Title: The complexities and versatility of the RAS-to-ERK signalling system in normal and cancer cells

Article Type: SI:Signaling in Cancer Initiation & Dev.

Keywords: Systems biology; Dynamic modelling; MAPK cascade

Abstract: The intricate dynamic control and plasticity of RAS to ERK mitogenic, survival and apoptotic signalling has mystified researches for more than 30 years. Therapeutics targeting the oncogenic aberrations within this pathway often yield unsatisfactory, even undesired results, as in the case of paradoxical ERK activation in response to RAF inhibition. A direct approach of inhibiting single oncogenic proteins misses the dynamic network context governing the network signal processing. In this review, we discuss the signalling behaviour of RAS and RAF proteins in normal and in cancer cells, and the emerging systems-level properties of the RAS-to-ERK signalling network. We argue that to understand the dynamic complexities of this control system, mathematical models including mechanistic detail are required. Looking into the future, these dynamic models will build the foundation upon which more effective, rational approaches to cancer therapy will be developed.

Authors' affiliations: Systems Biology Ireland, UCD School of Medicine, University College Dublin, Belfield, Dublin 4, Ireland (all authors); Conway Institute, University College Dublin, Belfield, Dublin 4, Ireland (all authors)
The complexities and versatility of the RAS-to-ERK signalling system in normal and cancer cells

Dirk Fey1,2,*, David Matallanas1,2,3, Jens Rauch1,2,3, Oleksii S. Rukhlenko1,2, Boris N. Kholodenko1,2, *

1 Systems Biology Ireland, UCD School of Medicine, University College Dublin, Belfield, Dublin 4, Ireland
2 Conway Institute, University College Dublin, Belfield, Dublin 4, Ireland

*Correspondence
DF: dirk.fey@ucd.ie
BNK: boris.kholodenko@ucd.ie

Abstract

The intricate dynamic control and plasticity of RAS to ERK mitogenic, survival and apoptotic signalling has mystified researches for more than 30 years. Therapeutics targeting the oncogenic aberrations within this pathway often yield unsatisfactory, even undesired results, as in the case of paradoxical ERK activation in response to RAF inhibition. A direct approach of inhibiting single oncogenic proteins misses the dynamic network context governing the network signal processing. In this review, we discuss the signalling behaviour of RAS and RAF proteins in normal and in cancer cells, and the emerging systems-level properties of the RAS-to-ERK signalling network. We argue that to understand the dynamic complexities of this control system, mathematical models including mechanistic detail are required. Looking into the future, these dynamic models will build the foundation upon which more effective, rational approaches to cancer therapy will be developed.

Keywords: Systems biology; Dynamic modelling; MAPK cascade

1. Introduction

Cancer is caused by excessive and uncontrolled cell proliferation. The behaviour of healthy cells is controlled by intracellular signalling networks that process the intra- and extracellular cues. Nearly all aspects of cancer pathophysiology, including cancer initiation, development, progression and metastasis are driven by the dysregulation of one or more signalling networks[1]. The underlying causes for these dysregulations are usually genetic aberrations, such as mutations or chromosomal rearrangements. This has led to the view of cancer as a genetic disease. However, an oncogenic aberration can be suppressed in normal tissues[2]. Furthermore, different genetic aberration can sometimes result in the same phenotype, especially when they affect the same pathway. Thus,
whether a particular genetic background results in cancer depends on the context, which is determined by the dynamics of the underlying signalling networks and the environment[3, 4]. By integrating the genetic background with environmental triggers, for example, from the tumour micro-environment, the dysregulated signalling network is the ultimate driver of malignant cell behaviour and cancer.

The RAS-to-ERK pathway is a key signalling system for deciding cell fate[3, 5]. The backbone of the RAS-to-ERK network consists of the GTPase RAS on the top, followed by a three-tiered signalling cascade of kinases: RAF-MEK-ERK and their phosphatases. But in reality, many other players are involved (Fig. 1). Upon stimulation with growth factors, or other inputs, the RAS-to-ERK system can promote cell-proliferation, differentiation, or migration. Which cell-fate decision is made depends on many factors, including the relative activation strengths of the different network components and their temporal activation profiles.

Perhaps unsurprisingly, the RAS-to-ERK pathway is also one of the most frequently affected pathways in cancer[6]. Because RAS and RAF mutations are particularly prevalent, many drugs target these molecules. This direct approach to cancer therapy does not always work, as evidenced by disappointing outcomes of RAS drugs in clinical trials, the paradoxical activation of ERK in response to RAF inhibitors, and a rapid development of resistance to other targeted therapies[6, 7]. A more quantitative, dynamic understanding of the targeted networks is required, similar to Cybernetics; the “control and communication in the animal and machine”[3, 8]. Such a systems-level understanding has begun to be unravelled. Early models revealed that the MAPK cascade exhibits so-called zero-order ultrasensitivity, a systems-property that facilitates cellular decision making[9]. Coupled appropriately with negative and positive feedback, ultrasensitivity brings about complex dynamics, such as oscillations and bistability[5, 10] believed to drive proliferation and differentiation, respectively[11, 12]. Many dynamic models of ERK signalling have been developed[3, 5, 13-16], lately moving to the single-cell level and trying to explain the response variability in cell populations[17], or including thermodynamic aspects in order to explain the paradoxical activation of RAF in response to inhibitors[18]. Yet, current models still lack many biological and pathophysiological aspects of RAS signalling.

Sections 2 and 3 of this review discuss RAS and RAF biology, with a particular focus on the complexities that have been largely missed in current systems-level models. Section 4 focusses on the downstream network and describes how higher systems-level properties emerge from the dynamic ERK signalling system.

2. Revisiting RAS signalling networks: a fresh look at old and recent concepts

2.1 RAS GTPases – Background

RAS signalling has been the focus of intensive research for over three decades[19]. The mammalian RAS protein family constitutes four members: HRAS, the splice isoforms KRAS4A and KRAS4B, and NRAS[19]. All these RAS proteins are small GTPases that can switch between an inactive GDP-bound form and an active GTP-bound form, thereby conveying different upstream stimuli to a plethora of effector proteins. Structurally, RAS contains three domains that are highly conserved between the isoforms: the so called Switch I, Switch II, and effector domains; and a hypervariable region at the C-
terminal end that largely differs between isoforms. These domains, and in particular the Switch I and II domains, are essential for performing RAS’s function as a signalling switch. Moreover, the effector domain is important for downstream effector binding and signal transmission, whereas the hypervariable region is thought to regulate isoform-specific functions.

RAS plays a central role in the regulation of cell homeostasis and mediate different and frequently opposite biological functions, such as cell proliferation, differentiation or apoptosis. Importantly, RAS proteins are proto-oncogenes, and mutations of RAS genes are one of the most common events in cancer with a prevalence of over 30%. Despite intensive research over the last four decades, treatments of RAS-driven tumours have limited success[7]. This is related to important gaps in our knowledge of the RAS activity regulation, the molecular mechanisms governing RAS signalling networks and their biological functions. Systems biology models can help us understand network functions, and several models focusing on different parts of the RAS network have been developed[20]. Yet, most of these models have limited coverage of the RAS signalling network, and future models will have to include critical missing RAS effectors and mechanisms. In this section we want to highlight several concepts to be considered in future mathematical models to recapitulate the complexity of RAS signalling under physiological and pathological conditions.

2.2 Old concepts struggle with the new data revealing the complexity of RAS signalling

**Downsides of current RAS models.**

Mathematical models of signal transduction networks are abstract representations of the reality that allow us to study complex biological networks, obtaining predictions of the systems behaviours. So far, most RAS models had a very simplistic view of RAS signalling. But renewed efforts such as the NCI RAS project[21], the identification of oncogenic RAS genetic signatures[22, 23] and RAS mutant synthetic lethality screens[24] are beginning to map the complexity of RAS signalling with ever-increasing detail. Thus, the time has come for systems biologists to go beyond simple models by incorporating the up-to-date experimental and clinical information.

**RAS isoforms and specificity.**

Most mathematical models do not distinguish between the four different RAS isoforms[19]. Although these isoforms have a high degree of sequence homology and functional redundancy, they also mediate specific functions that are most likely regulated by the hypervariable region, and not fully characterised posttranslational modifications[25]. For instance, KRAS mediates a pro-apoptotic signal through the activation of RASSF-MST2 pathway or blocking BCL-XL antiapoptotic functions[26, 27]; while HRAS but not NRAS has been shown to regulate cell motility due to its ability to activate the p38-MAPK pathway[28]. Thus, we should identify the RAS isoforms involved, depending on the function we study.

**Multiple RAS activators and deactivators: GEFs and GAPs.**

Another level of complexity of RAS signalling is related to multiple inputs that affect RAS (de)activation kinetics. All RAS proteins cycle between an inactive state (bound to GDP) and an active state (bound to GTP) and each single molecule behaves as a switch[19, 29]. The transition between these states is catalysed by two family of proteins, the guanidine exchange factors (GEFs) that promote the activation of RAS and the GTPase activating proteins (GAPs) that accelerate the intrinsic RAS GTPase activity[29, 30]. The diverse members of the GEFs and GAPs families contribute to the
mechanistic complexity of RAS signalling. They are regulated by specific stimuli and can have specificity for different RAS isoforms[31-33]. For instance while SOS1, the best characterised GEF, activates all RAS isoforms, RAS-GRF1 only activates HRAS[33]. Importantly, GEFs and GAPs also determine the duration of RAS signalling, which may affect which RAS effectors become active[26, 29].

**RAS localisation.** Another observation that has been largely missed by mathematical modelling is RAS sub-localisation that varies for the different RAS isoforms[7, 25]. The three RAS isoforms are localised in the plasma membrane but they have been shown to localise in different lipid domains resulting in clusters of the different isoforms[5, 34] which may result in different interactors in each location[25]. In addition, HRAS and NRAS are activated in the endoplasmic reticulum and Golgi apparatus where KRAS4B is absent[25]. KRAS and NRAS have also been shown to localise to the mitochondria where they regulate pro- and anti-apoptotic signals, respectively[27, 33]. Hence, RAS can signal from different cellular compartments and plasma membranes domains, and the same isoform is likely to induce location-specific signalling networks.

**2.3 RAS effectors pathways, beyond the conservative view**

To date there are only four effectors downstream of RAF that are classified as bona fide RAF effectors: the RAF kinase family, PI3K, RAL-GDS and PLCε[35]. Bona fide RAS effectors are defined as proteins that i) have higher affinity for RAS-GTP; ii) bind to RAS through the effector domain (mutations within this domain disrupt the binding); iii) mediate a known biological function that depends on the interaction with active RAS[36]. Additionally, several proteins have been described as putative RAS effectors that fulfil some, but not all of the three bona fide criteria[34]. The finding that the two bona fide RAS effectors RAF and AKT mediated important biological and oncogenic functions meant that much research has focused on these canonical pathways (see Sections 2,3). In contrast, very little is known about the putative RAS effectors: Is the interaction with RAS important for their functions? And what are the network-structures mediating these functions?

There is clear evidence that over 10 putative RAS effectors mediate RAS dependent functions through direct interaction (Figure 1)[34]. For example, KRAS interacted with endogenous RASSF1A, and regulated the MST2/Hippo pro-apoptotic signal downstream of KRAS in response to oncogenic stress and EGF [26, 37]. But point mutations in the effector domain did not decrease the RASSF1A-RAS interaction, and thus RASSF1 is not classified as bona fide effector. Other observations challenge the view that the inactive Ras-GDP form does not mediate signals[36]. For example, RAS-GDP has been shown to bind and regulate the transcription factor IKZF3/Aiolos, thereby affecting expression of the anti-apoptotic protein BCL-2[38]. Considering these findings, one could expect that, similar to the bona fide effector RAF, at least some of the putative RAS effectors are also highly regulated and interconnected by crosstalk, feedforward and feedback loops[34, 35].

**2.4 Oncogenic RAS signalling networks**

When studying RAS signalling, most of the attention has focussed on the HRAS isoform as a general model for RAS signalling. Yet, different RAS isoforms and mutants activate different effectors and have different biological functions and transforming capacities[34]. For instance, KRAS and NRAS differentially affect cell-proliferation, cell-differentiation and tumour-progression in colon
cancer[39], and whereas HRAS preferentially binds PI3K, KRAS is a better activator of RAF1[40]. Further, despite HRAS being the most studied RAS family member, most of the RAS mutations in cancer actually occur in KRAS (85%) and NRAS (14%). The general model is that oncogenic mutations render RAS constitutively active, but this might not always be the case. For instance, it was shown that the activity of the KRAS-G12D mutant depends on EGF stimulation[41]. Likewise, the activity of KRAS-G12C could be reduced by inhibition of upstream GEFs[42]. Thus, it seems that these mutations do not completely lock RAF in an active state. Instead, these mutations might impair the deactivation of RAF by hampering GTPase activity, thus rendering the mutated protein hypersensitive to upstream inputs (Figure 2). It has yet to be confirmed whether this hypersensitivity is a common feature, or whether there are two different types of mutant RAS: a constitutively-activated version, and a stimulus-dependent version. In any case, the hypersensitive RAS mutation could potentially explain the differential responses to anti-EGFR therapy in RAS-mutated colorectal cancer[43].

There is evidence for crosstalk between the mutant RAS isoform and the other wildtype RAS isoforms [44]. For example, wildtype KRAS signalling was necessary to prevent the activation of the pro-apoptotic RASSF1A/MST2 pathway by mutant KRAS in colorectal cancer[26]. Similarly, wildtype KRAS and NRAS signalling was also required to prevent the activation of the G2 DNA damage response by ATR/CHK1 during mutant KRAS-induced tumourigenesis[45]. To induce the activation of the wildtype isoforms, mutant KRAS allosterically activated SOS (a major GEF activating RAS, Figure 2)[46]. Thus, it seems that mutant KRAS can elicit different oncogenic functions through both direct downstream signalling and cross-activation of KRAS and NRAS. Yet, much remains to be learned about oncogenic RAS. In particular, the relative balance between direct and indirect, and constitutively-active versus growth-factor-induced RAS signalling is intriguing and should be incorporated into future models.

3. RAF signalling: the complexity of complexes

3.1 RAF kinases – Background

Downstream of RAS, RAF kinases form the top level of the canonical three-tier MAPK cascade of RAF, MEK, and ERK. RAFs are MAPK kinase kinases (MAP3Ks) comprised of three conserved regions (CR1, CR2, and CR3) and contain an N-terminal regulatory domain and the C-terminal kinase domain (Figure 3). While RAF activation is mediated through interaction with GTP-bound RAS and translocation to the plasma membrane[26], RAFs phosphorylate and activate the MAPK kinases MEK1 and MEK2, which in turn activate the MAPKs ERK1 and ERK2.

In mammals, three RAF isoforms (ARAF, BRAF, and CRAF) activate MEK. In comparison to CRAF and ARAF, BRAF has the highest basal kinase activity, which is attributed to a motif called the N-region (negative charge regulatory region)[47]. Despite 30 years of research, so far, the only bona fide physiological RAF substrates are MEK1 and MEK2[26]. Besides their enzymatic role as kinases, RAF proteins were also shown to have MEK-independent functions including control of apoptotic pathways by crosstalk with the pro-apoptotic kinase ASK1[48], and the Rho effector kinase ROKalpha[49]. In addition, RAF1 and ARAF were shown to efficiently control the pro-apoptotic MST2/Hippo pathway by binding and sequestration of MST2 [20, 26, 50-53].
The MAPK pathway is hyperactivated in approximately 30% of human cancers commonly caused by mutations in RAS, BRAF, CRAF, or MEK1/2. Indeed, RAF proteins were the first serine/threonine kinases discovered to have oncogenic mutations with BRAF being the most frequently mutated isoform in human cancers[54]. The great majority of BRAF mutations is located in its kinase domain leading to increased kinase activity, while others are known to activate MAPK signalling through enhanced dimerisation with and transactivation of wildtype RAF[55]. By far the most prevalent oncogenic mutation of BRAF in human cancer are found at position V600, with V600E being the most common one[56].

Since the discovery of these oncogenic BRAF mutations in a large number of tumours in 2002[57] including a high prevalence of mutations (>50%) in melanoma[58] and papillary thyroid carcinoma[59], RAF kinases were put into the limelight for therapeutic intervention. Therefore, development of drugs focused on the development V600 mutation-selective BRAF small molecule inhibitors. Mutation-selective drugs such as Vemurafenib/PLX4032[60] and Dabrafenib/GSK2118436[61] initially show impressive clinical responses in the treatment of melanoma patients. However at moderate doses, instead of MAPK inhibition, these inhibitors paradoxically lead to the activation of the MAPK pathway in a RAS-dependent manner by promoting homo- and/or heteromerisation between various RAF isoforms, specifically in wildtype RAF cells[62, 63].

Next, we describe novel aspects of RAF activation, crosstalks and drug-resistance mechanisms, highlighting gaps in our current understanding and scientific challenges ahead.

3.2. “Zooming-in” on RAFs – Activation processes

**RAF dimerisation**

For a long time, RAF proteins were considered as monomeric kinases that activate MEK. The discovery that under physiological conditions RAF kinases form homo- and heterodimers to drastically enhance the kinase activity[64] changed our view of MAPK signalling (reviewed in[56, 65]). Dimerisation is a feature of many other kinases, such as MEK[66], ERK[67], and JAK[68]. ATP-competitive inhibitors can facilitate kinase dimerisation, where the protomers acquire different drug affinities. The emergence of different drug affinities between monomers and protomers in a dimer has been enigmatic, but is explained by thermodynamics[18].

RAF dimerisation is a highly regulated process by RAS and 14-3-3 proteins allowing for the fine-tuning of signalling strength within the modular MAPK pathway. Several BRAF inhibitors enhance BRAF dimerisation in a RAS-dependent manner thereby activating MAPK signalling[62]. Novel inhibitors are being developed to evade this phenomenon[69, 70]. RAF dimers are functionally asymmetrical [71] with one RAF kinase functioning as an activator to stimulate activity of the second RAF kinase (receiver, Figure 3). Importantly, the activating RAF does not require kinase activity but an N-terminal phosphorylation that functions allosterically to induce cis-autophosphorylation of the receiving RAF. Structural studies of dimeric[72] and monomeric[73] RAF corroborated these findings and demonstrate that RAF proteins dimerize in a side-to-side conformation. Based on these structural data, molecular dynamic simulations revealed that phosphorylations of the so-called N-terminal acidic (NtA) motif facilitate RAF dimerisation by introducing crucial salt bridges, thus
providing a structure-based mechanism for the auto-transactivation and asymmetry of RAF dimers[74, 75].

Activation of MEK requires RAF binding, at least transiently, to allow for phosphorylation. While this was known for some time, only recently, the crystal structures of BRAF–MEK[72] and KSR2–MEK[76] were solved, thus providing novel insights: Both RAF and KSR form heterotetramers with MEK, where either BRAF or KSR2 back-to-back homodimers associate with two MEK1 molecules. While these structural data provide a better understanding on RAF function, available crystal structures are limited to the kinase domain lacking the N-terminal part of RAF. The full-length structure is still unsolved, but would provide crucial information for designing novel drug interventions such as interface inhibitors.

While most of the initial work has focused on RAF1 and BRAF (after the discovery of its oncogenic mutations), relatively little is known about the third RAF family member ARAF. With its low kinase activity towards MEK, a role outside the canonical MAPK pathway by sequestering the pro-apoptotic kinase MST2 is in line with our current understanding[77]. But recent data have again implicated ARAF in promoting MAPK activity. Knockdown of ARAF prevents MEK activation and cell migration in a cell-type dependent manner, and importantly, dimerisation seems to be required for ARAF kinase activity[78].

Modifications
RAF proteins are highly regulated by posttranslational modifications such as phosphorylations. While a large number of activating, inhibiting, and feedback phosphorylations of RAFs have been characterised for some time, the exact dynamics and processes are still not fully understood. These phosphorylations are highly dynamic and trigger conformational changes within the molecule thereby enabling the RAF activation cycle where an inactive RAF switches to an enzymatically active state and returns to basal state. Intrinsically, in addition to the intramolecular changes, dynamic phosphorylations also allow for differential binding of proteins leading to distinct RAF complexes. Due to the short half-life and technical difficulties with isolation, the exact composition of these RAF complexes is still enigmatic. Recently, the combination of mass-spectrometry based proteomics with genetic approaches allowed to map cell-context specific BRAF phosphorylation and ubiquitination sites as well as the cell-context specific BRAF interactomes[79].

In summary, recent findings change our view of RAF by revealing the dynamic complexities of RAF-dimer formation, intramolecular conformational changes and posttranslational modifications. Despite more than 30 years of research, there are still gaps in our knowledge of the RAF activation cycle, which includes highly dynamic phosphorylation steps, feedback mechanisms, and orchestrated protein interactions[65]. In addition, the majority of the published data is based on the idea of RAF as a monomer. Updated, “dimeric” versions of this knowledge are necessary.

3.3. “Zooming-out” from RAFs – Neighbours and effectors
The MAPK pathway is not a linear signalling pathway, but has branching points allowing crosstalk with other pathways. A number of scaffold proteins including KSR, IQGAP, Arrestin, and Paxillin organize RAF and MAPK functions, serving as binding platforms and controlling the spatial and temporal aspects of the signalling flux (reviewed in[80, 81]). While these scaffolds were investigated in parallel to RAF over the years, the dynamic complexes they generate remain enigmatic. It is not
clear how an upstream signal is propagated through the maze of scaffold-mediated signalling complexes. Scaffolds are notoriously difficult to work with because both overexpression and knockdown can lead to the disruption of the signalling complex, a phenomenon known as combinatorial inhibition[80]. Eliminating the side-effects of classical overexpression systems, modern approaches including Crispr/Cas for “endogenous tagging” will offer the solution to isolate, visualize, and investigate scaffold complexes under physiological conditions. One of the novel scaffolds suggested for BRAF and CRAF function was a RAF family member itself. ARAF was shown to scaffold and stabilize CRAF-BRAF cocomplexes in RAF-inhibited cells to ensure efficient signalling[82]. While this observation is in line with the idea that many ARAF functions are kinase-independent, these data might also suggest that RAfs not only form dimers, but trimeric or higher order complexes. How this is accomplished from a structural point of view is unclear so far.

While RAF kinases were primarily considered as activators of MEK/ERK signalling, more and more evidence suggested a direct involvement in other cellular processes. In addition to its anti-apoptotic role in the MST2 pathway, ARAF was also identified as a binding partner and inhibitor of the glycolytic enzyme pyruvate kinase M2 (PKM2)[83]. The ARAF function and localisation are organised by the scaffold KSR2[53] which has strong links to AMPK and energy metabolism. AMPK also regulates the BRAF-KSR complex through phosphorylation[84], suggesting an interesting cross-talk between apoptosis, energy homeostasis, MAPK signalling, and glycolysis. How this complex function is regulated, is unclear but its presence strongly suggests that RAF kinases are organised into distinct complexes within a cell. Kang and colleagues recently demonstrated a “synthetic lethal” interaction between oncogenic BRAF and the ketogenic enzyme 3-hydroxy-3-methylglutaryl-CoA lyase (HMGCL), suggesting RAF control and feedback to ketogenesis[85]. In cancer cells, BRAFV600E upregulates the expression of HMGCL and increased acetoacetate production, which, in turn, selectively enhances the BRAFV600E-MEK complex thereby promoting activation of MAPK signalling. While not direct, this feedback mechanism nicely demonstrates how signalling networks are embedded within their cellular functions.

3.4 RAF inhibitors and resistance mechanisms

While initially, RAF inhibition can be a very effective treatment, for example for BRAFV600E driven malignant melanoma, acquired resistance commonly develops within 6-9 months. Many resistance mechanisms overcoming the blockage of RAF and reactivating the ERK signalling have been identified (recently reviewed in [86]), including increased upstream RTK activity [87]; acquisition of secondary mutations of NRAS[88] and MEK1[89]; the expression of an alternative MEK activator[90]; and the activation of YAP/TAZ[91, 92]. A shortened BRAFV600E splice variant that causing enhanced RAF dimerisation was also identified [93].

Rather than focusing on directly suppressing RAF kinase activity, targeting RAF’s protein-protein interaction by small-molecule-inhibitors might be a valuable option. Recently, Athuluri-Divakar and colleagues reported a novel inhibitor, Rigosertib, which binds to the RAS-binding domain of RAS, thereby interfering with the RAS-RAF interaction and supressing the ERK signalling [94]. But while we currently have a good idea about the short-term RAF signalling and the impact of RAF inhibitors, the gap remains of how acquired resistance to RAF inhibition occurs within 6-9 months and how these processes are linked. Clues might be found in the tumour microenvironment, where BRAF inhibition
was shown to generate a fibroblast-derived niche which mediates therapeutic escape[95] through integrin-mediated signalling[96].

Although the great majority of scientific data for RAF focused on its oncogenic functions, RAF mutations might not always cause cancer. Interestingly, it was discovered that normal human skin also harbours a high percentage of somatic BRAF mutations (among other driver mutations)[2], suggesting that the normal cellular functions are maintained in the presence of these cancer-promoting mutations. It is currently unclear if and how the normal skin suppresses the oncogenic functions of these mutations, or which other factors have to contribute for the development of skin cancer. It seems that the RAF signalling output is more important than the mere presence of a RAF mutation.

4. The ERK signalling system: computing cell-fate

The most studied biological readout of the RAS/RAF signalling system is ERK activity. But the upstream signal through RAS, RAF and MEK is not only a mere activator of ERK. Rather, ERK activity and ERK function are tightly regulated by fine-tuned signalling mechanism involving different feedforward structures, pathway crosstalks and feedback loops, expertly reviewed elsewhere [3, 5, 13]. These loops act on different timescales thus fine-tuning the activation and de-activation dynamics of ERK. The biochemical mechanisms of this regulation predominantly involve phosphorylation and dephosphorylation, but also other posttranslational mechanisms [34, 97]. Through these modifications many protein functions such as kinase activity, binding affinities (specificity to upstream activators, downstream effectors, scaffolds etc.), target specificity, subcellular location and protein stability are regulated.

4.1. Dynamic networks and their feedback regulation

A prototypical example of how different feedback structures bring about different ERK activation dynamics resulting in different cell-fate decisions comes from rat adrenal gland pheochromocytoma cells (PC12), a convenient cell system for studying neuronal differentiation [12, 17, 98]. Stimulation with EGF triggers a sharp peak of ERK activity that quickly decays on the timescale of 20 to 30 minutes, and causes the cells to proliferate. In contrast, NGF treatment causes prolonged ERK activity that is sustained for many hours, and triggers neuronal differentiation. It was shown that EGF and NGF employ negative and positive feedback loops to regulate the transient and sustained responses, respectively[12]. These findings suggested an appealing picture that EGF-induced negative feedback causes transient responses, whereas NGF-induced positive feedback causes sustained responses. This picture is generally valid, confirmed in other cell-types[99], and extended to other, related MAPKs[4, 100, 101]. However, this simple picture does not explain where the different feedback wiring comes from. Here, the underlying biology is more complicated.

4.2. Feedback wiring put into the network context

In particular, the ERK system dynamics are not regulated by a single gene or protein acting as master controller. Rather, the control over the ERK response is distributed within the network, and influenced by crosstalk from parallel pathways such as PLC and PI3K-AKT signalling[98, 99]. In particular, the balance between ERK and AKT activation seems to be important for several reasons.
Firstly, AKT signalling regulates the cytoplasmic-nuclear localisation of ERK via PEA15, thereby influencing the relative strength by which the positive and negative feedback are employed[98]. Secondly, the cell-fate outcome depends on where the relative activation of ERK and AKT falls on a two-dimensional ERK-AKT response map. A curved boundary between ERK and AKT separates differentiating (high ERK, low AKT) from proliferating (high ERK, high AKT) PC12 cells[102], whereby a negative crosstalk from PI3K to RAS keeps the cells close to the proliferation-differentiation boundary.

Although this PI3K-to-RAS crosstalk is sometimes labelled as negative feedback loop, describing it as crosstalk would be more accurate. A real feedback requires a reciprocal effect from RAS to PI3K and back to RAS that closes the loop. Many such closed loops involving ERK and its upstream activators or inhibitors have been described[5, 13, 97]. Interestingly, recent evidence suggests that crosstalk might actually regulate the feedback wiring or strength. For example, Chen[102] described a feedback loop from RAS to ERK and from ERK via regulation of RASA2 expression back to RAS. Because the RAS-GAP activity of RASA2 depends on PI3K signalling[102], it is likely that the strength of this ERK negative feedback is regulated by PI3K. Further support comes from a mathematical model that could recapitulate the observed single cell ERK dynamics only when a feedforward regulation of the feedback strength was included[17]. Although both studies concerned PC12 cell fate in response to EGF and NGF, the timescales of the feedforward crosstalks do not match. The RASA2 loop acts slowly (12h), and can therefore not explain the fast-acting loop (10 min) inferred in [17]. Although the fast-acting loop regulating the feedback gain remains biochemically uncharacterised, this loop could be mediated by one of the many existing crosstalks. But how exactly are the actions of the different feedforward and feedback structures coordinated, and how do they give rise to higher systems-level functions?

4.3. The emergence of systems-level functions

Engineering tells us that systems-level functions such as feedback-amplification and noise-rejection are important for the proper processing and transmission of signals (Cybernetics, [8]). For example, negative feedback is critical for steering a system into a desired state, damping unwanted perturbations, and maintaining the desired state in the presence of noise. Perhaps unsurprisingly, the ERK network confers such systems-level functions (Figure 4). For these systems-level functions, the static network structure is only one ingredient. Much more important are the dynamic responses to perturbations and how they play out over time on the network level. Ultimately, these network dynamics control the system’s high-level functions and signal-processing properties.

System dynamics

Complex dynamic behaviour for instance, arises from the combination of ultrasensitivity with different types of feedback loops, reviewed in[5, 13]. Briefly, ultrasensitivity refers to a sigmoidal shape of the systems input-output relationship (“S” shape of the dose response)[9]. When coupled with negative feedback, ultrasensitivity can give rise to excitable behaviour or sustained oscillations, whereas the combination with positive feedback promotes bistability[5, 10].

Perfect adaptation

Another high-level system property is perfect adaptation, which is an ERK activity response that peaks and returns back to the exact pre-perturbation baseline even in if the perturbation persists.
Perfect adaptation can be achieved by an incoherent feed-forward loop with the correct parameter values\[103\]. A more robust way to achieve perfect adaptation independently of the exact parameter values is by using integral feedback\[104, 105\]. Integral feedback means feeding-back the accumulated response, which is the area under the response curve, as opposed to the acute (instantaneous) response value. Integral feedback was for example observed for CFOS induction\[99\]. Interestingly, the nuclear ERK response and many immediate early genes also exhibit perfect adaptation, but whether integral feedback plays a role has not been confirmed\[97\].

**Negative feedback amplifier.**

It was shown that the ERK systems exhibits properties of a negative feedback amplifier (NFA) due to a strong negative feedback loop from ERK to RAF\[106\]. This negative loop (i) speeds up the response time, (ii) converts intrinsic switch-like activation kinetics into graded linear responses, (iii) conveys robustness to changes in rates of reactions within the NFA module, (iv) stabilises outputs in response to drug-induced perturbations, and also has therapeutic implications (Figure 4). It was shown that the ERK negative feedback diminishes the effectiveness of RAF and MEK inhibitors \[107\], but also that the drug-effect is quite different depending on the inhibitor’s mechanism of action\[108\]. In negative feedback configurations, competitive inhibitors tend to be more effective compared to gene knockdowns\[108\]. The reverse is true for positive feedback configurations, where non-competitive allosteric inhibitors and knockdowns tend to be more effective.

**Robustness.**

Robustness is a systems property that refers to the ability of a system to maintain its function in the presence of noise and other undesired perturbations\[105\]. In linear systems the overall robustness is conserved, because the overall sensitivity can only be shifted around by making specific components less sensitive at the expense of others. In contrast, nonlinear systems such as the ERK system can generate robustness, albeit this also comes at the cost of sensitivity. A good example is the NGF induced bistable switch, which generates a discrete and robust yes-or-no answer at the expense of quantitative sensing.

Robustness is an inherently dynamic property that comes in many flavours and definitions\[105\]. For example, robustness can refer to the filtering-out noise on fast timescales, or the rejection of parameter changes on the slow timescales. Filtering-out noise can be achieved by distributing the control over several loops. A good example, are the two coherent feedforward loops from ERK to CFOS (on the transcriptional and phospho-protein level) which filter out time-dependent growth-factor fluctuations in a frequency dependent manor\[99\]. Fast variations, presumably constituting noise, are rejected, whereas slow variations, presumably constituting the signal, are transmitted. On the other end of the dynamic spectrum, parameter-robustness arises from the strong negative feedback from ERK to RAF discussed earlier. The negative feedback amplification quickly compensates for slow (or static) changes in the total ERK concentration\[106\]. Thus, the combination of different feedforward and feedback loops make the ERK system response remarkably robust to both fast-acting external noise and slow-acting internal parameter changes.

**4.4. Reading out ERK activity: which dynamic features matter?**

While ERK can encode a variety of dynamic behaviours and systems-level functions, it is as of yet mostly unclear which dynamic features actually matter: Which aspects of ERK activity do
downstream targets sense to elicit their biological responses? Amplitude, duration and integral of the ERK response have all been correlated with cell-fates in different contexts (Figure 5)[109]. The probably best controlled study used genetically engineered light-activatable C-RAF in rat kidney epithelial cells to specifically manipulate ERK activity in a pulsatile versus sustained fashion. Sustained ERK upregulated 100 genes, whereas pulsed ERK upregulated only 10 genes[110]. Five of the ten genes did not respond to sustained ERK, making them specific sensors of pulsatile ERK activity. But so far, there have been few systematic studies comparing a variety of several growth factors and cell-types, and those who did either focussed on the steady-state, thus neglecting the dynamic aspects of the signalling response[111, 112]; or found that the associations between dynamic signalling features and phenotypes were complex and uncertain[113]. A notable exception is [99], which showed that the characteristic decay time of the acute ERK response is the best predictor for CFOS expression. This holds true in both MCF7 and PC12 cells for a variety of growth factors and dosages, giving some validity to the general idea that sustained ERK signalling promotes differentiation, whereas a transient peak of activity promotes proliferation. A critical limitation of this study is the exclusive use of bulk measurements, which limits the conclusions that can be drawn concerning the cell-fate of individual cells.

4.5. Robust single-cell decisions in noisy, heterogeneous populations.

There is remarkable heterogeneity on the cell-to-cell level that impacts cell fate decisions[11, 17, 102]. One of the best examples that ERK dynamics matter on the single cell level comes from MCF10A cells[11]. These cells exhibit spontaneous repetitive pulses of ERK activity, each lasting for about 20 minutes. The frequency and duration of these pulses is stimulus-dependent and correlates with the probability and timing of S-phase entry. Considering that the amplitude of these ERK pulses is constant, the same correlation could be achieved with an integral ERK readout. Indeed, such integral readouts can be achieved by transcriptional targets of ERK that are slowly induced and relatively stable, such as DUSPs[99]. Another example comes again from PC12 cells, which can be tricked into differentiating by applying multiple, repetitive input pulses of EGF, a growth factor that normally leads to proliferation[17]. Whilst even repetitive EGF stimulation cannot trigger a sustained ERK response, pulsing at the optimal frequency maximises the integral under the ERK response curve, thus suggesting that the cumulative ERK response triggers differentiation.

At a first glance, it seems difficult to reconcile these two observations into a coherent picture. How can the same dynamic signal (repetitive ERK pulses) cause two mutually-exclusive cell-fates (proliferation and differentiation)? Clearly, the interpretation of the ERK signal by downstream targets is different[114]. Apart from the typical reference to cell-type specificity, one way to provide context is by regulating the localisation of the ERK signal. Nuclear and cytosolic ERK responses differ[98, 115, 116]. Other possible explanations are timing (homeostasis vs acute stimulation) and the activation of additional parallel pathways (e.g. AKT, PLCγ, Figure 1). Activation of these parallel pathways might over time bring the entire system into a new homeostatic state in which the ERK downstream network responds differently to the same signal.

4.6. Population context

The population context also provides critical inputs to the ERK system[110, 117]. It was shown that the output of the ERK signal depends on the cell density in rat kidney epithelial cells that exhibited...
random ERK pulses[110]. Here, the ERK signal propagated from one cell to the next, thus increasing the likelihood of an ERK pulse in the receiving cell. Another study showed complex interactions between mutated tumour cells and un-mutated stroma cells that occurred in a reciprocal fashion[117]. The KRAS G12D mutation in pancreatic ductal adenocarcinoma (PDA cells) signalled to adjacent heterotypic fibroblasts, which in turn engaged reciprocal signalling in the PDA cells. This signalling loop via fibroblast doubled the number of KRAS-G12D dependent signalling nodes in the PDA tumour cells, thereby affecting tumour cell proliferation and apoptosis. Although confined to 2D cell cultures, these studies provide nice examples demonstrating that deciphering the cell-context can lead to important insights. A lot of work remains (in both 2D and 3D culturing systems), especially considering that tumours consist of a heterogeneous, organised mix of cells embedded in complex tumour microenvironments.

4.7. The clinical relevance of intra-cellular dynamic fine-tuning

The environmental noise, population heterogeneity and system-level robustness of the cellular signalling raises questions to what extent the exact quantitative behaviour matters. Does the ultimate fate of a cancer patient really depend on intracellular fine-tuning? Although little is known with regards to ERK signalling, parallels may be drawn from another, related MAPK signalling network; the JNK system. Computational modelling has revealed that the nonlinearity and feedback in this pathway can amplify small, individual changes in gene expression to cause markedly different responses on the systems-level[4]. Applying this model to tumour data from Neuroblastoma patients yielded patient-specific simulations that correlated with tumour stage, amplification of the MYCN oncogene, and patient survival (three independent retrospective cohorts, n=709 patients). To generate these patient-specific simulations, the measured gene expression values for all five model components were used as static parameters in the model. These personalised models allowed predictions of the dynamic JNK response to cellular-stress (chemotherapy) over time. Crucially, dynamic properties – in particular the switch-likeness of the response – correlated best with patient survival. The work provides a rationale for modelling the ability to activate the pathway, as opposed to directly predicting its activity at the point of diagnosis or surgery per se. It seems that the ability to activate the pathway is more important, because it can actually predict what might happen later on when the patient is exposed to drugs or increased cellular stress from the tumour microenvironment. But prospective validation is still outstanding, and it is currently unknown whether these findings translate to other cancers and signalling networks, such as the ERK pathway.

5. Concluding remarks

The time has come to address the bewildering complexity of RAS signalling. Inspired by Cybernetics, we should decipher how the RAS system is controlled, thus gaining a quantitative, dynamic understanding that can help us to develop rational approaches to therapy. Such a systems-level understanding in terms of mathematical models is beginning to be unravelled[3]. The RAF/MEK/ERK signalling system is best studied, with a whole range of dynamic models[5, 15-17]. This is why we know most about the dynamic complexity of ERK signalling. But, most of the biological detail discussed in Sections 1 and 2 has yet to be incorporated into these models. One aspect is increasing the resolution of the models using “zoomed-in” details such as RAS isoform-specificity and BRAF complex-formations. Another is widening the scope by “zooming out” and, for example, modelling
the many non-bona-fide RAS pathways. Such modelling should help deciphering the mechanistic details that constitute the “context” underlying seemingly unexplainable response differences. Dynamic modelling can provide clinically useful information, as was recently demonstrated in a proof-of-principle study, where simulating the precise activation dynamics of the JNK pathway was prognostic of neuroblastoma patient survival\[4\]. Emerging evidence that signalling dynamics regulate cell fate decisions calls for a mechanistic, model-guided understanding of the complexity of signalling networks. This attractive goal of fundamental research will have major clinical applications.

Acknowledgements

This work was supported by the European Union FP7 grant “SynSignal” (No. 613879) and Science Foundation Ireland under grant No. 14/IA/2395. We apologize to authors whose contributions could not be included because of space constraints.

Figure 1. The RAS signalling network. Simplified scheme neglecting many intermediates, crosstalks, and feedbacks. Red – bona fide RAS effectors, blue – putative RAS effectors.
**Figure 2. Dynamic effects of RAS mutations.**

\[
\frac{dx}{dt} = k_1 \cdot GEF \cdot \frac{1 - x}{K_1 + 1 - x} - k_2^{WT/Mut} \cdot \frac{x}{K_2 + x}
\]

\[k_1 = 1, k_2 = 0.25, K_1 = 0.1, K_2 = 0.1, \alpha = 0.1\]

**Cross activation of WT RAS by mutant RAS**

\[
\frac{dx}{dt} = k_1 \cdot SOS \cdot \frac{1 - x}{K_1 + 1 - x} - k_2 \cdot \frac{x}{K_2 + x}
\]

\[SOS = \text{input} + \alpha RAS^{Mut}, \alpha = 0.2\]

**Figure 3. RAF kinase family, dimerization mechanisms, inhibitor-induced resistance, and selected signalling crosstalks.**
Figure 4. Higher level properties arising from the ERK negative feedback amplifier.

\[
\frac{d}{dt}p\text{ERK} = k_{1,\text{input}} \frac{1 - p\text{ERK}}{K_1 + 1 - p\text{ERK}} \gamma(p\text{ERK}) - k_2 \frac{p\text{ERK}}{K_2 + p\text{ERK}}
\]

\[
\gamma(p\text{ERK}) = \frac{K_1}{K_1 + p\text{ERK}}
\]

input = 0.64, \( k_1 = \begin{cases} 0.4 & \text{without feedback} \\ 1 & \text{NFA} \end{cases} \), \( K_1 = 0.1, k_2 = 0.1, K_2 = 0.1 \), \( f = \begin{cases} 100 & \text{without feedback} \\ 0.5 & \text{NFA} \end{cases} \)

Figure 5. Single-cell ERK dynamics affect cell-fate. (A) The frequency and duration of ERK pulses determines the occurrence and timing of cell-cycle entry in MCF10A cells; The duration of the ERK response in PC12 cells differs between EGF and NGF stimulation that cause proliferation and neuronal differentiation, respectively; The characteristic decay rate of the ERK response to either EGF, NGF or HRG correlates with CFOS expression. (B) Regulation of positive and negative feedback strength (gain) controls the ERK-response dynamics (transient versus sustained responses), ERK-response heterogeneity between cells in a population, and PC12-cell differentiation.
References


