Overline: Computational Biology

One sentence summary: Bayesian inference-based modeling identifies the most likely paths through a signaling network.

Editor's Summary:

Picking the Right Path

Signaling networks have become increasingly complex as large-scale analysis and experiments in multiple systems add new potential connections and players. Xu *et al*. present a mathematical approach to rank the possible paths through a signaling pathway in order to develop hypotheses that can be rationally tested. They call their approach BIBm for Bayesian inference-based modeling and apply BIBm to explore the signaling pathways by which epidermal growth factor (EGF) stimulates extracellular signal-regulated kinase (ERK). Using a limited set of biochemical experiments, the authors test four models and find that the one that relies on two Raf family members to rank the highest. This model was then experimentally validated in two cell lines to show that both Raf-1 and B-Raf contribute to ERK activation in response to EGF.

Inferring Signaling Pathway Topologies from Multiple Perturbation Measurements of Specific Biochemical Species

Tian-Rui Xu[1,+], Vladislav Vyshemirsky[2,+], Amélie Gormand[1,+,&], Alex von Kriegsheim[3,$], Mark Girolami[2,*], George S Baillie[1], Dominic Ketley[2], Allan J Dunlop[1], Graeme Milligan[1], Miles D Houslay[1], Walter Kolch[1, 3,$*]

1. Faculty of Biomedical and Life Sciences, University of Glasgow, Glasgow, G12 8QQ, UK
2. Department of Computing Science, University of Glasgow, Glasgow, G12 8QQ, UK
3. The Beatson Institute for Cancer Research, Glasgow, G61 1BD, UK

[+] These authors made equal contributions

[&] Current address: Department of Experimental Medical Science, Lund University, BMC C11, SE 221 84 Lund, Sweden (email Amelie.Gormand@med.lu.se)

[$] Current address: Systems Biology Ireland, University College Dublin, Belfield, Dublin 4, Dublin, Ireland (email walter.kolch@ucd.ie)

*. To whom correspondence should be addressed.

E-mails: MG: girolami@dcs.gla.ac.uk; WK: walter.kolch@ucd.ie

ABSTRACT

The specification of biological decisions by signaling pathways is encoded by the interplay between activation dynamics and network topologies. Although we can describe complex networks, we cannot easily determine which topology the cell actually uses to transduce a specific signal. Experimental testing of all plausible topologies is infeasible due to the combinatorially large number of experiments required to explore the complete hypothesis space. Here, we demonstrate that Bayesian inference-based modeling provides an approach to explore and constrain this hypothesis space permitting the rational ranking of pathway models. Our approach can use measurements of a limited number of biochemical species when combined with multiple perturbations. As proof-of-concept, we examined the activation of the extracellular signal-regulated kinase (ERK) pathway by epidermal growth factor (EGF). The predicted and experimentally validated model shows that both Raf-1 and, unexpectedly, B-Raf are needed to fully activate ERK in two different cell lines. Thus, our formal methodology rationally infers evidentially supported pathway topologies even when a limited number of biochemical and kinetic measurements is available.

INTRODUCTION

Mathematical modeling in biology has a long and illustrious history, of providing physical explanations of biological phenomenon-- think of Hodgkin &

Huxley's model of neuronal action potential propagation (*1*). A major tenet of systems biology is to interpret and inform experimental work through an iterative process of mathematical modeling and experimental validation. When considering the structure of signaling pathways, experimental evidence reveals functional or physical interactions that give rise to hypotheses suggesting how a signal is transduced through the pathway. However, experiments done under different experimental conditions typically suggest several plausible pathway topologies, and it becomes difficult to distinguish which connections cells actually use as the number of potential connections grows. The comprehensive map of the epidermal growth factor (EGF) signaling network illustrates this problem, because when all possible connections are described (*2*), it is impossible to tell which ones are realized in a particular situation. It was demonstrated by Wilkinson (*3*) that, in principle, competing mathematical models of biochemical pathways can be evidentially ranked in an objective manner using experimental data. Thus, we suggest and employ a Bayesian Inference-Based modeling (BIBm) approach, founded on Wilkinson's assumption, that provides the means to rank alternative hypotheses about pathway structures on the basis of the evidence provided by experimental data.

BIBm takes a number of working hypotheses about the structure and dynamics of a biological process, which have attached to them preliminary levels of confidence in their validity. These levels of confidence, or prior probabilities, are informed by existing knowledge and can be formally represented as a calibrated probability distribution over all hypotheses under consideration (*4*). An abstracted mathematical model explicitly capturing all

assumptions and describing the main phenomena of the system then formally represents each hypothesis (*5*). Unlike other computational systems biology modeling approaches in which a single model is considered and subsequently refined, BIBm considers all of the plausible enumerated hypotheses in the form of mathematical models based on the amount and quality of evidential support for each of the instantiated models. This formalizes the 'exploration of hypothesis space' in that different plausible alternatives are evaluated in terms of confidence based on the strength of evidential support for each one (*5*). Thus, the Bayesian probabilistic framework (*3-5*) provides a rigorous formal methodology for characterizing and propagating uncertainty based on experimental data during model evaluation and provides the means to update the levels of confidence in each hypothesis in light of the information content of experimental data (see Supplementary Material, sections 1 and 4). Furthermore, the Bayesian hypotheses testing methodology used in the BIBm framework accounts for model complexity (6). We do not take a single optimal point of the likelihood, i.e. a single plausible value, but rather integrate (marginalize) over all of the model parameters taking into account their likely distribution as measured or estimated based on prior knowledge (prior distribution) and the likelihood. Thus, overly complex models are automatically penalized, and BIBm assigns the highest ranking to the simplest model(s) that suffice to explain the experimental evidence. The updated probabilities (posteriors) may re-rank and select between prior hypotheses, or suggest further experimentation, if preconceived models have similar weights of posterior probability (*5, 6*). This procedure is iterated until the obtained inferences allow to draw conclusions in respect to the original hypothesis

about the biological system closing the cycle of hypothesis generation, model development, model assessment, and experimental validation (Fig. 1). Thus, BIBm permits evaluation of pathway models based on a limited set of measurements and multiple perturbations, and it accommodates dynamic models within a probabilistic framework.

We tested BIBm on the real biological network of the activation of the extracellular signal-regulated kinase (ERK) pathway by EGF, with ERK activation as the readout. We perturbed the pathway by pharmacologically inhibiting or activating adenosine 3',5'-monophosphate (cAMP) signaling, which has several points of crosstalk with the EGF-mediated activation of ERK (*7, 8*). The ERK pathway is activated by surface receptors, such as the EGF receptor (EGFR), which activate a Ras family guanosine triphosphatase (GTPase) at the cell membrane by recruiting guanine nucleotide exchange factors (GEFs) that exchange GDP for GTP (*9*). Ras-GTP binds to and initiates the activation of kinases of the Raf family. Raf then phosphorylates and activates mitogen-activated protein kinase kinase (MEK), and MEK phosphorylates and activates ERK (*9-11*). ERK participates in the regulation of fundamental cellular processes (*12*), and its deregulation is implicated in the pathogenesis of many diseases, especially cancer (*10*). Deregulation typically occurs at the level of Ras and Raf activation, a complex part of the pathway with many possible isoforms and regulatory inputs. Ras activates all three Raf isoforms A-Raf, B-Raf, and Raf-1, whereas B-Raf can also be activated by another GTPase Rap1 (*8, 13*). However, activation of EPAC, a Rap1 GEF, by the cAMP analog 8CPT-2Me-cAMP did not activate ERK (*14*),

and it is unclear which growth factors and under what conditions Rap1 activates B-Raf (*15*). cAMP signaling can inhibit Raf-1 and activate B-Raf (*7*) confounding the analysis of cAMP effects on ERK activity in cells where both Raf-1 and B-Raf are present. Exploring all possibilities experimentally would require a prohibitively large number of experiments. Thus, we used BIBm to analyze the most plausible pathway topologies based on a limited, and hence feasible, set of experimental measurements.

RESULTS

Application of BIBm to analyze the EGF activated ERK pathway

On the basis of the published literature, we selected four different equally plausible pathway topologies for EGF-mediated activation of ERK (Fig. 2A). Model details are described in section 3 of the Supplementary Materials. In Model 1, activation of ERK proceeds through Ras (*16, 17*) and the model includes the possibility that Ras activates both Raf-1 and B-Raf (*18*). In Model 2, two branches of ERK activation occur through Ras and Rap1 (*8, 13*). Models 3 and 4 correspond to Models 1 and 2, but include EGFR desensitization, which may critically regulate ERK activity (*19*).

By assessing the informativeness of measurements of different species in the pathway, we found that ERK activity provided the most informative data with an eigenvalue of 0.2127, which was 0.57 times greater than the next best data set (MEK, with an eigenvalue of 0.1216) (see Materials and Methods and Supplementary Materials, section 1). Therefore, we measured the kinetics of ERK activation in PC12 cells under 11 distinct conditions and perturbations,

which represented a total of 168 experimental measurements (Supplementary Materials, section 2.3). Each time point and each replicate is considered a separate measurement. To perturb the pathway, we used drugs affecting the activity of the cAMP signaling system, which has multiple points of crosstalk with the EGF-stimulated ERK activation pathway (Fig. 2A and Fig. S1). We targeted protein kinase activated by cAMP (PKA) with the cAMP analog 6-Benz-cAMP, EPAC with the cAMP analog 8pMeOPT-2-O-cAMP, and phosphodiesterase 3 (PDE3) with cilostamide, which we determined was the most active PDE in PC12 cells (Fig. S2D). We simulated model dynamics with ordinary differential equations (ODEs; Supplementary Materials, section 3) with the following simplifications: (i) The manipulation of the abundance of cAMP by different means is considered as perturbations that are not part of the topology models; (ii) Cilostamide was modeled as direct 'activator' of both EPAC and PKA because it increases cAMP by inhibiting PDE3 (*20*); (iii) The activation processes from receptor to adaptors to effector proteins were defined as EGFR→GRB2→SOS →Ras and EGFR→CRK→C3G→Rap1 pathways (*16*).

To rank the four alternative hypotheses, we computed the posterior probability of every model given the experimental data (Figs. 2C and 2D; and Supplementary Materials, section 4) to obtain the Bayes' factors. The resulting Bayes' factors are the ratio of likelihoods that a given model is consistent with observed experimental data, and hence can be used for model ranking. The Bayes' factor is a summary of the evidence provided by the data in favor of one scientific hypothesis, represented by a mathematical model, in comparison to another. Posterior odds for preferring one model to another can

be obtained by multiplying the Bayes' factor by the prior odds. Thus, we can base our ranking on probability distributions, which give us exact information on the probability that our conclusions are correct.. Comparing the four alternative models gave Bayes' factors of $\log_{10}\{B_{2,1}\} \approx 7.45 \pm 2.60$, $\log_{10}\{B_{2,3}\} \approx 23.61 \pm 1.47$, and $\log_{10}\{B_{2,4}\} \approx 4.56 \pm 1.81$. We used the mean ± standard error format to report the estimated values of the Bayes' factor, which indicated a significant preference (*6*) of the two-branched model 2 over model 4 and the single branch models. Model 2 is about 100 times more likely than Model 4, >3 million times more likely than Model 1, and $>10^{22}$ more likely than Model 3 (Fig. 2D). Thus, the Bayes' factors show that EGF uses both Raf-1 and B-Raf to activate ERK, and that EGF receptor desensitization is a minor contributor to ERK activation kinetics in response to EGF and cAMP signaling. Only the experimental data from the 168 perturbation measurements of ERK phosphorylation were used to compute the marginal likelihoods of the alternative models, and these were sufficient to select Model 2 as the most plausible one for the structure of the signaling pathway (Figs. 2C and 2D).

Experimental validation of BIBm predictions

These predictions of pathway topology were validated by siRNA experiments in PC12 and HEK293 cells. Both Raf-1 and B-Raf are present in both of these cell lines, and the cells share a similar regulation of ERK activation kinetics in response to EGF (*21*). Due to the low transfection efficiency of PC12 cells, we used HEK293 for the detailed mechanistic experiments (Figs. 3 and 4), and validated salient results in PC12 cells (Fig.

5). In HEK293 cells (Fig. 3), knocking down either the GRB2 adaptor protein (by ~70%), which mediates Ras activation, or the CRK adaptor protein (by ~50%), which mediates Rap1 activation, significantly attenuated EGF-stimulated ERK activation compared to control cells. Knocking down both GRB2 (by ~80%) and CRK (by ~40%) further reduced ERK activation and was slightly, but significantly, more efficient than the single knockdowns (Fig. 3B-D).

We also knocked down Raf-1 and B-Raf, which are the relevant Ras and Rap1 effectors (Figs. 3E-H). Knocking down Raf-1 (by ~80%) significantly attenuated EGF-stimulated ERK activation at all timepoints, whereas knocking down B-Raf (~90%) had no significant effect on ERK activation (Fig. 3F and G). However, the B-Raf knockdown significantly enhanced the effects of the Raf-1 knockdown at all timepoints tested (Fig. 3G). These results suggest that Raf-1 is the main transducer of EGF-mediated ERK activation with an additional contribution from the Rap1 - B-Raf branch. These experimental results are consistent with a B-Raf function for which Raf-1 is rate limiting, such as the formation of Raf-1- B-Raf heterodimers that enable B-Raf to signal through Raf-1 (*22, 23*). In fact, we observed that a small proportion of Raf-1 and B-Raf coimmunoprecipitated in unstimulated cells and that EGF increased the Raf-1 – B-Raf interaction (Fig. 4A). To explore the role of heterodimerization, we knocked down upstream components and measured dimer formation and ERK phosphorylation. Knockdown of either GRB2 or CRK reduced Raf-1 – B-Raf heterodimer formation and the amount of phosphorylated ERK (Fig. 4B). The double knockdown had a stronger effect on both parameters, suggesting that

mechanistically the contributions of the Ras and Rap1 branches converge on Raf-1 – B-Raf heterodimerization and subsequent ERK activation. Therefore, we also assessed the activation of Raf-1 and B-Raf by EGF with antibodies against critical activating phosphorylation sites, S338 and S445, respectively (*11, 24*). Both the CRK, as well as the GRB2, knockdown reduced the activation of Raf-1 and B-Raf with the double knockdown having a stronger effect (Fig. 4C). Raf-1 activation was more severely affected than B-Raf activation by the single and double knockdowns. These results are consistent with observations that Raf-1 is strongly activated by heterodimerization with B-Raf (*22, 23*), and indicate that Raf-1 serves as the rate-limiting enzyme that determines the output of the heterodimer towards ERK activation.

Rap1 was reported to selectively activate B-Raf (*25*), but either to not affect (*15*) or even inhibit (*26*) Raf-1 in Rat-1 cells. Therefore, it was surprising that the CRK knockdown interfered with Raf-1 activation. However, a dominant-negative Rap1 mutant also reduced Raf-1 activation by EGF, confirming the siRNA results (Fig. 4D). In addition, activation of endogenous Rap1 by a constitutively active EPAC mutant did not interfere with Raf-1 activation (Fig. S3). Therefore, the negative effects of silencing the CRK branch on Raf-1 activation are most likely due to the interference with Raf-1 - B-Raf heterodimerisation rather than a direct negative effect of Rap1 on Raf-1.

Using confocal microscopy and biochemical cell fractionation, we examined whether knocking down GRB2 or CRK, or both, would alter the subcellular localization and EGF mediated membrane recruitment of Raf-1 and B-Raf, but no significant changes were observed (Fig. S4). This apparent

lack of effect on localization could be because these techniques fail to capture changes in transient relocalization events, or because Raf proteins may have alternative modes of associating with membranes, for example through lipid-binding domains, which are independent of the activation of Ras and Raf (*27-29*). This two-branched pathway topology would not have been revealed by experiments with individual siRNA knockdowns of Raf-1 or B-Raf, this dual use of the Raf proteins was revealed from analyzing the plausibility of the models.

To validate the predicted branched pathway model in PC12 cells we nucleofected PC12 cells with siRNAs targeting Raf-1 and B-Raf (Fig. 5). In these cells, downregulation of either Raf isoform reduced EGF-stimulated ERK activation with B-Raf having a more pronounced effect. The double knockdown did not further enhance the effect of the B-Raf knockdown, suggesting that in terms of EGF-stimulated ERK activation B-Raf is the rate-limiting Raf isoform in PC12 cells. These results are consistent with the B-Raf having greater kinase activity than does Raf-1 (*18*). Despite differences in the mechanistic details between EGF-stimulated ERK activation in HEK293 and PC12 cells, BIBm correctly identified the most likely hypothesis in terms of pathway structure. Furthermore, BIBm highlighted a non obvious pathway topology and produced predictions that could guide experimental analysis.

DISCUSSION

The results that the EGFR uses at least two different complexes of adaptor proteins and GEFs and two different Raf isoforms to activate ERK was unexpected. The prevalent opinion in the field is that B-Raf is the major

activator of the ERK pathway, although data from Raf isoform knockout mice have shown that either Raf-1 or B-Raf may function, depending on cell type and context (*30*). Our method provides a rational and tractable framework for capturing such complex relationships. For instance, the cAMP crosstalk with the EGF-activated ERK pathway is highly specific. In PC12 cells, PKA activation slightly inhibits ERK activation, whereas EPAC activation, or even more efficiently, PDE3 inhibition increases ERK activation by EGF (Fig. 2B). Although PKA is distributed throughout the cell, PDE3 resides at the membrane and hence has a localized effect on cAMP abundance (*31*). Thus, increases in cAMP can have entirely different effects, depending on how and where it is generated (*32, 33*). This intricate functional differentiation could be implicitly captured by BIBm analysis.

The technical demands of Bayesian inference for large-scale ODE models are particularly challenging. and pose questions regarding efficient posterior sampling and computation of marginal likelihoods. Firstly, for a fixed computational resource such as compute time, it is difficult to construct a Markov Process that will converge to the actual posterior distribution and then, secondly, explore the parameter space according to this distribution. Thirdly obtaining unbiased low-variance estimates of the marginal likelihood which appears in the expression for the Bayes factor is notoriously challenging in general. This difficulty arises due to the high dimensionality of the parameter space and the complex likelihood function induced by the dynamics of the ODE model, see (*34*) for examples. To overcome these practical issues in this work we have employed new methods, presented in the supplementary material and for further analysis see (*34*). These are based on Population-

based Markov Chain Monte-Carlo sampling integrated with Stratified Path Sampling of the Thermodynamic Integral which provides the means to simultaneously obtain unbiased estimates of the marginal likelihood and perform posterior sampling.

As this study demonstrates, methods to reduce the computational load and focus on informative parameters enabled the application of a consistent Bayesian inferential framework to model-based reasoning and effective design of investigative experimental strategies. Inferring signaling networks is one of the biggest challenges in systems biology due to the complexity and dynamic nature of these networks. Identifying the correct network topology is also constrained by the usually limited data sets available, and the effort and resources needed to provide accurate biochemical readings at various time points and conditions. Thus, the development of enhanced analysis methods allows more targeted biochemical experimentation. In this study, weakly informative prior distributions over kinetic model parameters were employed to reflect the uncertainty about plausible parameter values. However, the proposed methodology also permits the use of more confident priors in situations when prior experimental evidence about plausible parameter values is available.

Another approach to infer network topology is Modular Response Analysis (MRA) (*35*), which employs systematic perturbations of every network component to reconstruct the network topology. Although powerful when applied to small networks (*36*), the experimental work can become unfeasible, as every single component needs to be perturbed. Whereas BIBm is based on probabilistic evaluation, MRA lacks probabilistic semantics and

cannot provide measures of uncertainty with which to reason. Both methods require prior knowledge of network components. Although MRA can infer network topologies independent of an extensive prior knowledge of connectivity, the large number of perturbations required makes it experimentally costly, especially for larger networks. BIBm requires preconceived models, that is hypotheses, to compare, but can do so with a limited number of experimental observations. This ability of BIBm to use a limited type of experimental data is particularly relevant because it is compatible with high-throughput assay platforms, which typically measure a single output under multiple perturbation conditions. Thus, MRA and BIBm seem highly complementary, and it will be interesting to explore whether they can be combined for enhanced pathway mapping.

MATERIALS AND METHODS

Implementation of BIBm

The statistical and technical challenges in computing the probabilities of the dynamical system models are formidable, especially when only sparse data sets are available (3, 5, 7). BIBm overcomes this problem with a combinatorial strategy that is briefly described here and fully described in the Supplementary Materials, sections 1 and 4. BIBm first assesses the informativeness of measured biochemical species by calculating the sensitivity of the potential data sets to small parameter variations within the plausible parameter range (Supplementary Materials, section 1). The inner product of these sensitivity matrices is an approximation of the Fisher Information matrix conditional on the corresponding data set. Then we

evaluated eigenvalues of this Fisher Information matrix. The largest eigenvalues correspond to the most informative set of species to be measured.

At the next stage, we employed Markov Chain Monte Carlo (MCMC) sampling on models of biochemical systems described by Ordinary Differential Equations (ODEs) to compute Bayes' factors (5, 6) as ratios of marginal likelihoods (MLs) for any possible couples of models. For a mathematical model S_i, the marginal likelihood to reproduce the empirical data D is $p(D|S_i) = \int p(D|S_i,\theta)p(\theta)d\theta$ where θ is a vector of all unknown model values and parameters. Nonlinear systems of ODEs are used to define mathematical models of biochemical systems, and, therefore, the likelihood function $p(D|S_i,\theta)$ is a nonlinear function that cannot be evaluated analytically. We employed methods based on thermodynamic integration to obtain stable low variance estimates of the marginal likelihoods, as well as samples from the posterior of plausible parameter values that explain the empirical data. The starting point for this work is the thermodynamic integral that defines the logarithm of the marginal likelihood in terms of a series of intermediate canonical distributions interpolating between the prior distribution of the model parameters and the posterior distribution using an inverse temperature parameter $\beta \in [0,1]$ such that

$$I = \log p(D|S_i) = \int_0^1 E_{\theta|D,\beta}\{\log p(D|S_i,\theta)\}d\beta$$

where the expectation is taken with respect to the power-posterior

$$p(\theta|D,S_i,\beta) = \frac{p(D|S_i,\theta)^\beta p(\theta)}{\int p(D|S_i,\theta)^\beta p(\theta)d\theta}.$$

We developed a MCMC procedure based on a Stratified Path Sampling strategy to estimate the thermodynamic integral (Supplementary Materials, section 4). We converted the unit interval for β into N-1 subintervals such that $\sum_{n=1}^{N-1}\Delta\beta_n = 1$ where each $\Delta\beta_n = \beta_{n+1} - \beta_n$. Then we obtained a stratified path sampling estimate by making M_n draws from each of the n strata such that $\tilde{\beta}_n \sim p_n(\beta)$ and $\theta_m^n \sim p(\theta|D, S_i, \tilde{\beta}_n)$ where $p_n(\beta)$ denotes the sampling density of the n'th stratum:

$$\hat{I}_S = \sum_{n=1}^{N-1} \frac{1}{M_n} \sum_{m=1}^{M_n} \frac{\log p(D|S_i, \theta_m^n)}{p_n(\tilde{\beta}_m)} .$$

Setting each of the sampling densities of the strata to be uniform, $p_n(\beta) = U[\beta_n, \beta_{n+1}] = \Delta\beta_n^{-1}$ the estimator becomes

$$\hat{I}_S = \sum_{n=1}^{N-1} \frac{\Delta\beta_n}{M_n} \sum_{m=1}^{M_n} \log p(D|S_i, \theta_m^n) .$$

Samples $\theta_m^n \sim p(\theta|D, S_i, \tilde{\beta}_n)$ were obtained using MCMC based on the Metropolis-Hastings scheme.

Mathematical models and Bayesian inference

Detailed mathematical descriptions of the four ODE models are presented in the Supplementary Materials, section 3. Most of the reactions are described by the Michaelis-Menten kinetic law either directly or in a modified form where the enzyme concentration is taken into account similar to the modeling approach employed by Brown *et al.(16)*. There are about 20 species and 50 parameters involved in the considered models. We formulate our prior distributions of model parameters on the presumption that all the biochemical

processes are slow enough to be observed. Thus, the value for the limiting rate should be small, and the value for the Michaelis constant should be on the same scale as our initial concentrations. Because rate parameters are strictly positive, we employ the Gamma distribution for the limiting rate of the reaction $\Gamma(1.1,9.0)$, where $\Gamma(a,b)$ denotes a Gamma distribution (*37*) with shape and inverse scale parameters a and b, respectively. To keep the reactions slow enough the Michaelis constant should be on the scale of substrate concentration. Therefore, we assume a $\Gamma(2.0, 3333.0)$ prior distribution. The prior distribution over alternative models was uniform, reflecting no *a priori* preference for any of the hypotheses. The choice of Gamma prior distributions also shifts the weight of the *a priori* belief for the reaction rates away from zero, therefore inducing an assumption that all of the reactions have a noticeable contribution to the dynamics of the model, and preventing reduction of Models 3 and 4 to Models 1 and 2, respectively, by penalizing an assignment of the receptor desensitization rate to zero. Our choice of priors for this work was based on general assumptions about desirable properties of the models, such as reaction rates being non-zero and contributing to the output of the model. More informative priors may be used within the described framework when additional information is available. For example, kinetic rates of some reactions may be estimated with in vitro assays. In such a case, an informative prior can be defined as centered at the estimated value for a kinetic rate parameter and having appropriate variance to express the confidence about such estimates. We did not have such detailed information available *a priori*, and, therefore, weakly informative priors with large variance were employed. The choice of weakly informative

priors causes larger variance of the posteriors to be obtained at the parameter inference stage of the study. These variable posteriors correspond to the multitude of possible alternative solutions of the inference problem. It is very important that we integrate over all of the possible solutions when estimating the marginal likelihoods, because otherwise model uncertainty would not be taken into account properly. Therefore, BIBm integrates over all of the possible solutions when ranking alternative hypotheses.

We performed Bayesian model comparison to assess the proposed hypotheses about the pathway structure embodied in models 1, 2, 3 and 4. We describe a Stratified Path Sampling strategy with Population MCMC, which was used to obtain estimates of the required marginal likelihoods (see Supplementary Materials. Section 4.1 for details).

Perturbation analysis in PC12 cells

The experimental data were measured as relative concentrations of double phosphorylated, i.e. activated ERK (ERK-PP) in comparison to the total ERK (Supplementary Material, section 2). Absolute protein concentrations for many of the species cannot be measured easily and reliably. In fact, the vast majority of biochemical data represents relative values describing changes of rates or concentrations. This motivated our choice of the initial concentrations for the species of the mathematical models. We used arbitrary concentration units similar to the Brown et al. model (17). Attention was paid to ensure that relative sizes of protein pools are preserved, i.e. that Raf-1 is less abundant than ERK. All of the model predictions were also produced on the scale of relative concentrations preventing possible errors due to the choice of initial concentrations.

We performed 168 individual experiments with PC12 cells in 11 different conditions Supplementary Materials, section 2.3). Activation of ERK was quantified from the amount of activating phosphorylation detected by Western blotting. Cells were grown to ~80% confluency in normal growth medium (Dulbecco's Modified Eagle's Medium containing 10% horse serum and 5% foetal bovine serum) before being serum starved for 3 hours prior to any of the experiments. Cells were then pre-treated for 10 minutes with cAMP analogues: 10 μM 8pMeOPT-2′-O-Me-cAMP (Biolog), and 10 μM 6-Benz-cAMP (Biolog) to activate specifically EPAC or PKA, respectively. Cilostamide (Sigma) was applied at 10 μM to inhibit PDE3. Following pretreatment, the cells were stimulated with 100 ng/ml EGF for 0, 2, 5, 10, 20 and 40 minutes, and then the cells were lysed in lysis buffer (25 mM Hepes, 50 mM NaCl, 10% glycerol, 1% Triton containing protease and phosphatase inhibitors). Cell debris was removed by centrifugation. The protein concentration of each cell lysate was measured and normalized to the same concentration in each experiment in order to load the same amount of protein in each gel. Proteins were separated by NuPAGE®Novex 4-12% Bis-Tris gels electrophoresis and transferred on nitrocellulose membranes. The membranes were immunoblotted with specific antibodies directed against phosphorylated (mouse monoclonal Phospho-p44/42 MAPK (Thr202/Tyr204), Cell Signaling Technology) or nonphosphorylated ERK (rabbit polyclonal ERK1/2 antibody, Cell Signaling Technology). Primary antibodies were detected with fluorescent secondary antibodies (AlexaFluor® 680 Anti-Rabbit or Anti-Mouse IgG Molecular Probes; IRDye™ 800 Anti-Rabbit or Anti-Mouse IgG, Rockland

Inc.). Fluorescence associated with the nitrocellulose membranes were analyzed with an infrared scanner (LICOR, ODYSSEY). The two different antibodies were simultaneously quantified on the same gel by scanning two different wavelengths simultaneously, allowing a relative ratio of phosphorylated ERK to total ERK to be calculated.

Antibodies, cells, and siRNAs for biochemical validation experiments

Monoclonal antibodies for GRB2, CRK, and Raf-1 were from BD. Monoclonal C3G and β-actin antibodies were from Santa Cruz. Polyclonal B-Raf antibody was from Millipore. Monoclonal p44/42 MAPK (ERK1/2), phosphorylated 44/42 MAPK (phospho-ERK1/2), phospho-Raf-1 (Ser^{338}), and phospho-B-Raf (Ser^{445}) antibodies were from Cell Signaling Technology (Hitchin, Hertfordshire, UK). Lipofectamine 2000 transfection reagent was from Invitrogen and used according to manufacturer's instructions. Human EGF was from Promega. cAMP analogues were from Biolog. All other materials were supplied by Sigma.

HEK293 cells were maintained in Dulbecco's modified Eagle's medium supplemented with 0.292 g/liter L-glutamine and 10% (v/v) newborn calf serum at 37°C in a 5% CO_2 humidified atmosphere. PC12 cells were maintained in RPMI-1640 plus 10% horse serum and 5% FBS.

Small interfering RNA duplex oligonucleotides targeting human GRB2 and CRK for use in HEK293 cells were purchased from Ambion. The sequence of the siRNA targeting GRB2 was GGUUUUGAACGAAGAAUGU-dTdT. The sequence of the siRNA targeting C3G was GGGAAAACCAGCUGAGGUG-dTdT. The sequence of the siRNA targeting

CRK was GGGAUAGGUAUCUUGCCUC-dTdT. The Silencer control siRNA1 from Ambion was used as negative control. Each siRNA (100nM final concentration) was transfected into HEK293 cells using Lipofectamine 2000 reagent (Invitrogen) according to Invitrogen's instructions. 36 hours after transfection, cells were serum starved for 12 h before treatment with 10 ng/ml EGF.

The RNA oligonucleotides targeting human Raf-1 and B-Raf were from Thermo Scientific. The siGENOME SMARTpool Raf-1 siRNA sequences were GCACGGAGAUGUUGCAGUA, GCAAAGAACAUCAUCCAUA, GACAUGGAAAUCCAACAAUA, and CAAAGAACAUCAUCCAUAG. The ON-TARGETplus SMARTpool B-Raf siRNA sequences were CAUGAAGACCUCACAGUAA, UCAGUAAGGUACGGAGUAA, AGACGGGACUCGAGUGAUG and UUACCUGGCUCACUAACUA. HEK293 cells were transfected as above and serum starved 48 hours later.

For rat PC12 cells, smart-pool siRNAs (Dharmacon) were used to knockdown Raf-1 and B-RAF, and a nonTargeting siRNA pool (Dharmacon) was used as control. 40 pM siRNA oligonucleotides were introduced into PC12 cells by nucleofection (Amaxa Biosystems, Germany) according to the manufacturer's instructions. Cells were seeded on collagen, 24h later serum starved overnight, and then stimulated with 20 ng/ml EGF.

ERK1/2 phosphorylation assay in validation experiments

Hek 293 cells transfected with siRNA were rendered quiescent by serum starvation for 6 or 12 hours prior to stimulation with 10 ng/ml EGF for 0, 2, 5, 10, 20, or 40 minutes. These periods of serum starvation gave identical ERK

activation profiles in response to EGF stimulation, and therefore were used alternatively depending on logistical requirements. After stimulation cells were immediately placed on ice and solubilized directly in Laemmli loading buffer. The samples were sonicated and then heated for 15 min at 100°C. ERK1/2 phosphorylation was detected by protein immunoblotting with the phospho-ERK1/2-specific antibody. After phospho-ERK1/2 detection the nitrocellulose membranes were stripped of immunoglobulins and reprobed with the ERK1/2 antibody to normalize the loading of proteins. The same set of samples were separated on SDS-PAGE and blotted with different antibodies to examine the efficiency of the knockdowns.

Raf-1 – B-Raf coimmunoprecipitation experiments

Cells were lysed in immunoprecipitation buffer (150 mM NaCl, 0.01 mM NaPO4, 2 mM EDTA, 1mM Na_3VO_4, 0.5% Triton X-100 and 5% glycerol plus protease inhibitor cocktail tablets). Samples were then centrifuged for 15 minutes at 20,000*g* at $4\,^{\circ}\mathrm{C}$, and the supernatant was transferred to a fresh tube with Protein G beads (Sigma) to preclear the samples. After incubation on a rotating wheel for 1 hour at $4\,^{\circ}\mathrm{C}$, the samples were re-centrifuged at 20,000*g* at $4\,^{\circ}\mathrm{C}$ for 1 min, and the protein concentration of the supernatant was determined. Samples containing equal amounts of protein were incubated for 1 hour with the Raf-1 antibody (BD Bioscience) and Protein G for another 2 hours at $4\,^{\circ}\mathrm{C}$ on a rotating wheel. Samples were then washed four times with immunoprecipitation buffer and resolved by SDS-PAGE and immunoblotted with the indicated antibodies.

SUPPLEMENTARY MATERIALS

REFERENCES

1. A. L. Hodgkin, A. F. Huxley, A quantitative description of membrane current and its application to conduction and excitation in nerve. *J Physiol* **117**, 500-544 (1952).

2. K. Oda, Y. Matsuoka, A. Funahashi, H. Kitano, A comprehensive pathway map of epidermal growth factor receptor signaling. *Mol Syst Biol* **1**, 2005 0010 (2005).

3. D. J. Wilkinson, Stochastic modelling for quantitative description of heterogeneous biological systems. *Nat Rev Genet* **10**, 122-133 (2009).

4. A. Gelman, J. B. Carlin, H. S. Stern, D. B. Rubin, Bayesian Data Analysis. *Chapman & Hall/CRC, Boca Raton, Florida, USA* (2003).

5. V. Vyshemirsky, M. A. Girolami, Bayesian ranking of biochemical system models. *Bioinformatics* **24**, 833-839 (2008).

6. R. E. Kass, A. E. Raftery, Bayes Factors. *Journal of the American Statistical Association* **90**, 773-795 (1995).

7. M. D. Houslay, W. Kolch, Cell-type specific integration of cross-talk between extracellular signal-regulated kinase and cAMP signaling. *Mol Pharmacol* **58**, 659-668 (2000).

8. P. J. Stork, J. M. Schmitt, Crosstalk between cAMP and MAP kinase signaling in the regulation of cell proliferation. *Trends Cell Biol* **12**, 258-266 (2002).

9. M. M. McKay, D. K. Morrison, Integrating signals from RTKs to ERK/MAPK. *Oncogene* **26**, 3113-3121 (2007).

10. A. S. Dhillon, S. Hagan, O. Rath, W. Kolch, MAP kinase signalling pathways in cancer. *Oncogene* **26**, 3279-3290 (2007).

11. C. Wellbrock, M. Karasarides, R. Marais, The RAF proteins take centre stage. *Nat Rev Mol Cell Biol* **5**, 875-885 (2004).

12. Y. D. Shaul, R. Seger, The MEK/ERK cascade: from signaling specificity to diverse functions. *Biochim Biophys Acta* **1773**, 1213-1226 (2007).

13. S. Kao, R. K. Jaiswal, W. Kolch, G. E. Landreth, Identification of the mechanisms regulating the differential activation of the mapk cascade by epidermal growth factor and nerve growth factor in PC12 cells. *J Biol Chem* **276**, 18169-18177 (2001).

14. J. M. Enserink, A. E. Christensen, J. de Rooij, M. van Triest, F. Schwede, H. G. Genieser, S. O. Doskeland, J. L. Blank, J. L. Bos, A novel Epac-specific cAMP analogue demonstrates independent regulation of Rap1 and ERK. *Nat Cell Biol* **4**, 901-906 (2002).

15. F. J. Zwartkruis, R. M. Wolthuis, N. M. Nabben, B. Franke, J. L. Bos, Extracellular signal-regulated activation of Rap1 fails to interfere in Ras effector signalling. *Embo J* **17**, 5905-5912 (1998).

16. K. S. Brown, C. C. Hill, G. A. Calero, C. R. Myers, K. H. Lee, J. P. Sethna, R. A. Cerione, The statistical mechanics of complex signaling networks: nerve growth factor signaling. *Phys Biol* **1**, 184-195 (2004).

17. B. Schoeberl, C. Eichler-Jonsson, E. D. Gilles, G. Muller, Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF receptors. *Nat Biotechnol* **20**, 370-375 (2002).

18. R. Marais, Y. Light, H. F. Paterson, C. S. Mason, C. J. Marshall, Differential regulation of Raf-1, A-Raf, and B-Raf by oncogenic ras and tyrosine kinases. *J Biol Chem* **272**, 4378-4383 (1997).

19. S. Traverse, K. Seedorf, H. Paterson, C. J. Marshall, P. Cohen, A. Ullrich, EGF triggers neuronal differentiation of PC12 cells that overexpress the EGF receptor. *Curr Biol* **4**, 694-701 (1994).

20. J. A. Beavo, L. L. Brunton, Cyclic nucleotide research -- still expanding after half a century. *Nat Rev Mol Cell Biol* **3**, 710-718 (2002).

21. C. Kiel, L. Serrano, Cell type-specific importance of ras-c-raf complex association rate constants for MAPK signaling. *Sci Signal* **2**, ra38 (2009).

22. M. J. Garnett, S. Rana, H. Paterson, D. Barford, R. Marais, Wild-type and mutant B-RAF activate C-RAF through distinct mechanisms involving heterodimerization. *Mol Cell* **20**, 963-969 (2005).

23. L. K. Rushworth, A. D. Hindley, E. O'Neill, W. Kolch, Regulation and role of Raf-1/B-Raf heterodimerization. *Mol Cell Biol* **26**, 2262-2272 (2006).

24. A. S. Dhillon, A. von Kriegsheim, J. Grindlay, W. Kolch, Phosphatase and feedback regulation of Raf-1 signaling. *Cell Cycle* **6**, 3-7 (2007).

25. M. R. Vossler, H. Yao, R. D. York, M. G. Pan, C. S. Rim, P. J. Stork, cAMP activates MAP kinase and Elk-1 through a B-Raf- and Rap1-dependent pathway. *Cell* **89**, 73-82 (1997).

26. S. J. Cook, B. Rubinfeld, I. Albert, F. McCormick, RapV12 antagonizes Ras-dependent activation of ERK1 and ERK2 by LPA and EGF in Rat-1 fibroblasts. *EMBO J* **12**, 3475-3485 (1993).

27. M. Hekman, H. Hamm, A. V. Villar, B. Bader, J. Kuhlmann, J. Nickel, U. R. Rapp, Associations of B- and C-Raf with cholesterol, phosphatidylserine, and lipid second messengers: preferential binding of Raf to artificial lipid rafts. *J Biol Chem* **277**, 24090-24102 (2002).

28. G. Muller, P. Storz, S. Bourteele, H. Doppler, K. Pfizenmaier, H. Mischak, A. Philipp, C. Kaiser, W. Kolch, Regulation of Raf-1 kinase by TNF via its second messenger ceramide and cross-talk with mitogenic signalling. *EMBO J* **17**, 732-742 (1998).

29. M. A. Rizzo, K. Shome, S. C. Watkins, G. Romero, The recruitment of Raf-1 to membranes is mediated by direct interaction with phosphatidic acid and is independent of association with Ras. *J Biol Chem* **275**, 23911-23918 (2000).

30. G. Galabova-Kovacs, A. Kolbus, D. Matzen, K. Meissl, D. Piazzolla, C. Rubiolo, K. Steinitz, M. Baccarini, ERK and beyond: insights from B-Raf and Raf-1 conditional knockouts. *Cell Cycle* **5**, 1514-1518 (2006).

31. Y. Shakur, K. Takeda, Y. Kenan, Z. X. Yu, G. Rena, D. Brandt, M. D. Houslay, E. Degerman, V. J. Ferrans, V. C. Manganiello, Membrane localization of cyclic nucleotide phosphodiesterase 3 (PDE3). Two N-terminal domains are required for the efficient targeting to, and association of, PDE3 with endoplasmic reticulum. *J Biol Chem* **275**, 38749-38761 (2000).

32. G. S. Baillie, M. D. Houslay, Arrestin times for compartmentalised cAMP signalling and phosphodiesterase-4 enzymes. *Curr Opin Cell Biol* **17**, 129-134 (2005).

33. M. Mongillo, T. McSorley, S. Evellin, A. Sood, V. Lissandron, A. Terrin, E. Huston, A. Hannawacker, M. J. Lohse, T. Pozzan, M. D. Houslay, M. Zaccolo, Fluorescence resonance energy transfer-based analysis of cAMP dynamics in live neonatal rat cardiac myocytes reveals distinct functions of compartmentalized phosphodiesterases. *Circ Res* **95**, 67-75 (2004).

34. B. Calderhead, M. Girolami, N. Lawrence, Accelerating Bayesian Inference over Nonlinear Differential Equations with Gaussian Processes. *Twenty-Second Annual Conference on Neural Information Processing Systems, NIPS 2008* MIT Press, Boston, USA (2008).

35. B. N. Kholodenko, A. Kiyatkin, F. J. Bruggeman, E. Sontag, H. V. Westerhoff, J. B. Hoek, Untangling the wires: a strategy to trace functional interactions in signaling and gene networks. *Proc Natl Acad Sci U S A* **99**, 12841-12846 (2002).

36. S. D. Santos, P. J. Verveer, P. I. Bastiaens, Growth factor-induced MAPK network topology shapes Erk response determining PC-12 cell fate. *Nat Cell Biol* **9**, 324-330 (2007).

37. A. Papoulis, S. U. Pillai, Probability, Random Variables and Stochastic Processes. *4th edition, McGraw-Hill Higher Education, New York, USA* (2002).

38. **ACKNOWLEDGEMENTS.** Expression vectors for Rap1b mutants, pcDNA3-HA-RAP1b-V12 and pcDNA3-HA-RAP1b-N17, were kindly provided by Dr S Yarwood. Funding Sources: This work was supported by grants from the UK Department of Trade and Industry (Beacon Project; AG, VV, WK), Medical Research Council (G0400053 and G0600765; GSB, WK MDH), the European Union (LSHB-CT-2006-037189; MDH), Science Foundation Ireland (06/CE/B1129), and by the Fondation Leducq (06CVD02; MDH and GSB). MG is funded by an EPSRC Advanced Research Fellowship, EP/E052029, and BBSRC project grant BB/G006997/1.

Author contributions: TRX, AG, AvK and AD performed the biochemical measurements and perturbation experiments. VV and WK wrote the manuscript. TRX, AG, VV, GM, GB and MH analyzed the data. VV and DK developed the models and performed the computational analysis.

Competing interests: None

Figure Legends

Fig. 1. Outline of the Bayesian Inference-Based modeling (BIBm) approach. BIBm uses prior knowledge and experimental data **(A)** to formulate alternative working hypotheses **(B)**. These working hypotheses are then formally defined using mathematical models **(C)**, and ranked according to a Bayesian inferential framework that extracts the information from experimental data and consequently derives plausible distributions of model parameters. Alternative models then can be compared by the weight of evidence from the experimental data supporting them **(D)**. This comparison and ranking of alternative models by evidence support allows refinement of

the set of working hypotheses, and selection of the most probable hypotheses for experimental validation **(E)**. The cycle of formulating and formalizing hypotheses, evaluating evidential support, and refining the set of hypotheses may be repeated several times until a validated model is found **(F)**.

Fig. 2. BIBm evaluation of the EGF-mediated ERK activation and crosstalk with the cAMP signaling system. (A) Schematic representation of the topology of the four plausible pathway models (hypotheses) and the cAMP-mediated input. Black arrows represent stimulatory inputs and dashed red lines represent inhibitory inputs. **(B)** Experimental ERK activation data observed in response to perturbations of the cAMP signaling system. **(C)** Prediction of models matched to the experimental data. The depicted graph demonstrates the predictions made using Model 2, but all models were matched to the data similarly (see Figs. S13-S16). **(D)** BIBm analysis based on these data selected the dual-path structure of Model 2 as the significantly most plausible one.

Fig. 3. Experimental validation of the BIBm analysis in HEK293 cells by knocking down critical components predicted by BIBm of the ERK-activating pathways. HEK293 cells were transfected with the indicated siRNAs and serum starved prior to stimulation with 10 ng/ml EGF for 5 min. ERK-PP is phosphorylated ERK1 and 2 detected with a phosphorylation-specific antibody, and ERK is total ERK1 and 2 detected with an antibody that recognizes all ERK1/2. **(A)** Scheme of siRNA knockdowns of EGFR-mediated signaling complexes. **(B)** Knocking down GRB2, CRK, or both reduces ERK activation. Protein loading was monitored by reprobing the same membrane

with the ERK1/2 antibody. Co, control siRNA. **(C)** Bar graph depicting the ratio of P-ERK1/2 to ERK1/2 in cells with the indicated molecules knocked down. Data shown represent the mean ± standard deviation from at least 5 separate knockdown experiments. Asterisks denote statistical significance as measured by t-test ($p \leq 0.0012$) **(D)** The knockdown efficiency was detected by blotting the same set of ERK activation samples shown in (B) with the respective antibodies. Protein loading was monitored by blotting for β-actin. **(E)** Scheme of siRNA knockdowns of Raf isoforms. **(F)** ERK activation in Raf-1, B-Raf, or double knockdown cells was examined by blotting for P-ERK1/2 in cells transfected with siRNA and stimulated as indicated.Potein loading was monitored by reprobing the same membrane with ERK1/2 antibody. **(G)** Bar graph depicting the ratio of P-ERK1/2 to ERK1/2 in cells transfected with the indicated siRNAs and then exposed to EGF for various times. Data shown represent the mean ± standard deviation from at least 6 separate experiments (*$p<0.05$, **$p<0.01$). **(H)** The knockdown efficiency of Raf-1 and B-Raf as detected by Western blotting. Protein loading was monitored by reprobing with the β-actin antibody.

Fig. 4. Dual input through CRK and GRB2 is required for EGF-induced Raf-1 and B-Raf heterodimerization and activation. **(A)** HEK293 cells were serum starved overnight and stimulated with 10 ng/ml EGF for the time indicated. Endogenous Raf-1 was immunoprecipitated (IP) and blotted (IB) for the presence of B-Raf. Whereas little B-Raf coimmunoprecipitated with Raf-1 in the absence of EGF, EGF transiently increased Raf protein heterodimerization. Lysates were blotted for B-Raf to assure that the abundance of B-Raf did not change during the experiment. **(B)** HEK293 cells

were transfected with GRB2 and CRK siRNAs for 36 hours, serum starved, and stimulated with EGF for 5 minutes. Raf-1—B-Raf heterodimerization was assayed as in (A). The amount of phospho-ERK (ERK-PP) correlated with the extent of Raf protein heterodimerization. **(C)** Cells were transfected with siRNAs targeting GRB2 and CRK, serum starved, and stimulated with EGF as in (A). Lysates were blotted with the indicated antibodies that detected the indicated proteins or phosphorylated proteins. **(D)** Cells were transfected with activated (Rap1b-V12) or dominant-negative (Rap1b-N17) Rap1 mutants and analyzed as in (C).

Fig. 5. Experimental validation of the BIBm analysis in PC12 cells. PC12 cells were transfected with control (Cont) and siRNAs as indicated. Twenty-four hours later cells were serum starved overnight and stimulated with EGF. **(A)** Cell lysates were Western blotted with antibodies to detect the indicated proteins and quantified with fluorescent secondary antibodies and a LICOR scanner. **(B)** The graph represents the average of three independent experiments. Error bars indicate standard deviation.

Fig. 1

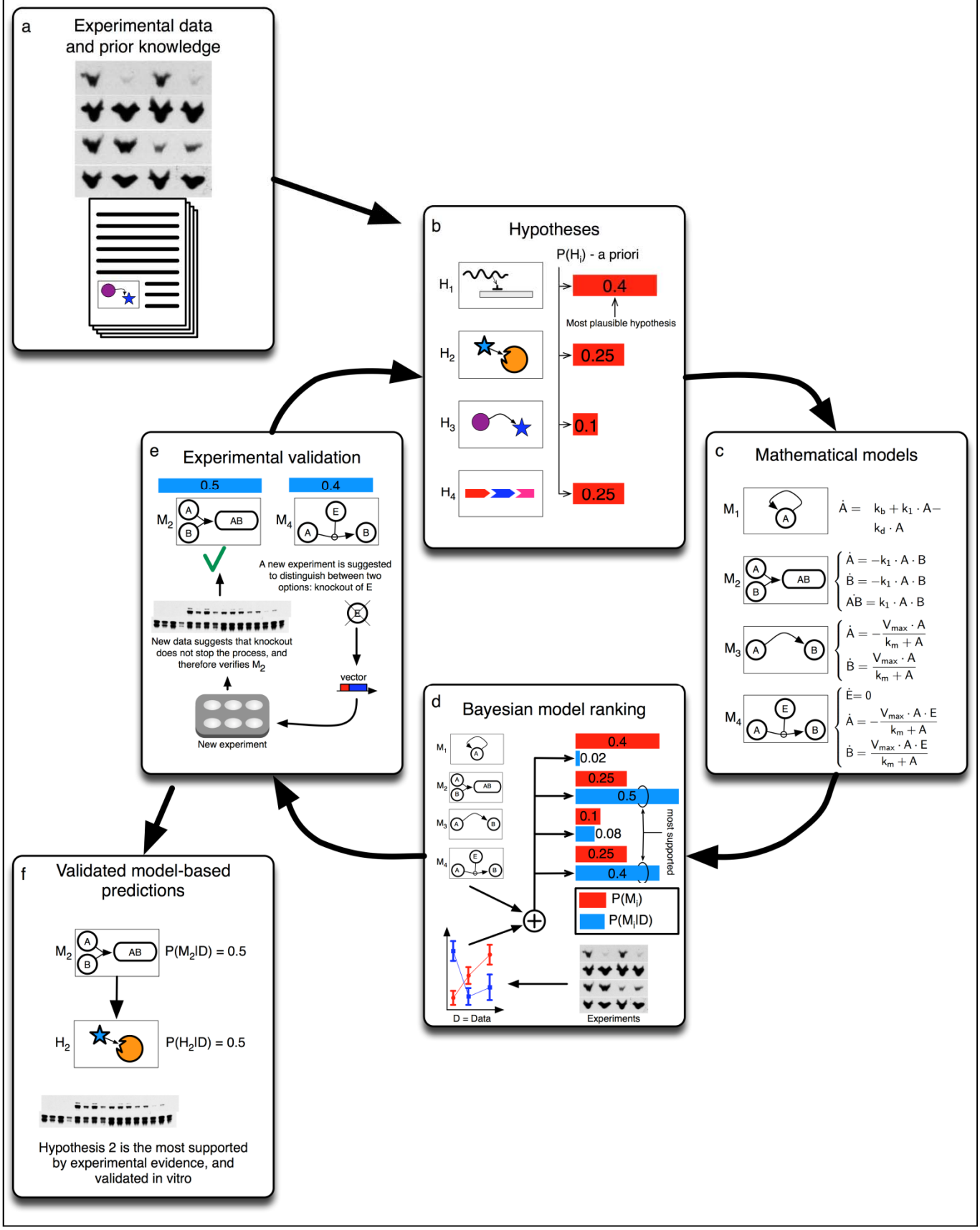
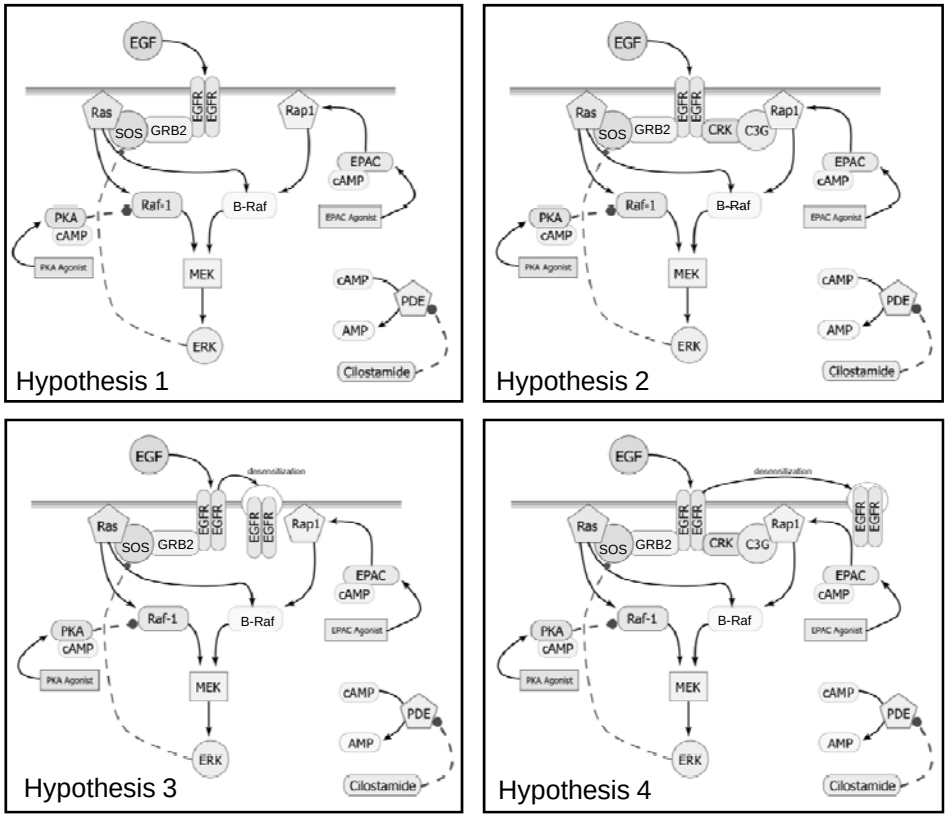
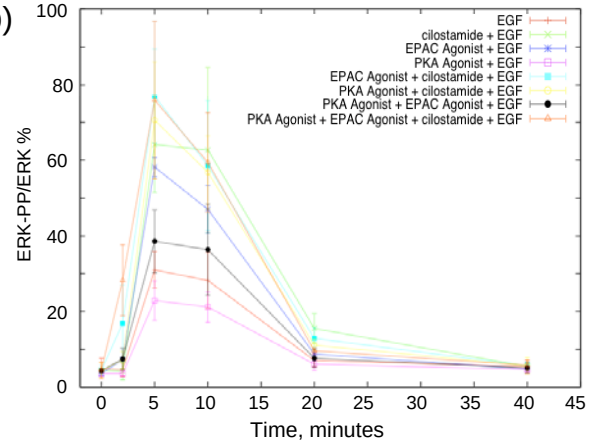

a
Experimental data and prior knowledge
b
Hypotheses
P(H_i) - a priori
H1
0.4
Most plausible hypothesis
H2
0.25
H3
0.1
H4
0.25
c
Mathematical models
M1
A
Ȧ = kb + k1 · A − kd · A
M2
A
B
AB
Ȧ = −k1 · A · B
Ḃ = −k1 · A · B
ȦḂ = k1 · A · B
M3
Ȧ = −Vmax · A / km + A
Ḃ = Vmax · A / km + A
M4
E
Ė = 0
Ȧ = −Vmax · A · E / km + A
Ḃ = Vmax · A · E / km + A
d
Bayesian model ranking
M1
M2
M3
M4
0.4
0.02
0.25
0.5
0.1
0.08
0.25
0.4
most supported
P(Mi)
P(Mi|D)
D = Data
Experiments
e
Experimental validation
0.5
0.4
M2
M4
A new experiment is suggested to distinguish between two options: knockout of E
New data suggests that knockout does not stop the process, and therefore verifies M2
vector
New experiment
f
Validated model-based predictions
M2
P(M2|D) = 0.5
H2
P(H2|D) = 0.5
Hypothesis 2 is the most supported by experimental evidence, and validated in vitro

Fig. 2

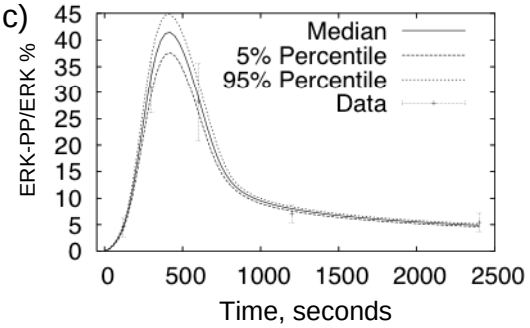
a)


EGF
Ras
SOS
GRB2
EGFR
Rap1
EPAC
cAMP
Raf-1
B-Raf
PKA
MEK
ERK
PDE
AMP
Cilostamide
CRK
C3G
Hypothesis 1
Hypothesis 2
Hypothesis 3
Hypothesis 4

b)


ERK-PP/ERK %
Time, minutes
EGF
cilostamide + EGF
EPAC Agonist + EGF
PKA Agonist + EGF
EPAC Agonist + cilostamide + EGF
PKA Agonist + cilostamide + EGF
PKA Agonist + EPAC Agonist + EGF
PKA Agonist + EPAC Agonist + cilostamide + EGF

c)


ERK-PP/ERK %
Time, seconds
Median
5% Percentile
95% Percentile
Data

d)

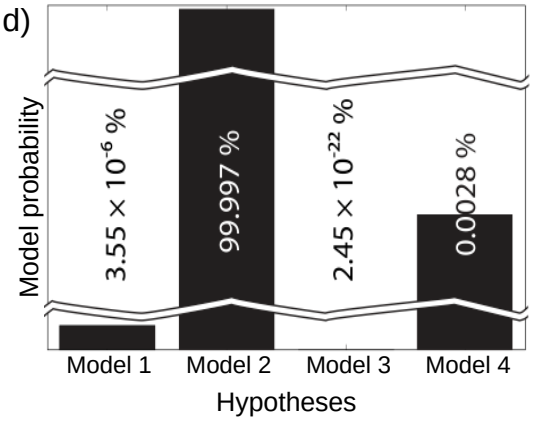

Model probability
3.55 × 10-6 %
99.997 %
2.45 × 10-22 %
0.0028 %
Model 1
Model 2
Model 3
Model 4
Hypotheses

Fig. 3

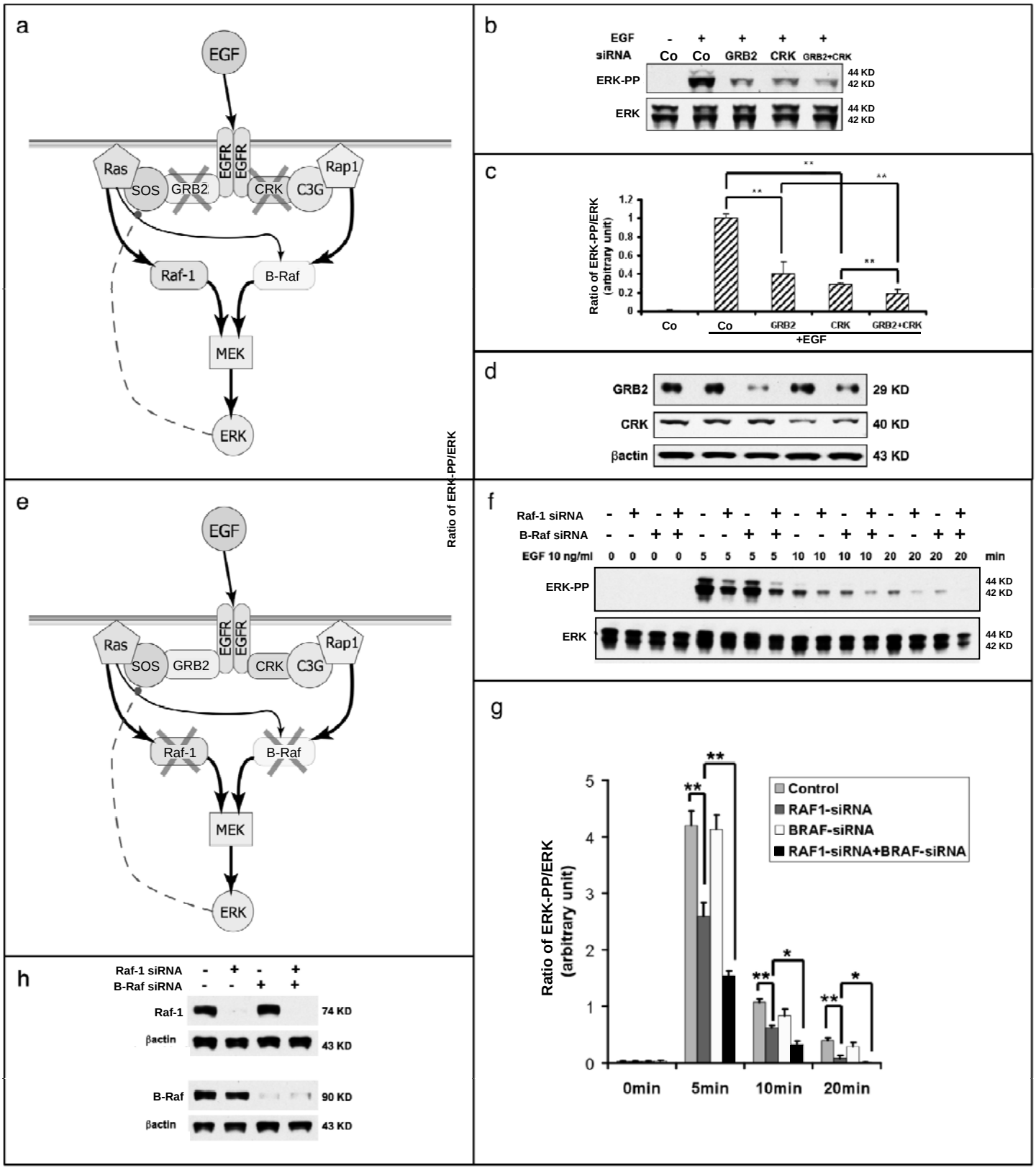

a
EGF
EGFR
EGFR
Ras
SOS
GRB2
CRK
C3G
Rap1
Raf-1
B-Raf
MEK
ERK
b
EGF - + + + +
siRNA Co Co GRB2 CRK GRB2+CRK
ERK-PP
ERK
44 KD
42 KD
c
Ratio of ERK-PP/ERK (arbitrary unit)
Co Co GRB2 CRK GRB2+CRK
+EGF
d
GRB2
CRK
βactin
29 KD
40 KD
43 KD
Ratio of ERK-PP/ERK
e
EGF
EGFR
EGFR
Ras
SOS
GRB2
CRK
C3G
Rap1
Raf-1
B-Raf
MEK
ERK
f
Raf-1 siRNA
B-Raf siRNA
EGF 10 ng/ml 0 0 0 0 5 5 5 5 10 10 10 10 20 20 20 20 min
ERK-PP
ERK
44 KD
42 KD
g
Ratio of ERK-PP/ERK (arbitrary unit)
Control
RAF1-siRNA
BRAF-siRNA
RAF1-siRNA+BRAF-siRNA
0min 5min 10min 20min
h
Raf-1 siRNA
B-Raf siRNA
Raf-1
B-Raf
βactin
74 KD
90 KD
43 KD

Fig. 4

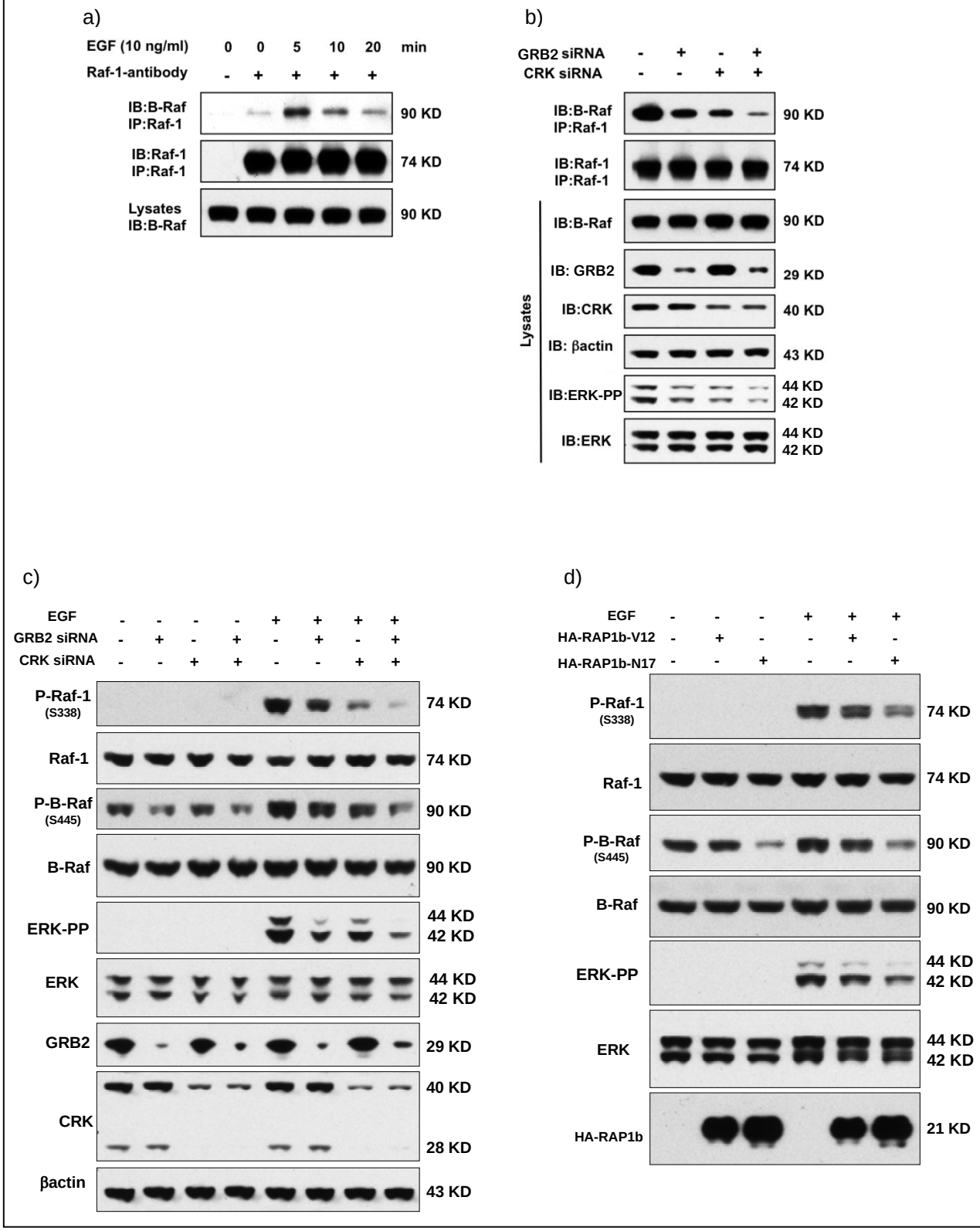
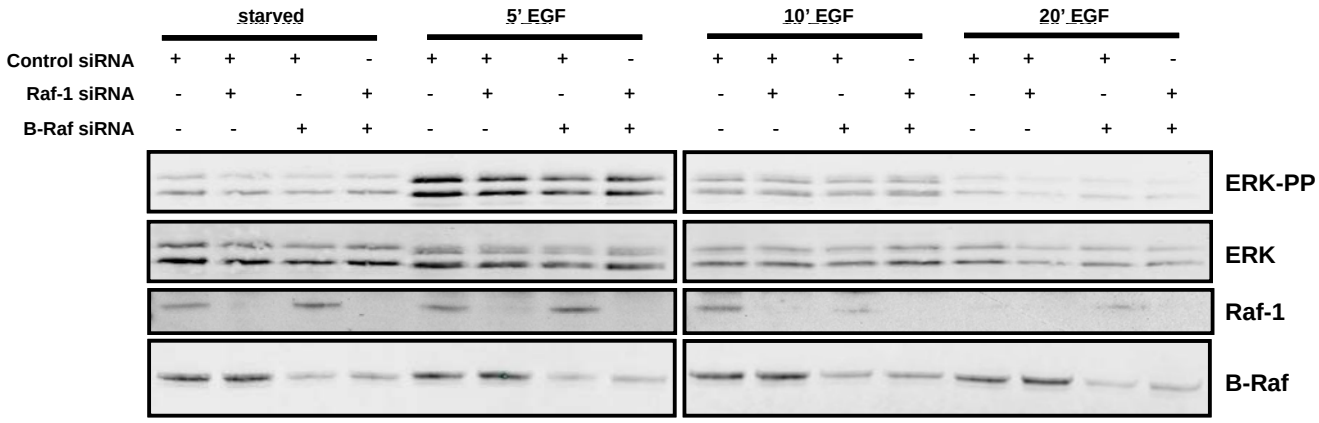
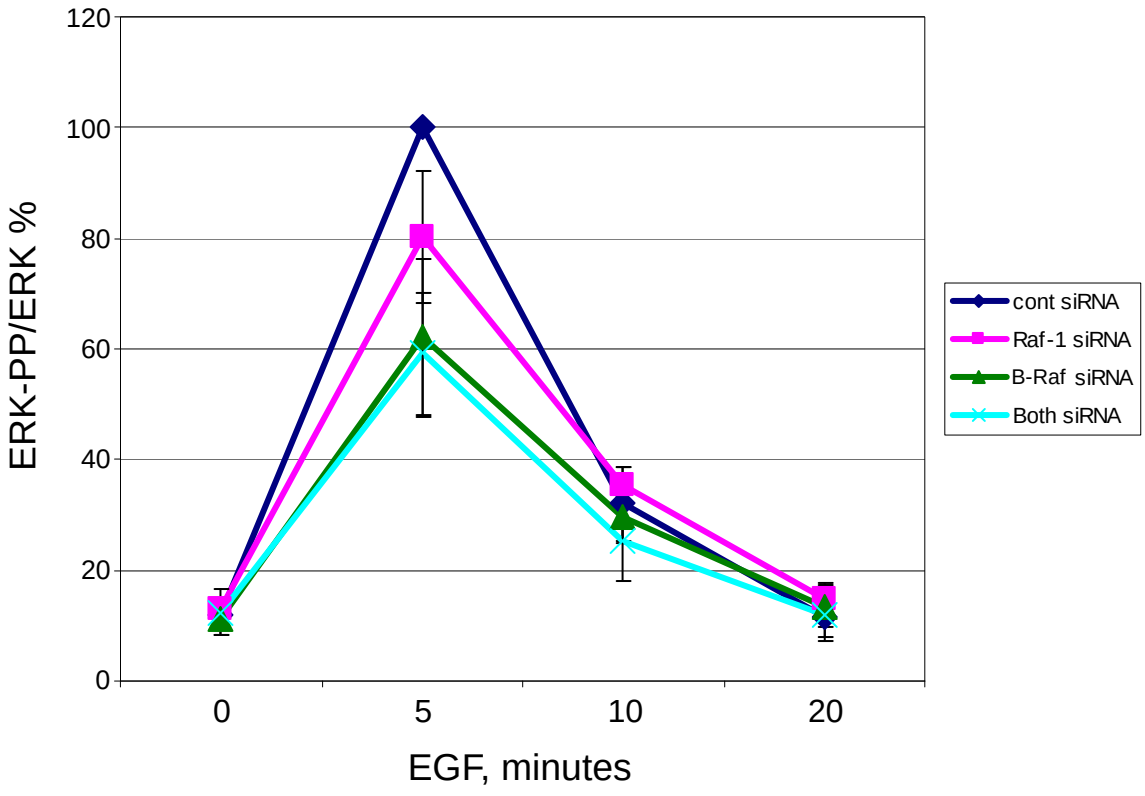

a)
EGF (10 ng/ml) 0 0 5 10 20 min
Raf-1-antibody - + + + +
IB:B-Raf IP:Raf-1
90 KD
IB:Raf-1 IP:Raf-1
74 KD
Lysates IB:B-Raf
90 KD
b)
GRB2 siRNA - + - +
CRK siRNA - - + +
IB:B-Raf IP:Raf-1
90 KD
IB:Raf-1 IP:Raf-1
74 KD
Lysates
IB:B-Raf
90 KD
IB: GRB2
29 KD
IB:CRK
40 KD
IB: βactin
43 KD
IB:ERK-PP
44 KD
42 KD
IB:ERK
44 KD
42 KD
c)
EGF - - - - + + + +
GRB2 siRNA - + - + - + - +
CRK siRNA - - + + - - + +
P-Raf-1 (S338)
74 KD
Raf-1
74 KD
P-B-Raf (S445)
90 KD
B-Raf
90 KD
ERK-PP
44 KD
42 KD
ERK
44 KD
42 KD
GRB2
29 KD
CRK
40 KD
28 KD
βactin
43 KD
d)
EGF - - - + + +
HA-RAP1b-V12 - + - - + -
HA-RAP1b-N17 - - + - - +
P-Raf-1 (S338)
74 KD
Raf-1
74 KD
P-B-Raf (S445)
90 KD
B-Raf
90 KD
ERK-PP
44 KD
42 KD
ERK
44 KD
42 KD
HA-RAP1b
21 KD

Fig. 5


starved
5' EGF
10' EGF
20' EGF
Control siRNA
Raf-1 siRNA
B-Raf siRNA
ERK-PP
ERK
Raf-1
B-Raf


ERK-PP/ERK %
120
100
80
60
40
20
0
0
5
10
20
EGF, minutes
cont siRNA
Raf-1 siRNA
B-Raf siRNA
Both siRNA