Biology Using Engineering Tools: The Negative Feedback Amplifier

Marc R. Birtwistle\textsuperscript{1,2} and Walter Kolch\textsuperscript{1}

\textsuperscript{1}Systems Biology Ireland, University College Dublin, Belfield, Dublin 4, Republic of Ireland
\textsuperscript{2}Corresponding author: Marc Birtwistle; marc.birtwistle@ucd.ie

Key Words: negative feedback, signal transduction, quantitative modeling, mitogen activated protein kinase, spatiotemporal dynamics

Running Title: Biological Negative Feedback Amplifiers

List of Abbreviations: negative feedback amplifier (NFA), Extracellular-signal Regulated Kinase 1/2 (ERK1/2), Mitogen-activated protein kinase (MAPK), mitogen-activated protein kinase kinase kinases (MAP3K or MAPKKK), mitogen-activated protein kinase kinase (MAP2K or MAPKK), receptor tyrosine kinases (RTKs), Ras guanidine nucleotide exchange factors (RasGEFs), doubly phosphorylated, active ERK (ppERK), epidermal growth factor (EGF), EGF Receptor (EGFR/ErbB1), dual specificity phosphatase (DUSP), Raf-oestrogen receptor fusion protein (BXB-ER),
Abstract

Negative feedback is an ubiquitous feature of biological networks. Recent work from Sturm and colleagues\textsuperscript{1} presents experimental evidence that biological negative feedback can serve the same function as it does for engineered systems: robustness to perturbations within the feedback loop. Such behavior has important implications for how to attack deregulated signaling networks containing negative feedback in diseases such as cancer.
Main Text

How does a cell in an organism differentiate signals from noise while being immersed in bath of growth factors and hormones? This issue is one of the major challenges in understanding how the high fidelity and specificity of biological responses are generated. Engineers have considered similar problems for a long time. Can we borrow concepts from engineering to understand biology?

In the early 20th century, the reliable transmission of telephone signals became a growing problem as lines became longer. To transmit signals over longer distances, amplification was needed, but noise and distortion were added every time the signal was amplified. A solution to this problem was the negative feedback amplifier (NFA), an invention made in 1927 by Harold Stephen Black, an employee of Bell Laboratories, who proposed that the output of the amplifier be used to attenuate its input, creating a negative feedback loop. This proposal seemed so counterproductive that it took 9 years for a patent to be issued. Why would one want to decrease the magnitude of a signal that needs to be amplified? There are actually two good reasons for doing this. First, consider the effects of signal amplification on the dynamic range of the transduction system. With higher amplification (gain), the output will be saturated at lower input magnitudes, reducing the dynamic range. However, with negative feedback in place, the amplification is diminished, thus allowing the output to respond to a greater range of input magnitudes, in a more linear fashion. Second, consider that during signal transmission, there is a perturbation to the amplifier, such that its output magnitude is now suddenly too high. With a negative feedback loop in place, this increased magnitude is passed back to the amplifier input as increased negative feedback strength, consequently attenuating the input signal and mitigating the effects of the perturbation. Thus, although the amplifier gain is reduced by the negative feedback, the NFA affords increased resistance to perturbations within the amplifier and a greater range of responsiveness to different input strengths. These properties seem very useful for biological signaling networks that constantly have to deal with intrinsic and extrinsic noise while simultaneously responding to a sea of varied-concentration growth factors.
Recent work demonstrates that the mammalian Extracellular-signal Regulated Kinase 1/2 (ERK1/2) cascade has properties similar to the NFA found in electronic systems, namely, it amplifies input signals, makes input-output response curves more linear, and confers robustness to perturbations within the feedback loop. But how exactly does an enzymatic kinase cascade function as a NFA? And what are the implications for biology and medicine?

**MAPK cascades as signal amplifiers**

Mitogen-activated protein kinase (MAPK) cascades are conserved from yeast to mammals. Although originally discovered as kinases that are activated by mitogens, the acronym has become the name for a large family of kinases that respond to different signals including hormones, stress, radiation and mechanical cues. They share a common topology with a three-tiered module of kinases at the core. First, mitogen-activated protein kinase kinase kinases (MAP3K or MAPKKK) activate a mitogen-activated protein kinase kinase (MAP2K or MAPKK) by phosphorylating two serines in the MAP2K activation loop. The activated MAP2K in turn phosphorylates a threonine and tyrosine in the MAPK activation loop thereby activating the MAPK. The kinase module in general functions as a linear rather than a branched transducer, as MAP2Ks are highly selective, only featuring MAPKs as substrates. Although in stress activated MAPK cascades MAP2Ks can be more promiscuous, they still only phosphorylate MAPKs. For instance, the MAP2Ks MKK4 and MKK7 cooperate to activate JNK MAPK, but while MKK7 selectively activates JNK, MKK4 can activate both JNK and p38 MAPKs. In contrast, MAPKs typically have many substrates, making signal transduction promiscuous at the cascade output level. These various pathway outputs can be specified by different MAPK activation dynamics, and are thought to mediate various biological responses.

What is the reason for having a linear cascade of three or even four kinases? Why would cells favour a design encompassing several processes to accomplish a single communication task? One reason is that the cascade design enables large peak signalling magnitudes combined with stable “off” states. In addition, it offers more interfaces for independent regulation by feedback, cross-talk, and
scaffolding, which allows the integration of different inputs and the fine tuning of the activation kinetics. In a pathway where output specificity is controlled by the MAPK activation dynamics, the accuracy of the control mechanisms should increase the number of distinct responses.

Many of these theoretical considerations have been experimentally investigated in the ERK1/2 pathway, which is activated by many receptor tyrosine kinases (RTKs) and features Ras as the G-protein that activates the kinase module consisting of Raf-MEK-ERK (Fig. 1). Ras is activated by RTKs via Ras guanidine nucleotide exchange factors (RasGEFs), such as SOS, which exchange GDP for GTP. The ERK1/2 pathway regulates diverse cellular functions including proliferation, survival, migration, and differentiation. Due to overexpression or mutations of RTKs and mutations in Ras and B-Raf in many human cancers, the pathway has become a major target for cancer drug development. Although the amplifier function of the pathway has been suspected, only quantitative experimentation could confirm it. On average, a cell possesses approximately 20,000 Ras molecules, but only 0.3-5% become activated in response to 10% foetal calf serum, and only 30-50% in response to large mitogen doses. This is because only a percentage of total Ras is at the plasma membrane, where it can be activated by receptors, and then only a fraction of the correctly localised Ras will become activated. The number of ERK1/2 molecules that become activated (ppERK) in response to Ras activation, however, is much higher. Approximately 50% of the total 60,000 to 100,000 ERK1/2 molecules per cell become active in response to typical EGF doses, and even more than 50% in many instances as estimated from gel shift assays. A main reason for this signal amplification is that under many conditions, at each tier of the cascade the signal will be amplified. Thus, in many situations the sensitivity of the final output (ppERK) to the initial input (RasGTP) increases multiplicatively when a tier is added. This sensitivity amplification was demonstrated experimentally in Xenopus oocytes, where the Hill coefficient (which quantifies the steepness of a sigmoidal input/output response, i.e. the sensitivity) is greater for a three-tiered cascade than for a two-tiered cascade. The multi-site phosphorylation mechanism operating within the kinase module adds further to this sensitivity amplification, if the phosphorylations occur in a distributive manner, where each phosphorylation involves separate kinase and substrate binding and dissociation events, rather than in a processive manner, where kinase and substrate do not dissociate after the first
phosphorylation event\textsuperscript{27,28}. It has been shown \textit{in vitro} that the activation of ERK by MEK\textsuperscript{29} and deactivation of ERK by MAPK phosphatase 3 follow a distributive mechanism\textsuperscript{30}. Recent work shows that both mono- and bi-phosphorylated forms of ERK are observable after Epo stimulation in primary erythroid progenitor cells\textsuperscript{31}, implying that either ERK phosphorylation or dephosphorylation may be distributed \textit{in vivo}. However, it will be interesting to examine this issue in different cell types as the organisation of the cascade by scaffolding proteins\textsuperscript{32,33} may convert the distributive mechanism into an apparent processive one.

While the cascade structure of MAPK pathways allows for amplification of signals, it comes at an inevitable tradeoff: with higher amplification comes a shorter linear range of input doses to which the system output will respond\textsuperscript{13,34}. Also, amplification, as discussed here, does not take into account system dynamics, which are thought to be important for physiological effects elicited by MAPK signalling pathways\textsuperscript{10,11,35,36}. How do cells control MAPK dynamics and respond to large ranges of input doses? One answer to this question is by negative feedback.

\textit{Negative feedback: modulating input/output sensitivity, tuning dynamics, and conferring robustness to perturbations}

On top of the core ERK1/2 pathway lays a plethora of negative feedback loops that span multiple time scales (Fig. 2). The light blue lines in Fig. 2 denote short-term negative feedback loops, which begin to act nearly immediately upon activation of ERK. Activated ERK leads to the phosphorylation and inactivation of the RasGEF SOS\textsuperscript{37,38}, and modelling suggests that several phosphorylation sites on SOS all independently mediate strong negative feedback\textsuperscript{39}. This would afford a simple mechanism for a dosage dependent negative feedback. Gab1, a scaffolding protein involved in PI-3K and RasGEF recruitment to the plasma membrane, is also inhibited by ERK-dependent phosphorylation\textsuperscript{40}. In addition, Raf-1 is phosphorylated and inhibited by ERK\textsuperscript{1,41}, and this is one of the strongest negative feedbacks in mouse fibroblasts\textsuperscript{42}. Although underappreciated in the Ras-ERK pathway modelling literature, many studies show that activated ERK phosphorylates the EGF Receptor (EGFR/ErbB1) on T669\textsuperscript{43}, decreasing receptor internalization and substrate
specificity and mediating decreased EGFR kinase activity. The phosphorylation state of this site also negatively correlates with ERK inhibition-mediated increases of EGFR activity and plasma membrane retention. Moreover, these ERK inhibition-mediated effects on EGFR lead to increased epithelial-to-mesenchymal transition and migration. Interestingly, this threonine site and the peptide motif recognized by ERK (PLTP) is conserved on ErbB2 and ErbB4, but not on ErbB3, suggesting that ERK also feeds back to the ErbB2 and ErbB4 receptors, but not to ErbB3. Modelling reveals, however, that negative feedback to receptors can have competing effects, as receptors are responsible for the recruitment of both positive and negative effectors of ERK signalling. Thus, the classification of such feedback as strictly “negative” or “positive” is likely to be context-dependent.

Long-term negative feedback from ERK is depicted by the dark blue lines in Fig. 2. It only begins to take effect approximately 30 minutes after ERK activation, and depends on new protein synthesis. Active nuclear ERK leads to the transcription of multiple cytoplasmic and nuclear dual specificity phosphatase (DUSP) isoforms, which upon translation, dephosphorylate and deactivate ERK. In some instances, the stability and/or the phosphatase activity of these newly synthesized DUSPs is also controlled by ERK activity, resulting in an positive feedforward loop embedded into the negative feedback. The Sprouty and Spred family of proteins, which inhibit RasGEF recruitment and Raf activation, are also upregulated by Ras-ERK pathway activity. Yet another active ERK-mediated transcriptional negative feedback involves upregulation of Mig6/RALT, which not only inhibits the activity of various RTKs, but also leads to increased EGFR degradation in a manner apparently independent from the traditional ligand-stimulated pathway.

It is clear that a variety of negative feedback mechanisms have been uncovered, and likely there are more yet to be discovered. Moreover, many of these feedbacks are general, operating in many different cell lines. One obvious reason for many cell types having so many different negative feedbacks is robustness. If one or more feedbacks are compromised then others can compensate and the overall system behaviour would remain relatively unchanged. It may also be that different scaffolds allow local regulation of negative feedback, which can result in different ERK signalling patterns in different areas of the cell.
The collective effects of short-term negative feedbacks are typically very strong, with inhibition of ERK1/2 signalling causing massive upregulation of upstream components\textsuperscript{35, 57, 58}. Notably, as demonstrated by Sturm and co-workers\textsuperscript{1}, when immediate strong negative feedback is combined with the amplifier properties as described above, the ERK pathway adopts characteristics of the well-known NFA that is widely employed in engineered systems, as first theoretically predicted by Sauro and Kholodenko\textsuperscript{59} and eventually experimentally proven by Sturm et al\textsuperscript{1}. The experimental setup of Sturm and co-workers consisted of two different ERK1/2 cascade inputs: one endogenous, being stimulated by growth factors and containing the ERK dependent negative feedbacks; and the other exogenous, starting with a tamoxifen-inducible synthetic Raf-oestrogen receptor fusion protein (BXB-ER) that is not subject to negative feedback. This setup allowed analysis of ERK activation in the presence or absence of the immediate negative feedbacks. Negative feedback amplifiers as known from engineering have several properties that were reproduced experimentally using this setup. First, according to theory, when negative feedback is present, the relationship between system output and input becomes more linear, countering the effects of multiple kinase tiers and multi-site phosphorylation to amplify input/output sensitivity. Experimentally, when the ERK pathway was activated by increasing serum levels, the ppERK levels increased smoothly, in a linear fashion. However, when the ERK pathway was stimulated by increasing tamoxifen levels, ppERK levels increased sharply, with bimodal cell population responses. Thus, negative feedback makes the ERK response more linear with respect to increasing input magnitude. Second, again according to theory, negative feedback should make the system robust to perturbations within the feedback loop, but not against perturbations outside of it. Experimentally, when the ERK pathway was stimulated with EGF, ppERK levels were resistant to the effects of MEK inhibition (using the specific chemical inhibitor U0126), much more than when the pathway was stimulated with tamoxifen. These results show that the negative feedback indeed confers robustness towards perturbations to the amplifying module. However, inhibition of the EGFR (by 4557W) yielded a more linear, dose-dependent inhibition. Interestingly, this linear inhibition occurred despite the negative feedback from ERK to the EGFR. As the EGFR is outside of the amplifier module, these results suggest that the NFA effect depends on the combination of an amplifier with negative feedback. Alternatively, this might also mean that the
strongest negative feedbacks lie downstream of the EGFR, as was indeed found to be the case in mouse fibroblasts. In a more general sense these results indicate that the engineered NFA design captures the essence of the biological design of the ERK pathway. This then predicts that other salient properties of the engineered NFA should apply to the biological NFA.

A salient NFA property is that despite the reduced gain, it still will linearly amplify signals. Therefore, the ERK cascade should behave as a linear amplifier. In fact, linear increases in RasGTP after stimulation with growth factors correspond to linear increases in ppERK in different cell lines. However, the relationship between growth factor dose and ppERK is generally log-linear, with ten-fold increases in growth factor dose leading to linear increases in ppERK. One possible explanation for this is variable amplification of the growth factor signal into ppERK signals as a function of growth factor dose. Analysis of a network model for how ErbB receptors lead to activation of ERK through Ras suggested that the control of this variable amplifier strength lies with the amount of active Ras, the Raf-MEK association rate, and the MEK dephosphorylation rate. However, if the relationship between RasGTP and ppERK levels is linear, which as mentioned above has been observed in several cell lines, and is also a consequence of NFA-like behavior, then amplification within the ERK cascade is constant, not variable. This suggests that the mechanisms converting log increases in growth factor concentration to linear increases in RasGTP and ppERK would lie upstream of the Ras-ERK cascade, rather than inside the cascade as implied by model analysis. Thus, when RasGTP levels vary linearly with ppERK levels, it is unclear what gives rise to this amazing biological phenomenon that such a wide, logarithmic range of growth factor concentrations are able to induce linearly increasing yet appreciable changes in Ras-ERK cascade activation.

One potential explanation for this log-linear behavior is negative cooperativity of ligand-receptor binding. In general, one may write the following hill-type equation to relate ligand concentration, \( L \), to the levels of ligand-bound receptor (\( R^* \)),

\[
\frac{R^*}{R_{\text{max}}} = \frac{(L/L_{50})^n}{1 + (L/L_{50})^n}.
\]
Here, \( L_{50} \) is the ligand dose eliciting half-maximal receptor activation responses, \( R^*_{\text{max}} \) is the maximum level of ligand-bound (active) receptors, and \( n \) is the cooperativity coefficient (\( n<1 \) means negative cooperativity). In some cases, there is a linear relationship between the levels of active (tyrosine phosphorylated) receptor and ppERK levels\(^{37, 60}\). In such cases, \( R^* \) may be taken as a surrogate for ppERK levels. It was reported that \( n=0.31 \) for EGF binding to EGFR\(^63\), and plotting Eq. 1 with this value of \( n \) indeed gives a log-linear relationship between ligand dose and active receptor levels spanning approximately five decades (Fig. 3). Amazingly, this corresponds quite closely to the dose-response characteristics of EGF-stimulated ERK activation in several cell lines, where limit of detection occurs between \( 10^{-3} \) to \( 10^{-2} \) nM EGF stimulation, and saturation occurs around 10 nM EGF, with log-linear increases between [e.g. \(^{47, 62}\), MRB personal observations]. Thus, negative cooperativity of ligand-receptor binding may provide a simple yet effective mechanism for making cells responsive to a wide range of growth factor concentrations, which would work in synergy with the downstream negative feedback amplifier system to transduce these signals reliably. If this is the case, then one might expect similar behavior from other growth factor receptor systems. Indeed, negative cooperativity in ligand-receptor binding is observed for platelet derived growth factor type BB\(^64\), insulin\(^65\)\(^66\), and insulin-like growth factor\(^66\).

Strong negative feedback combined with a time-delay and strong stimulation can lead to oscillatory behaviour.\(^67\) The long-term negative feedbacks clearly have a significant time delay. Although many of the short-term feedbacks result from direct feedback phosphorylation from ERK and therefore have small time-delays, some are only dependent on ERK activity, and most likely are mediated by downstream kinases such as the RSKs\(^68\). Nevertheless, even the negative feedbacks involving direct ERK phosphorylation take time to travel from the substrate and through the pathway back to ERK, introducing some time-delay. A strong stimulation causes rapid and large activation of ERK, which, with a time-delay, causes strong negative feedback. This not only causes ERK activity to go back down but also the negative feedback, which then leads to increasing ERK activity, albeit not to levels as high as the original stimulation, as there is now negative feedback present. Thus, in this situation the amplitude of ERK activation in each activation/feedback cycle would be decreased.
relative to the one previous, leading to so-called “damped oscillations”. Such damped oscillations have been observed experimentally. Do such oscillations have a physiological role? This question is currently unanswered, but could lead a new line of studies about how the activation dynamics of the ERK pathway can specify cell fate decisions. If the ERK cascade exhibits bistability, delayed negative feedback mechanisms can cause sustained oscillations where ERK activity constantly switches between the low and high activity steady-states with a defined frequency. A closely related phenomenon was recently observed in human mammary epithelial cells stimulated with low EGF doses, where total ERK2 nuclear levels (as observed by lowly expressed ERK2-GFP) oscillated with a period of approximately 15 minutes after ligand stimulation. As primarily only activated ppERK is translocated to the nucleus, these results imply sustained oscillations in ppERK levels.

Another function of negative feedback is adaption, i.e. return of activity to near pre-stimulus levels despite the persistence of stimulus. This is also referred to as “transient” signalling, which is the topic of much study as whether ERK signalling is transient or sustained in response to growth factors is thought to play a major role in cell-fate decisions. In general, the stronger the negative feedback, the nearer Ras and ERK activity return to pre-stimulus levels. However, strong and fast negative feedback also reduces the peak signalling amplitude before adaption is complete, so some time delay is desirable. Since, as discussed above, too much time delay can lead to oscillations, there is a fine balance the strength and dynamics of the negative feedback and the system behaviour. Which of the negative feedback(s) are responsible for transient signalling and adaption is still largely unclear.

When ERK itself is responsible for direct negative feedback by phosphorylation, then the system will not exhibit perfect adaption, where Ras and ERK activity return exactly to pre-stimulus levels. A perfect adaptive behaviour is characteristic of an engineering design termed “integral negative feedback”, where the strength of the negative feedback is proportional to the time-integrated forward activity. Recent modelling work suggests that the long-term, transcriptional negative feedbacks might act as such integral negative feedback circuits, as mRNA responses of active ERK-dependent gene transcription are proportional to the total time active ERK spends in the nucleus. Thus, a distinguishing function of the long-term negative feedback versus the short-term negative
feedbacks may be perfect adaptation. Such feedbacks, however, can be saturated by oncogenic insults. For instance, DUSP6 expression induced by constitutively active Ras is unable to bring about adaptation of ppERK levels.\textsuperscript{73}

Part and parcel with adaption comes desensitization, or the ability of ERK pathway activity not only to return to pre-stimulus levels (adaptation), but also to reduce its sensitivity to subsequent changes in stimulus levels, despite the persistence of the stimulus. Expression of the various transcriptional negative feedback regulators, DUSPs, Mig6 and Sprouty, may help to accomplish desensitization, allowing ERK activity to respond similarly to a wide range of input strengths. This also may be controlled at the receptor level by changes in receptor trafficking and cell surface availability,\textsuperscript{74} and also, as discussed above, may be partly an inherent property of receptors with negative cooperativity in ligand binding. However, very little MAPK modelling work to date has been focused on desensitization, so it is unclear precisely how the various negative feedback mechanisms coupled with receptor trafficking and ligand binding may play a role.

\textit{Outlook for medicine: fighting against the negative feedback amplifier}

Given the multitude of strong negative feedbacks operating in many different cell types, treatment strategies focused on inhibiting entities within such feedback loops is most likely unavoidable. In fact, in breast cancer cell lines, feedback strength from the ERK1/2 pathway seems to determine whether or not the cells are susceptible to MEK inhibition.\textsuperscript{58} This is natural, however, as Sturm and coworkers have shown that the output activity of such negative feedback amplifier systems is resistant to inhibition within the system.\textsuperscript{1} By using a mathematical model of the ERK pathway, an elegant yet simplistic approach to fighting the negative feedback amplifier was derived. The partial inhibition of the negative feedback by using a Raf inhibitor weakened the negative feedback effects enough to allow a MEK inhibitor to function optimally.\textsuperscript{1} From a biological standpoint, doing such an experiment is counterintuitive. As the inhibitors target two sequential components in a linear pathway, by common biological sense they should have the same effects. However, the results of this experiment showed that biological intuition is a poor tool to grasp biological properties emerging
from design. A small amount of Raf inhibitor that insignificantly affected ppERK1/2 levels caused an order of magnitude increase in the sensitivity of ppERK1/2 levels to a MEK inhibitor.¹

Thus, applying mathematical modelling to analyse emergent properties arising from the design of signalling pathways may go a long way to explain unexpected phenomena. More importantly, the predictions originating from the models can help to design experiments probing the implications of the design structure of signalling pathways. More importantly, such predictions can explain drug failures and appropriate remedies. We envision that in the future such approaches will help to discover new drug targets rationally, and also to use our current drug repertoire in a more intelligent and effective manner.

Acknowledgements

MRB and WK acknowledge Science Foundation Ireland under Grant No. 06/CE/B1129, and MRB acknowledges a Marie Curie International Incoming Fellowship (236758) and an EMBO long-term fellowship (ALTF 815-2010).
Figure 1. The basic ERK/MAPK pathway backbone. Various external stimuli can cause changes in the levels of RasGTP (RasT), which induces activation of the three-tiered ERK1/2 cascade. Both MEK and ERK need two phosphorylations to become fully activated kinases. Doubly phosphorylated ERK (ppERK) enters the nucleus to affect gene transcription, but also has cytosolic substrates.
Figure 2. Various negative feedback loops superimposed onto the ERK/MAPK backbone. The ERK cascade backbone from Fig. 1 with short-term (light blue) and long-term (dark blue) negative feedback depicted.
Figure 3. Log-linear relationship between ligand dose and active receptors are explained by a simple negative cooperativity model. Eq. 1 is plotted here with $n=0.31$ as measured by Alvarado et al. for EGF binding to EGFR (solid black line), and with $n=1$ to give a reference for the case of no cooperativity (black dashed line). The log-linear range extends approximately 4 to 5 decades for the negative cooperativity system, and only 1 to 2 for the no cooperativity system.
References


