08	17.8	17.79	18.12	17.88	18.07	17.86
1	0.05	−0.31	0.26	0.09	0.02	−0.05	0.21	−0.03	−0.22
2	−0.01	0.00	−0.08	0.06	0.09	−0.04	−0.04	0.13	−0.11
3	0.04	−0.08	−0.06	0.02	−0.01	−0.06	0.03	0.03	0.09
4	0.06	0.03	−0.03	0.02	−0.07	0.01	0.02	0.00	−0.03
5	−0.06	0.00	−0.02	0.02	−0.02	0.02	0.04	0.01	0.00
6	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00