Transcription Factors Synergistically Activated at the Crossing of the Restriction Point between G1 and S Cell Cycle Phases. Pathologic Gate Opening during Multi-Hit Malignant Transformation

Nicoletta Castagnino, Massimo Maffei, Lorenzo Tortolina, Gabriele Zoppoli, Daniela Piras, Alessio Nencioni, Alberto Ballestrero, Franco Patrone, and Silvio Parodi

Department of Internal Medicine and Medical Specializations (DIMI), University of Genoa, Italy; IRCCS Azienda Ospedaliera Universitaria San Martino - IST, Genoa, Italy

Corresponding Author: Silvio Parodi; email: silvio.parodi@unige.it

Received 3 March 2016; Accepted 12 September 2016

Editor: Francesca Spyrakis

Copyright © 2016 Nicoletta Castagnino et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract. Transcription factors (TFs) represent key regulators of gene-expression patterns controlling cell behavior. TFs are active at nuclear – chromatin levels. TFs do not act in isolation; small sets of TFs cooperate toward the transcription of sets of mRNAs and consequently the translation of new proteins (the molecular phenotypes of a cell). Most TFs are activated through a cascade of biochemical reactions mediated by receptors expressed on the target cell surface. Nuclear Receptors (NRs) are transcription factors activated instead by small hydrophobic molecules capable of crossing the plasma membrane. The convergence of different pathways on TFs and their posttranslational modifications ensure that the external stimuli generate appropriate and integrated responses. The reconstruction of the molecular anatomy of these pathways through Molecular Interactions Maps (MIMs) can depict these intricate interactions. A mathematical modeling approach simulates/mimics their mechanism of action in normal and pathological conditions. We can simulate the effect of virtual hits in neoplastic transformation as mutations/alterations in these pathways. We can also simulate the effect of targeted inhibitors on these deregulated pathways. This strategy can help to guide an appropriate combination of targeted drugs in the treatment of a cancer patient, a major innovative perspective of incoming years.

Keywords: Transcription Factors; Nuclear Receptors; Signaling Networks; mathematical modeling; targeted inhibitors

1. Introduction

A cell recognizes and responds appropriately to diverse internal and external stimuli. The initiating events include activation of receptors, followed by an elaborate cascade of signals, interconnected in large regulatory networks, leading to a change in gene expression. These cascades of signals travel on interconnected signaling networks (network of biochemical interactions) which relate extracellular signals to the expression of specific genes [1, 2]. Transcription factors (TFs) represent key cellular components that control gene expression: their activities determine how cells function and respond to the environment. Their role is central in controlling many biological processes, from cell cycle progression
and maintenance of intracellular metabolic and physiological balances, to cellular differentiation and development [3]. TFs contribute to switch on and off genes at the correct time and in the correct place, in different cell types and conditions.

In this review, we describe different classes of TFs and their mechanism of activation. Our attention was focused on the role of biochemical interactions on the activation/repression of some TF activities, in normal and pathological conditions (especially in cancer).

Reductionist studies focused on each component of signaling networks concurred to better understand the effects of each component on TF functions and responses to growth factors/external stimuli. However, the role of each component cannot be understood considering the behavior of one or few molecules: new properties emerge from the interaction of multiple components. We need to define the network of "protein - protein interactions" to which a specific (signaling) protein belongs.

We applied an integrated approach. We organized available information about these networks of interactions, containing multi-protein complexes, protein modifications, and enzymes that are substrates of other enzymes through Molecular Interaction Maps (MIMs). This graphical notation supports the reconstruction of explicit models, subsequently amenable to dynamic simulations. We can introduce in these models, mutations/alterations in specific components (proteins) of these altered signaling networks and virtual inhibitors of these alterations. Each individualized model can represent a cancer patient, characterized by specific alterations in defined genes. This approach can suggest reasoned combinations of targeted inhibitors against specific altered pathways, according to a molecular characterization of each cancer patient.

2. Transcription Factors

Transcription factors (TFs) are present in all organisms and their number increases with genome size: larger genomes tend to have a higher fraction of TFs among their genes. In humans, TFs represent the largest family of proteins, accounting for around 10% of all genes [4].

TFs regulate gene expression acting as DNA sequence-specific binding factors, orchestrating cofactor recruitment and assembly of the transcriptional machinery. They bind to specific short DNA sequences (the promoter and enhancer sequences, named also TFBS – Transcription Factor Binding Sites). The TFBS:TFs complexes regulate transcription, either positively or negatively. TFs can have multiple DNA-binding domains that attach to specific sequences of DNA, while other DNA regions are responsible for stimulatory or inhibitory effects on transcription [5].

TFs regulate many genes, and each transcribed gene is in turn regulated by many TFs.

Two types of TFs have been identified: general and sequence-specific. General transcription factors are involved in transcription from all polymerase II promoters and therefore constitute part of the basic transcription machinery. General TFs are involved in the formation of the pre-initiation complex and act cooperatively with RNA polymerase II. These protein factors have been named (Transcription Factor - TF) TFIIA, TFIB, TFIID, TFIIE, TFIIF, and TFIIH: they direct RNA polymerase II to its promoters and help it to initiate transcription [1, 2, 6].

Transcription requires both general and sequence-specific TFs. Sequence specific TFs bind specific regulatory DNA elements (TFBS) at varying distances away from the polymerase II transcription start site, leading to distinct spatiotemporal patterns of gene expression [4]. The occupancy of transcription activators on a gene promoter (TFBS) recruits the transcription initiation machinery, which consists of RNA polymerases and more than 50 other components [1, 2, 6]. This is followed by promoter clearance, elongation, and the termination of transcription.

Not only genes coding for proteins, but microRNAs and non-small RNAs can also be regulated by TFs [7]. MicroRNAs and non-small RNAs will in turn cooperate to the regulation of mRNAs transcription and protein translation.

3. TFs: Mechanism of Action

Many different kinds of signaling-proteins can transmit information between the cells (of multicellular organisms) and their environment. These molecules act as ligands that bind a specific receptor in their target cells. Signal transmitters present a considerable variation in their structure, in their function and in their mode of action in a specific target cell. Some signaling molecules are able to cross the plasma membrane and bind to intracellular receptors in the cytoplasm or nucleus, whereas most bind to receptors expressed on the target cell surface.

TFs can be activated by two different mechanisms, depending from the ligand.

In most cases, the signal (proteins, small peptides, amino acids, and so on) is transduced from the cell membrane into the nucleus through the activation of trans-membrane ligand-binding receptors on the cell surface. Chains of biochemical interactions/reactions at the level of a complex cytoplasmic signaling-network are generated. Let consider the SMAD transcriptional complexes as an example of this class of TFs. The acronym SMAD was coined in reference to identification of human SMAD1 and its sequence similarity to Sma (the orthologs in C. elegans) and Mad (the Mad protein in Drosophila, which was the first identified member of this family) proteins. They are a component downstream of the Transforming Growth Factor - TGF-β (a ligand) pathway (Fig. 1). Briefly, upon TGF-β stimulation, with some intervening transmission of the signal, SMAD4 complexes with phosphorylated SMAD2 or SMAD3, can
bind to a SMAD binding site (TFBS) on MYC (v-Myc avian Myelocytomatosis Viral Oncogene Homolog: Myc gene), CCND1 (Cyclin D1 gene) and other genes, repressing their transcription [8–10]. Other signaling pathways further regulate SMAD activation and function, allowing versatility and diversification of TGF-β growth factor responses [11].

In other cases, the receptors are inside the target cell. A hydrophobic ligand (see Fig. 3), instead of binding with receptors at the plasma membrane can be able of crossing the plasma membrane, and of activating transcription starting with an inactive TF bound to cytoplasmic or nuclear proteins. The signal molecules must be sufficiently small and hydrophobic to diffuse across the plasma membrane [1, 2]. This direct interaction with TF inside the cell (or nuclear receptors - NRs) is mediated by lipophilic hormones, including steroids like testosterone, estrogen, progesterone and the corticosteroids, retinoids (vitamin A), thyroid hormones, and steroids of the Vitamin D family.

The NRs superfamily comprises a large class of TF, with very different modes of action. Members of this superfamily include not only receptors for known ligands as the classic endocrine receptors but also a large number of so-called orphan nuclear receptors, a set of proteins sharing significant sequence homology to known nuclear receptors, but for which the ligands or activators, target genes and function, were initially unknown or have not yet been identified [13, 14]. Most orphan nuclear receptors (such as retinoic X receptors - RXRs, peroxisome proliferator-activated receptors - PPARs and liver X receptors - LXR) have been ‘adopted’ through the identification of their ligands [15–17]. The event of ligand - TF binding generates a different structural conformation in this type of TFs - NRs. The molecule will lose affinity for the anchoring cytoplasmic/nuclear proteins. It will become finally capable of binding to specific transcription factor binding sites (TFBS) at a DNA/chromatin level, resulting in up or down regulation of tissue-specific sets of target genes, often related to differentiation events.

Multiple signaling pathways converge at the level of TFs, modifying these proteins with an array of post-translational modifications (phosphorylation, sumoylation, ubiquitination, acetylation, glycosylation) that are often interconnected in time, space, and combinatorial function. TFs can function as coincidence detectors in which two or more pathways must be activated before gene transcription is altered. Multiple phosphorylation sites represent points of convergence of signaling pathways initiating at the plasma membrane [18].

TFs are able to integrate different and several cellular signals. Cross talks between nuclear-receptor- mediated and signal-transduction pathways, stemming from the activation of signaling cascades controlled by extracellular receptors, define complex gene expressions in different cells.

Emerging evidence of the interaction between hormonal and growth factor signaling pathways are reported. For instance, Estrogen Receptor - ER activation in human cancers is promoted by a crosstalk between ER and receptor tyrosine kinases (such as ERBB2, epidermal growth factor receptor 1 and insulin-like growth factor 1 receptor), Phosphatidylinositol 3-Kinase - PI3K, tyrosine-protein kinase SRC and Mitogen-activated protein kinases - MAPKs [19]. In many instances, rather than a linear response mediated solely through estrogen responsive DNA elements, NRs activity can be modulated through growth factors and/or membrane-initiated responses [20].

Given the wide variety of processes controlled by TFs, mutations/alterations in the signaling molecules involved in their activation/repression can be deregulated and can contribute to cancer or other diseases.

We will consider in detail how TFs can be activated through these cytoplasmic signaling-networks cascades and we will touch more briefly the TFs - NRs activated directly by internalized ligands.

4. Transcription Factors Directly Activated by Hormones and Vitamins

The nuclear receptor (NRs) superfamily includes a related but diverse array of transcription factors, able to bind structurally diverse ligands. Among the TFs ligand-activated according to this general behavior, we mention:

- The estrogen receptor (ER).
- The androgen receptor (AR).
- The glucocorticoid receptor (GR).
- The thyroid hormone receptor (TR).
- The retinoic acid receptor (RAR).
- The vitamin D receptor (VDR).

In many cases, ligands (see Fig. 3) and corresponding NRs have been identified, but several “orphan receptors” remain, whose ligands, target genes, and physiological functions were initially unknown [22, 23].

Forty-eight NRs have been described in humans [23, 24]. The structural organization of NRs is shared among many of them:

- An amino terminus (the A/B region) with a trans-activation domain (AF-1), variable in length and in sequence in the different family members. Co-activators and/or other transcription factors are able to recognize it.
- A central DNA binding domain (DBD) responsible for targeting the receptor to highly specific DNA sequences (TFBS) comprising a response element.
- A carboxy-terminal ligand-binding domain (LBD) well conserved between the various family members. This LBD guarantees selective ligand recognition, with the ligand-induced activation function domain (AF2) crucially involved in transcriptional co-regulator interactions [14, 23].
Figure 1: A detail referring to the pathways downstream of the TGF-β growth factor depicted in the Molecular Interaction Map (MIM) depicted in [12] (reproduced with permission of 2008-2015 Impact Journals, LLC). The text explanation is slightly modified from the Supplementary Material - Annotation List of [12]. The symbol # followed by an Arabic number indicates an interaction or a contingency depicted in the MIM. Syntactic rules of the MIM are depicted in Fig. 2. Dimeric TGF-β (#108 – number in the annotation list of [12]) binds with high affinity (#110) type II (TGFβR-II) receptors dimers (#109), on the cell surface. Dimeric TGF-β type I receptor (TGFβR-I) (#111) binds the dimeric [TGFβR-II: dimeric TGFβ complex] (#112). In the hetero-tetrameric receptor complex, type II receptors phosphorylate a serine/threonine-rich region (the GS region), located in the kinase domain of TGFβR-I (#113), which then propagates the signal. An activated TGF-β complex can bind (#127) and phosphorylate Smad3 and Smad2 (#129, #130). Smad4 can bind activated Smad2 and Smad3 (#132) forming complexes that can translocate into the nucleus (#133). These SMAD complexes can then bind to a SMAD binding site [TFBS:SMAD] on MYC, repressing its transcription. SMAD4 (most likely complexed with phosphorylated SMAD2 or SMAD3) also binds (an additional example) to the promoter region of CCND1 upon a TGF-β stimulus, repressing transcription.
NRs can exist as monomers, homo-dimers, or heterodimers and recognize DNA sequences termed hormone response elements (HREs). HREs are bipartite elements that are composed of two hexameric core half-site motifs forming direct, indirect or inverted repeats, which consist of two half-sites separated by a short spacer [14]. They are a subclass of TFBS.

The NRs can been subdivided in two main classes (Fig. 4), based on their mode of action [16, 23]:

- **Type I**: ligand binding in the cytosol where NRs are sequestered in the absence of ligand (such as the androgen receptor, the estrogen receptor, and the progesterone receptor).

- **Type II**: ligand binding in the nucleus, where NRs are sequestered by co-repressors in the absence of ligand (such as the thyroid hormone receptor and the retinoic acid receptor).

- **Type III** nuclear receptors (such as Retinoid X receptor - RXR and Hepatocyte Nuclear Factor – HNF-4) function similarly to type I; however, type III nuclear receptors bind as homodimers to direct repeat instead of inverted repeat HREs.

- **Type IV** nuclear receptors (such as nerve growth factor-induced gene - NGFI-B and steroidogenic factor -SF-1) bind as monomers to a half site HREs. Most of the orphan receptors, belong to type III and IV [16, 23].

Type I receptors have a cytoplasmic heat shock protein type (HSP) associated with the inactive NR. When the receptor interacts with the ligand, the NR is released. Type I NRs may be found in homodimer or heterodimer forms. Free (that is, unbound) hormones enter the cell cytoplasm and interact with their NR - TF-receptor, anchored in the cytoplasm by chaperone proteins (HSPs). In this process, HSPs are dissociated, and the activated ligand-receptor complex trans-locates into the nucleus.

In the case of steroid hormones, after binding to the ligand, NRs often form dimers. In the nucleus, the complex
Figure 4: Schematic signaling by Type I nuclear receptors (such as the estrogen receptor - on the left) and Type II nuclear receptors (such as thyroid hormone receptor - on the right), depicted as a Molecular Interaction Map. Syntactic rules of the MIM are depicted in Fig. 2. The red cartouche represent the molecule involved in the biochemical interactions (HSP - heat shock protein; oxysterol is an intracellular ligand; RXR - retinoid X receptor).

associates with transcriptional co-activators and acts as a TF, increasing or suppressing transcription of particular genes by its action on DNA.

Type II receptors do not have HSP-binding and are located already in the nucleus, bound to their specific DNA response elements even in the absence of ligand.

Their ligands cross the cell membrane, the cytoplasm, the nuclear membrane, and enter the nucleus, where they activate the receptor without release of HSP. The activated NR interacts on DNA with hormone response elements (a variety of TFBS) and the transcription process is initiated as with type I receptors. All-trans and 9-cis retinoic acids are bioactive Vitamin A components. They bind and activate their cognate NRs to regulate target genes. NRs are used also by the thyroid hormones and Vitamin D. This in addition to steroid hormones.

There are actions of ligands taking place not within hours but within minutes, these are transcription independent.

As is usual for TFs, the binding to TFBS by these NRs can be modulated both by co-activators and co-repressors [25]. Some NRs have a low basal level of transcriptional activity in the absence of an activating ligand. Synthetic ligands, which reduce this basal level, are called inverse agonists.

Some synthetic ligands can stimulate or repress according to different tissues (selective NRs modulators).

5. TFs Activated by a G0-G1-S Cytoplasmic Signaling-Network

All living organisms are products of repeated rounds of cell growth and division + changed patterns of transcribed genes (differentiation). During the four phases / steps of the cell cycle process, each cell duplicates its contents and then divides. The cell cycle can be subdivided in two Gap phases (G1 and G2) during which the cell prepares to divide, an S - Synthesis phase, during which the DNA is duplicated and an M - Mitosis phase, in which the cell divides into two daughter cells, each with a full copy of DNA. The cell cycle process is tightly regulated to ensure that specific events (for instance the G1-S checkpoint and DNA-structure checkpoints) take place in an orderly manner [2, 26].

In a normal mammalian cell, growth factor signaling activation determines whether early G1 phase cells transit a restriction point “R” (START in yeast) to undergo eventual cellular division or exit the cell cycle, and enter into a quiescent state G0 (the resting phase). During progression to a neoplastic state, the cell becomes largely independent of growth factor signaling and can also avoid a G0 exit to a quiescent state. A cancer cell becomes capable of a growth-factor independent transition across the G1-S restriction point. It does this activating altered pathways not through growth factor stimulation, but through the onset of mutations.
Differentiation regulated transcription factor Proteins - DP (E2F1-3/DP1-2 for instance). Some of the E2F family members are required to control the timely activation and repression of specific target genes that are essential for ordered progression through the cell cycle, in different cellular contexts and differentiation conditions.

During the G1 to S transition, phosphorylation of the TF repressor pocket family proteins (Retinoblastoma or Rb for short, p107 and p130) by cyclin-dependent kinases (for instance CyclinD:Cdk4 and CyclinE:Cdk2) releases the inhibition of pocket proteins from transcription factors (E2F1-3/DP1-2 for instance). Some of the E2F transcription Factors - E2Fs1-5 need to hetero-dimerize with one of three Differentiation regulated transcription factor Proteins - DP that are capable of binding to DNA. Figure 5 depicts the network involved in the G1-S transcriptional activation.

In cooperation with other transcription factors (see Fig. 5), E2F family members are required to control the timely activation and repression of specific target genes that are essential for ordered progression through the cell cycle, in different cellular contexts and differentiation conditions.

E2F1-3:DP1-2 responsive genes can be positive cell cycle gene regulators. Among the positive regulations: CCND1, JUN, MYC, MYCN, CCNE1, CCNE2. Among the negative cell cycle gene regulators: E2F4-8, RB1, RBL1, TP53 [28].

CyclinE:Cdk2 phosphorylates further the pocket proteins, favoring the transcription of more Cyclin D/E, generating an initial positive feedback loop, further activating G1-S transcription [26].

Several feedback loops (including E2F4-8:DP2 activation) terminate the first wave of gene expression at the transition from G1 to S phase. An example of a negative feedback loop is the transcription of E2F6-8 downstream of E2F1-3:DP1 transcription factors. This happens shortly after the phosphorylation of E2F1-3 by CyclinE:Cdk2. Phosphorylated E2F1-3 detach from their promoters. The combination of these effects, represents a negative feedback loop, because E2F6-8 now can repress transcription at the same promoters previously activated by E2F1-3. Additional gene target of E2F1-3 is also the S-phase kinase-associated protein - SKP2, a protein involved in an ubiquitin ligase pathway degrading E2F1-3. This represents an additional component of the negative feedback loop [26, 29].

Transcription of G1-S cell cycle target genes is irreversibly decided at a G0 → G1 transition point, required not just by a quiescent-cell maintenance, but by a new cell replication.

This restriction point “R” divides the cell cycle into a growth factor dependent early G1 phase and a growth factor independent phase from late G1 through S phase. Transition of this point commits cells to enter the cell cycle and progress through it independently of signals from the environment. Mutations/alterations in driver-gene mutations that encode proteins belonging to the signaling network involved in the control of G0-G1 (see our MIM pathways upstream of MYC/CCND1 transcription, Fig. 5) can precisely disrupt (through a discrete number of hits) the ‘irreversible restriction point’ [27, 30]. The gate(s) to S phase and further cell cycle progression will remain open. This multi-hits event represent a hallmark of cancer.

6. Dynamic Simulations: Molecular Interaction Map

The cell cycle signaling network upstream of the transcription of genes required for the S phase has been modeled through a dynamic mathematical model [12], at the level of (bio)chemical interactions/reactions. From a basic perspective, chemical interactions/reactions, are at the very basis of the life of any cell, and therefore of life for short. Going back in time about four billions of years, we could advance the hypothesis that local chemical interactions/reactions were already central in the transition from a pre-biologic to a biologic world [31]. It was “logically/mathematically necessary” that molecular combinations and interactions endowed with properties favoring the local appearance of molecular combinations closer to the parental combinations would be "evolutionary favored (conserved)".

We studied and implemented dynamic simulations of multiple interconnected downstream pathways [32, 33]. We modeled a network sub-region related to the G0-G1-S cell cycle transition, downstream of TGFβ, WNT, HGF, EGF-family of growth factors [12]. Recently we have enlarged our MIM, including the ALK, the MET and the Extracellular Matrix Integrin receptors pathways (see Fig. 5 and Table 1) [29]. The signals downstream of them propagate through a complex network, involving cross talks among interacting pathways, and strong feedback loops on different levels. In our MIM, the pathways (Fig. 5) regulate five TFs (TCF7L2, SMAD4 complexes, AP1, TP53, E2F:DP1) and their activators/repressors (Fig. 6). TCF7L2 (transcription factor 7-like 2 (T-cell specific, HMG-box)) also known as TCF4 and Activator protein 1 (AP-1), a heterodimer composed of proteins belonging to the c-Fos, c-Jun, ATF families are TFs regulating gene expression in response to a variety of stimuli, controlling a number of cellular processes including proliferation. Tumor Protein TP53 is a TF acting as a tumor suppressor in many tumor types, inducing growth arrest or apoptosis depending on the cell type.

Defects in the control of this interconnected network of pathways play an important role in colorectal cancer (CRC) pathogenesis and other solid tumor too. We verified the
emergence of new properties originating from this integrated network of more than 100 signaling proteins.

Using a Molecular Interaction Map (MIM), we realized a comprehensive reconstruction of the “molecular anatomy” of the network (Figure 5). A MIM represents “a diagram convention that is capable of unambiguous representation of networks containing multi-protein complexes, protein modifications (phosphorylation for instance), and enzymes that are substrates of other enzymes. This graphical representation makes it possible to view all of the many interactions in which a given molecule may be involved, and it can portray competing interactions, which are common in bio-regulatory networks” [34–37].

Figure 2 provides the syntax for reading the MIM. A MIM includes also an Annotation List (see for instance the legend of Fig. 1) providing text explanation of the interactions reconstructed and their specific references and a Glossary [12] referring a brief explanation of the species depicted in the MIM.

Drawing our MIM, we move in the direction of an extensive representation of the crucial G0-G1-S transition, and therefore we should find in it most of the cancer gatekeeper mutations [38–40].

Tumor and non-tumor genomes contain thousands of somatic mutations (often called passenger mutations), but only a few of them “drive” the development of a cancer cell, by affecting genes which confer selective growth advantage to tumor cells. A discrete number of mutations/alterations (let say 4-8 hits) is required to make the cancer cell actually to tumor cells. A discrete number of mutations/alterations by affecting genes which confer selective growth advantage only a few of them “drive” the development of a cancer cell, somatic mutations (often called passenger mutations), but still correlated (unpublished data). Their number increased by about 50% as of June 2015 [29].

For dynamic mathematical simulations, an efficient software is available for achieving numerical solutions of ODEs. We can achieve a satisfactory representation of the temporal evolution of large networks of biochemical interactions. Simbiology was the Matlab toolbox used in our case [42]. After the introduction of a perturbation in the model, for instance through an altered/mutated signaling-protein, Simbiology is able to converge rapidly to a new equilibrium.

8. Parameterization of the Model

Extensive pre-training of our model, was based on the input of more than 100 pertinent articles. The parameters are the concentrations of molecular species (proteins) and the rate parameters governing their interactions. We performed a sort of patchwork of fitting together multiple protein concentrations and kinetic parameters, so that they could fit the inputs of all published papers utilized in the training phase. We achieved a semi-quantitative predictive model, statistically very significant [12]. We noticed that our parameters can be discretized in eight log-scale intervals (ranges) keeping the same good correlation with experimental results. A discretization in four log-scale intervals (ranges) is suboptimal, but still correlated (unpublished data).

In the near future, we could be able to improve our model further, introducing absolute quantification of proteins and phospho-proteins and improving information concerning kinetic constants. The protein quantification of the signaling proteins involved in our network could be carried out using for instance, nanoflow high performance liquid chromatography (nano-HPLC) coupled to high-resolution electrospray tandem mass spectrometry (HR-ESI-MS/MS). Recent observations suggest that protein concentration differences in different types of tumor, could explain, at least in part, different responses to altered pathways inhibitors [43].

A direct parameterization of rates requires known or derived (protein) structures of the involved molecules. Sophisticated computational pipelines are available to predict protein structure along with the prediction of protein-protein interactions, may guide the behavior of a cancer cell, and take a therapeutic advantage from this more advanced understanding.

Starting from the relationship between molecular species in specific pathways, depicted as MIM, we moved toward a dynamic mathematical model using Ordinary Differential Equations - ODEs. As of the end of 2014, the model involved 447 reactants (basic species, modified species, complexes and inhibitors), 348 reversible reactions and 174 catalytic reactions: \((348 \times 2 + 174 = 870 \text{ reactions})\) [12]. Their number increased by about 50% as of June 2015 [29].
Figure 5: Molecular Interaction Map of the pathways upstream of the TFs TCF7L2, SMAD4, AP1, TP53, E2F:DP1, involved in the control of G0-G1-S genes transcription (MYC and CCND1 genes as examples). A light yellow oval surrounded the cartouches of mutated/altered signaling-proteins in the colorectal HCT116 cancer line. The targeted inhibitors and their relative targets are indicated. Reproduced from [29].

interaction sites. We need extensive docking simulations to derive all possible interactions among the molecules. We can use molecular dynamics simulations to derive Gibbs free energies and to predict equilibrium concentrations. Introducing a general approximation at the level of association rates, we can finally derive from the Gibbs free energies both association and dissociation rates (Olaf Wolkenhauer, University of Rostock, personal communication, [29]).

Starting from an initial “physiological condition”, the model can be adapted to simulate individual pathologic cancer conditions, implementing (sequentially) discrete numbers
of alterations/mutations in relevant onco-proteins. We performed inhibitor treatments, measuring MYC and CCND1 mRNA levels and the amount of specific phospho-proteins (AKT and ERK). We verified some salient model predictions using the mutated CRC lines HCT116 and HT29 [12]. We also verified the good behavior of our pre-trained model against pre-clinical results published in [44].

Our dynamic model suggests that the duration of the initial G0-G1 cell cycle phase is short (30–60 min). As depicted in our MIM, during this phase, in order to respond to different signals or growth factors, a cell prepares transcription/translation of new proteins through a complex network of biochemical interactions and post-translational modifications, for instance phosphorylations – de-phosphorylations. At an experimental level, stimulation with FBS (fetal bovine serum) of a quiescent mouse epidermal cell line, showed AKT and GSK-3β phosphorylation after 15–30 min [45]. FBS represents the most widely used serum-supplement for the in vitro cell culture of eukaryotic cells, containing growth factors. Similar times for MAPK phosphorylation have also been reported [46].

During the last four hours of the G1 phase, new proteins required for entering the S phase of the cell cycle, are transcribed and synthesized. Multi-phosphorylation of RB1 protein, results in the release of E2F and the transactivation of target genes, in cooperation with other pro-transcriptional G0-G1 pathways. Additional proteins will be synthesized in the early S phase, including proteins required for DNA duplication and finally the duplication of the components of the entire cell mass [26, 28, 29].

In the G0-G1 cell cycle phase that we have reconstructed, we do not have a transit through multiple waves of
subsequent new transcriptions/translations + accompanying protein degradations. Our model simulates only the G0-G1 cell cycle phase depicted through our MIM, at a biochemical interactions level. The complexity of the network can still be kept reasonably under control.

**9. Dynamic Simulations: Synergism among Five Important Driver Gatekeeper Mutations in CRC**

It has been recently developed an epithelial organoid culture system in which human intestinal stem cells (ISCs) indefinitely self-renew and form crypt-like organoid structures in Matrigel [47, 48]. Using CRISPR-Cas9–mediated engineering of human intestinal organoids they can be progressively transformed in neoplastic lesions [49, 50]. In the organoids papers, the authors considered the five (gatekeepers) high frequency mutations in CRCs: TP53 (48%), APC (44%), KRAS (30%), PIK3CA (18%), SMAD4 (11%), as reported in the COSMIC release of May 2016 [51].

We introduced in our dynamic model, all possible permutations (32) for 0 (1), 1 (5), 2 (10), 3 (10), 4 (5), 5 (1) mutations (Fig. 7). We simulated the effect of the possible hits in a neoplastic transformation as mutations in the pathways depicted in Fig. 5 and listed in Table 1, and we verified their effect on these deregulated pathways, at the level of mRNAs transcriptions immediately downstream of the G0-G1 restriction point [29], referring to the transcription algorithm of [12].

Mutations in genes encoding proteins along the WNT, MAPK, TGF-β, TP53, and PI3K pathways depicted in Figure 5 (and in Table 1) are recurrent in human colorectal tumors.

When we introduce the 10 combinations of two mutations:

- The less synergic combinatorial couple of mutations is [PI3K – TP53; 3.13], along pathway #1 [ErbB-family receptors – PI3K – PTEN – AKT – MDM2 – TP53 – TFBS] (see Table 1). Our results support the hypothesis that for PI3K and TP53 mutations, the signal seems to travel largely along the same pathway. A major result of each of the two mutations is a loss of TP53 function.

- The most synergic combinatorial couples are [KRAS + TP53; 10.39] and [KRAS + PI3K; 9.86]. These mutations belong to the pathways #1 and #2 respectively (see Table 1): [ErbB-family receptors – Shc – Grb2 – SOS – GAP – KRAS – BRAF – MEK – ERK – AP1 – TFBS] (transcription antagonist) and [ErbB-family receptors – PI3K – PTEN – AKT – MDM2 – TP53 – TFBS] (transcription antagonist). As depicted in our MIM (Fig. 5), these two pathways seem largely independent between themselves. The introduction of the mutations tend to give gain of function in the pathway #1, and loss of inhibitory function in the pathway #2. The strongest two-mutations synergisms could be explained through the independence of the two pathways #1 and #2.

The pathway #6 [TGFB receptors – SMAD2/3 – SMAD4 – TFBS] (transcription antagonist) appears a pathway independent from the previous mentioned pathways already starting from the level of the cytoplasmic membrane receptors (for instance pathways #1, #2, #5 see corresponding Table 1), therefore its loss of inhibitory function can be efficiently synergic in a three-mutations combination. Maximum observed transcription efficiency of this triple combination of mutations and altered pathways at the G1-S restriction point: 20.30 times the physiologic condition (mutations in the KRAS + TP53 + SMAD4 pathways).


In the presence of four mutations, we find a lower synergism when mutations of PI3K and TP53 are both present. The two less synergic couples of mutations (PI3K + TP53) + SMAD4 & APC, 12.69; + KRAS & APC, 17.84; + KRAS & SMAD4, 21.94.

As expected, in the presence of a TP53 altered pathway: [ErbB-family receptors – PI3K – PTEN – AKT – MDM2 – TP53 – TFBS] (transcription antagonist), the addition of a PI3K mutation when TP53 is already inactivated is not relevant in respect to the best combinations of four mutations. Transcription level of the five mutations: 34.11). Notice that TP53 is the most frequent driver gate-keeper mutation in CRC.

**10. Dynamic Simulations: Combinations of Targeted Inhibitors in CRC**

The strategy of matching targeted agents with genomic alterations is associated with encouraging results in the treatment of patients with various cancers [52, 53]. Targeted agents are intended to exploit the phenomenon of “oncogene
Table 1: List of pathways depicted in the MIM.

1. Pathway [ErbB-family receptors – PI3K – PTEN – AKT – ramifications a, b, c, d];
3. Pathway [ErbB-family receptors – E-Cadherin (Cadherin/Catenin adhesive complex)];
4. Pathway [ErbB-family receptors – PLCγ – PI2P – PKC – BRAF – MEK – ERK – AP1– TFBSAP1, transcription agonist]; the terminal parts of pathway 2 and 4 are the same;
6. Pathway [TGFβ-receptors – SMAD2/3 – SMAD4 – TFBSSMAD, transcription antagonist];
7. Pathway [TGFβ-receptors – TAK-1 – TAB2 – NLK – TCF7L2 – TFBSCTF7L2 (TCF7L2 binding site), transcription agonist], converging with 8;
9. Pathway [Integrins – ILK – PIP3 – AKT – ramifications a, b, c, d]
10. Pathway [Integrins – FAK – AKT – ramifications a, b, c, d]
11. Pathway [Integrins – FAK – PI3K – AKT – ramifications a, b, c, d]
13. Pathway [Integrins – FAK – SRC – ramifications e and f]
14. Pathway [ALK – SRC – ramifications e and f]
15. Pathway [ALK – PI3K – ramifications a, b, c, d]
18. Pathway [c-Met – SRC – ramifications a, b, c, d];
19. Pathway [c-Met – SRC – ramifications e and f];

AKT Ramifications
b. Pathway [AKT – mTOR – p70S6K]
c. Pathway [AKT – MDM2 – TP53 – TFBSTP53, transcription antagonist]

SRC Ramification
f. Pathway [SRC - PI3K – AKT – ramifications a, b, c, d]

List of biochemical interactions / reactions pathways depicted in the MIM (Fig. 5). Reproduced from [29].

addiction” [54, 55], in all cases in which an “oncogene addiction” effect is present. Cancer cells, along their evolutionary process, can become dependent for their survival on the constitutive activity of a defined mutated (or functionally drastically altered) oncogene. Therefore the consequence of its inhibition can in principle become lethal (or more severe) for the malignant cell, than for the corresponding healthy cell. Combinations of specific inhibitors of two – three altered pathways can do something similar to bringing back the 4-5 multiple hits that induced the malignant transformation to 1-2 residual hits. This condition of incomplete transformation could be incompatible with the life of a malignant cell.

The combined introduction of drugs that target signaling-proteins characterized by a functional excess, within the
biochemical networks controlling many cellular functions, represents an important improvement in the treatment of a cancer patient [56].

We introduced in the model and we simulated the behavior of a virtual organoid carrying four specific mutations (APC, KRAS, PI3K and SMAD4) treated with specific virtual inhibitors of onco-proteins, alone or in combinations of 0 (1), 1 (5), 2 (10), 3 (10), 4 (5), 5 (1) inhibitors (Fig. 8). The inhibitors that we simulated were PanERB inhibitor, CI1040 (MEK inhibitor), PI103 (PI3K inhibitor), Perifosine (AKT inhibitor), Nutlin-3 (MDM2 inhibitor) [29]. Some of them are already in clinical trials. For instance, MEK inhibition appears to have a strong rationale in the presence of an excess function of this pathway as a consequence of a mutation in KRAS or in BRAF [57].

Figure 8 shows how the transcription levels decrease as the number of targeted inhibitors used in the simulation model increase, a behavior in agreement with the emerging suggestion of using therapies consisting of simultaneous combinations of inhibitors of independent protein signals [58].

Analyzing the effects of a single inhibitor treatment, the dynamic simulation showed that a PanERB inhibitor treatment (1 of Fig. 8) did not decrease the transcriptional gene’s level, in the presence of downstream mutations in the PI3K and KRAS pathways (for instance as simulated in our model, mutations in PI3K and/or KRAS proteins respectively). The reason is that PanERB inhibitor’s target is upstream along the PI3K and KRAS pathways, as depicted in our MIM (Fig. 5 and Table 1).

On the other hand, the treatment with CI1040 (2 of Fig. 8), a MEK inhibitor, showed a significant effect (the MEK inhibitor’s target is downstream of the KRAS mutation), synergic with the inhibition of MDM2 (2-5 of Fig. 8). An association of PI103 (PI3K inhibitor) and Perifosine (AKT inhibitor) was also synergic (3-4 of Fig. 8 - more complete inhibition of the same altered pathway?). An additional synergism is observed with the 2-3-4 combination (inhibition of MEK + PI3K + AKT: equivalent to the inhibition of two important complementary and both altered pathways).

11. Conclusions

We have considered two different sets of TFs:

- TFs as Nuclear Receptors, directly activated by stimulating molecules coming from outside the cell, capable of entering the cell, and controlling differentiation behaviors dependent both by hormones and hydrophobic vitamins.
- TFs regulated by a complex signaling-network active in the cell cycle phase immediately preceding the G1-S restriction point, above the irreversible decision by the cell of entering the S phase. In this case, we considered growth factors exerting their stimulation through trans-membrane receptors, and activating a cytoplasmic signaling-network downstream.

Biochemical interactions are not linear pathways but rather an interconnected network rich in ramifications and
Figure 8: We introduced in our dynamic model, all the possible permutations of the five-targeted inhibitors PanERB inhibitor (1), CI1040 (MEK inhibitor) (2), PI103 (PI3K inhibitor) (3), Perifosine (AKT inhibitor) (4), Nutlin-3 (MDM2 inhibitor) (5), in the presence of the four mutations (and mutated pathways): APC (WNT pathway), KRAS (MAPK pathway), SMAD4 (TGF-β pathway) and PI3K (PI3K pathway). In the ordinates, we show the increased transcription rates of the complex described in Fig. 6. The numbers in the figures indicate the combinations of inhibitors whose effects differ most from the median value. Reproduced and slightly modify from [29].

feedbacks. The downstream TFs link and integrate several signaling cascades controlled by extracellular receptors. Interactions among the effects of NRs and TFs downstream of a signaling-network have also been reported.

We have treated in deeper detail an important section of the second type of regulation. We reconstructed the biochemical interactions of a network upstream of the G1-S restriction point through a Molecular Interaction Map (MIM). We depicted (and modeled by ODEs) most of the interconnected pathways acting on gene’s transcription, required to reach the restriction point. Our model is capable of simulating what is going on when a virtual cancer cell, carrying specific mutations, decides to cross the G1 – S restriction point (a crucial decisional point). Mutated proteins (of a driver – gatekeeper type), especially the most important in bringing a normal cell to cancer, were introduced in our mathematical model.

Targeted drug molecules directed to inhibit signaling-proteins affected by excess of function within an altered pathway were simulated as “normal chemically interacting” molecules in our network of biochemical interactions. It is possible to introduce in the model a multiplicity of combination therapies, suggested by an interrogation of the model and databases of targeted drugs. Our dynamic model can simulate a combination of two, three, four, five specific inhibitors of distinct altered pathways. This strategy can suggest an appropriate combination of targeted drugs in the treatment of a cancer patient molecularly characterized, an innovative perspective of incoming years.

The complexity of the model goes well beyond the breaking point of what our mind is capable of connecting and following intuitively in its consequences: a main reason of the usefulness of the model. Signaling-network dynamic simulations represent an essential tool to understand better the molecular complexity of cancer and provide new avenues for targeted combinatorial treatments.

**Author Comment**

N. Castagnino, M. Maffei, and L. Tortolina contributed equally to the review article.

**Competing Interests**

The authors declare that they have no competing interests.

**Acknowledgements**

This work was supported by Funding Italian Ministry of Economic Development “Industry 2015 - Made in
References

[33] N. Castagnino, L. Tortolina, A. Balbi, R. Pesenti, R. Montagna, A. Ballester, D. Sconini, E. Moran, A. Nencioni, and S. Parodi, Dynamic simulations of pathways downstream of ERBB-family, including mutations and treatments: Concordance with


