Review

The structural network of inflammation and cancer: Merits and challenges

Emine Guven Maiorov, Ozlem Keskin, Attila Gursoy, Ruth Nussinov

Keywords: Sackler National Science Institute, Center for Cancer Research Nanobiology Program, Frederick, MD 21702, United States

Corresponding author: National Cancer Institute, Frederick, MD, United States.

ARTICLE INFO

Proteins perform their functions via interactions; among these, interactions with other proteins are the most common, and they control cell life. Thus, there is no wonder that a major effort by the scientific community has targeted the modeling of protein–protein interactions. The availability of structural data relating to the mode of protein interactions is crucial to the understanding of how function is executed, and its control. When tiled together in the framework of the cell, it helps obtain the information flow from the extracellular space, through the membrane, the cytoplasm, and into the nucleus, to turn genes on and off. Signaling is not linear: pathways merge and diverge, signals integrate, and the proteins, which are the nerve centers of these pathway crossings, are frequently involved in cancer. Mutations deregulating these proteins are often responsible for turning on, and keeping on, entire pathways. They can also deregulate proteins that act as repressors of over-expression and activity.

The community has invested much effort in constructing pathways and in devising simple and clear ways of presenting them. The most common among these is the so-called nodes-and-edges representation, where nodes are the proteins and edges their interactions. This representation has been extremely useful since it essentially provides an overview of which proteins interact, in much the same way as a reference guide. At the same time, such pathway diagrams also suffer from limitations [1], since the information that they provide is incomplete, and does not allow in-depth understanding of the processes and their regulation. In contrast, structural networks which piece together the structures of individual proteins are much more powerful [2]. Structural networks allow understanding of how pathways cross through key hub proteins; which domains are involved in the interactions and which residues. They provide insight into the question of whether the

ABSTRACT

Inflammation, the first line of defense against pathogens can contribute to all phases of tumorigenesis, including tumor initiation, promotion and metastasis. Within this framework, the Toll-like receptor (TLR) pathway plays a central role in inflammation and cancer. Although extremely useful, the classical representation of this, and other pathways in the cellular network in terms of nodes (proteins) and edges (interactions) is incomplete. Structural pathways can help complete missing parts of such diagrams; they demonstrate in detail how signals coming from different upstream pathways merge and propagate downstream, how parallel pathways compensate each other in drug resistant mutants, how multi-subunit signaling complexes form and in particular why they are needed and how they work, how allosteric events can control these proteins and their pathways, and intricate details of feedback loops and how kick in. They can also explain the mechanisms of some oncogenic SNP mutations. Constructing structural pathways is a challenging task. Here, our goal is to provide an overview of inflammation and cancer from the structural standpoint, focusing on the TLR pathway. We use the powerful PRISM (Protein Interactions by Structural Matching) tool to reveal important structural information of interactions in and within key orchestrators of the TLR pathway, such as MyD88.

© 2013 Elsevier Ltd. All rights reserved.
pathways are controlled allosterically; how constitutive gain- and loss-of-function mutations can turn the protein on or off; and how parallel pathways can take over during drug resistance. Structural pathways are also very useful in drug discovery, because they can help forecast the global effects on the cell [3–6]; and they can help in selecting drug targets which are not the direct ‘ailings’ dysfunctional protein [7,8]. Such drugs exploit conformational ensembles [9,10]; they may also make use of network dynamics [11].

Constructing the structural network of major pathways in the cell is challenging: it entails putting together available structures (determined by crystallography, NMR, and high quality models) when data about their interactions exist in the literature, and when these data are missing, predicting them. It necessitates modeling the linear pathways as well as pathway integration and branching; it also necessitates predicting if the protein can act as a dimer – homo or hetero, and possible oligomerization modes. To grasp the enormity of the challenge, recall that hub proteins, such as p53, can interact with tens or hundreds of partners; Raf can work as a monomer and as a dimer; and EGFR has two dimeric states, symmetric (inactive) and asymmetric (active). However, the interaction surfaces are often unknown [1,12–14].

So why are structural pathways important? Structural pathways predict new interactions, not observed in ‘classical’ pathway maps; in particular those involving scaffolding proteins; they allow identification of parallel pathways in drug resistant mutants; they may discover positive feedback loops, altering core processes, as in the case of inhibitors of apoptosis (IAP) [15]. They also provide interaction details, which allow identification of oncogenic mutations; help drug discovery; and allow assembling multi-subunit signaling complexes, such as the key complex MyD88.

A current goal in our lab is to take steps toward accomplishing the overarching task of modeling protein interactions in cancer-related pathways and the cancer network in the cell and mapping oncogenic mutations associated with these. Within this framework, here we aim to provide an overview of inflammation and cancer from this structural standpoint. Below, we first provide a broad description of the inflammation and cancer link that we aim to model, and some examples as they relate to this link. We then explain the challenge in modeling protein interactions on a large scale, particularly when seeking to also detect new interactions, which are not known a priori. We follow by explaining PRISM [16,17], the method that we use to model these pathways, its advantages and shortcomings.

Computational structural biology helps experiment by providing more complete information and leads. It is our belief that a complete map of key cellular pathways with structural data - including multimolecular associations of adaptor proteins - is critical to fully understand cancer cell biology with the abnormality originating from deregulation of signaling pathways.

1. Inflammation and cancer

Inflammation by innate immunity, which is required to fight microbial infections, heal wounds, and maintain tissue homeostasis, can lead to the hallmarks of cancer [18–20]. Several recent studies suggested that inflammation has an important role in all phases of tumor development, including tumor initiation, tumor promotion, invasion, metastatic dissemination, and evade the immune system [18,19,21]. Inflammation causes cellular stress and may trigger DNA damage or genetic instability [21], and chronic inflammation can contribute to primary genetic mutations and epigenetic mechanisms that initiate malignant cell transformation [20–22]. Tumor-promoting effects of inflammation alter tissue homeostasis, predisposing individuals to cancer [21,23,24]. The connection between inflammation and cancer has been established by Wirchow already in 1863, who noticed that cancers originate at sites of chronic inflammation and tumors have “lymphoreticular infiltrate” [20,21,25,26]. Inflammation establishes a tissue microenvironment, which tolerates tumor growth and metastasis by setting immunosuppressive mechanisms [21]. Therefore, inflammation not only induces carcinogenesis but also makes immune cells incapable of destroying tumor cells. A key inhibitor of a major pathway leading to inflammation is the TLR pathway (displayed in Fig. 1).

Inflammatory cells supply growth factors to maintain proliferation, survival factors to escape from apoptosis, pro-angiogenic factors and extracellular matrix (ECM) modifying enzymes that enable angiogenesis, invasion and metastasis [18]. Inflammatory cells can also secrete reactive oxygen species (ROS) that induce mutations, lead to failure of DNA repair, activation of oncogenes, and ultimately cancer [18,20,21]. ROS further activates inflammatory genes and takes part in tumorigenesis which is regulated by c-MYC, K-Ras, and Wnt signaling pathways [18].

The first evidence that connected inflammation with cancer was that inflammatory diseases such as inflammatory bowel disease (IBD) increase cancer susceptibility [27]. Additional evidence comes from the tumor microenvironment, which has inflammatory cells, cytokines and chemokines. In addition, long term administration of non-steroidal anti-inflammatory drugs (NSAIDs) leads to a decrease in the number of relapses and newly acquired tumors [18]. Moreover, pathways of inflammation function downstream of oncogenic mutations (MYC, RAS, BRAF, and RET), suggesting that these oncogenic mutations lead to activation of an inflammatory pathway [28]. Last but not least, inhibiting inflammatory cytokines and chemokines such as TNF-α, IL-1β, or essential transcription factors, such as NF-kB and STAT3, decrease the appearance of cancer [24]. Most of these proteins that connect inflammation and cancer are also members of TLR-pathway.

The TLR pathway in Fig. 1 is presented in the classical node-and-edge form. As we noted above, such pathway diagrams are informative and have been enormously useful in compiling and laying out the interaction maps. As can be seen in Fig. 1, they provide the interconnectivity between the proteins and outline the major signaling flow. Nonetheless, the information that they are able to provide is incomplete. Structural pathways can fill in gaps; add missing proteins and interactions, and provide the interaction details, which can help in figuring out parallel pathways, and the conformational mechanisms that control them. These are important for understanding function under physiological conditions and oncogenic mutations and drug resistance in disease.

The link between inflammation and cancer appears to stem from two pathways [29], the intrinsic and extrinsic inflammation pathways, as shown in Fig. 2. Intrinsic inflammation is initiated by mutations that lead to activation of oncogenes and inactivation of tumor suppressors (tumor-promoter role) [24]. On the other hand, in the extrinsic pathway, infection or inflammation precedes cancer and increases the risk of cancer (tumor-initiator role) [24]. The similarity between cancer tissue and inflamed tissue involves angiogenesis and tissue infiltrating leukocytes, such as lymphocytes, macrophages and mast cells [18,19].

RAS, MYC, RET, and BRAF are among the oncogenes that stimulate inflammation [21]. Besides their roles in tumor initiation and promotion, oncogenes impact the relationship between the tumor and its microenvironment. The RAS–RAF signaling pathway is involved in most cancer types, and the activated proteins along the pathway induce the secretion of pro-inflammatory cytokines and chemokines [30]. Another oncogene, MYC, has a critical role in tissue remodeling of ECM [24,31]. RET, a tyrosine kinase, can be constitutively activated by mutations even in the absence of its ligands.
As oncogenes, tumor suppressor proteins also modulate the production of pro-inflammatory elements. Transforming growth factor-1 beta (TGF-β), a tumor suppressor protein, accelerates tumorigenesis by increasing tumor-promoting inflammation. Other examples of such tumor suppressors are phosphatase and tensin homologue (PTEN) and von Hippel Lindau (VHL) proteins [24,28,32].

Tumors are not composed of solely homogenous cancer cells, but are rather heavily infiltrated by cells, both the innate and adaptive immune cells, like macrophages, and mast cells [19,23]. In addition to immune cells, epithelial cells, endothelial cells, fibroblasts and stromal cells, infiltrate tumor tissue [18]. Mast cells produce compounds that preserve chronic inflammation and support tumor development [33]. Some of the inflammatory cells, namely, regulatory T cells (TregS), myeloid-derived suppressor cells (MDSCs) and tumor-associated macrophages (TAMs) have immune suppressive roles [19,24]. MDSCs and TGF-β released by cancer cells may inhibit cytotoxic T cells and natural killer cells [19]. Inflammatory cells also secrete epidermal growth factor, EGF (tumor growth factor), vascular endothelial growth factor, VEGF (pro-angiogenic factor), and matrix degrading enzymes, which help tissue-remodeling, angiogenesis and metastasis [19].

2. TLRs, key adaptors, and key transcription factors

Toll-like receptors (TLRs) regulate both the innate and the adaptive immune systems and have an essential role in inflammation [22]. TLRs detect the presence of pathogens and initiate inflammation. TLRs belong to pattern recognition receptors (PRRs) and they recognize highly conserved motifs, pathogen-associated molecular patterns (PAMPs), such as lipopolysaccharide of gram-negative bacteria and damage-associated molecular patterns (DAMPs), such as high mobility group box-1 (HMGB1) protein or heat shock proteins [20,22,34,35]. Hence, pathogen-derived compounds and endogenous molecules of damaged tissue serve as ligands for TLRs. TLRs are single-pass trans-membrane receptors with
activates two extracellular TLRs (TLR9) almost others and the immune system, including innate immune cells, T and B lymphocytes (adaptive immune cells), endothelial cells, epithelial cells, fibroblasts and cancer cells [20]. Up to now, 11 mammalian and 13 murine TLRs have been identified [20]. Some of the TLRs are found on plasma membrane (TLR1, TLR2, TLR4, TLR5, TLR6, TLR10), whereas others are located on endosomal membranes (TLR3, TLR7, TLR8, TLR9) [20,36]. Upon binding with their ligands on extracellular LRR, TLRs form homo- or hetero-dimers, recruit some TIR domain-containing adaptor proteins to their cytoplasmic TIR domains and activate downstream signaling pathways [20]. TLRs have five TIR-containing adaptor proteins, Myeloid differentiation factor 88 (MyD88), TIR domain-containing adaptor inducing interferon-β (TRIF), adapter-related adaptor molecule (TRAM), also known as TICAM-2) [20] (see Fig. 1), and sterile α and heat-shock motifs (SARM) [37]. MyD88 is the common adaptor for almost all TLRs, except TLR3 and also IL-1R family members [38]. Mal serves as a bridging adaptor between TLR3 and TRIF. Binding surfaces of both MyD88 and TLRs are electropositive, thus they repel each other. However, Mal, with an electronegative surface, brings MyD88 and TLRs together [38–40]. Likewise, TRAM acts as a bridge between TLR3 and TRIF.

When stimulated, TLRs follow one of two pathways, as displayed in Fig. 1. In the first, MyD88-dependent pathway, TLRs associate with MyD88 and Mal proteins through their TIR domains, recruiting serine/threonine kinases IRAKs (interleukin-1 receptor-associated kinases) to stimulate the TRAF6/IκB complex and MAPKK, which result in activation of MAP kinases (such as ERK, JNK, and p38 [41,42]), N-κB and AP-1 [20,34]. If N-κB is bound to its inhibitor IκB, it is located in the cytoplasm. When IκB-kinase (IKK) phosphorylates IκB, N-κB dissociates from IκB and translocates to the nucleus to activate transcription of cytokines and chemokines, such as tumor necrosis factor-alpha (TNF-α), interleukin-6 (IL-6), IL-12, IL-1β and adhesion molecules [24,34,41,42]. However, in the second, TRIF-dependent pathway, stimulated TLR recruits TRIF and TRAM and results in the dimerization of interferon regulatory factors (IRFs), producing interferon alpha and beta (IFN-α and IFN-β) [20]. Structural pathways should be able to capture the dimerization and the altered interactions, which is not the case in the node-and-edge representation. Normal immune defenses of the host mediate NF-κB activation and cytokine and chemokine expression, leading to acute inflammation. If not finely tuned, it can contribute to the onset of cancer [20].

Fig. 3 displays the structural details of MyD88 signaling model. MyD88 interacts with Mal through its TIR domain, and IRAK4, IRAF, TRAF6, and FADD through its DD. MyD88 interactions with IRAK4, IRAF, and TRAF6 can co-exist since they use different interfaces. But MyD88 cannot interact with both IRAK4 and FADD simultaneously because they use overlapping interfaces and thus have a steric clash (Fig. 4). This means that either apoptosis or NF-κB pathway is activated by MyD88 interaction, but not both. Similarly, Fig. 5 demonstrates interactions of TRAF6. TRAF6 associates with RANK (Receptor activator of NF-κB) and CD40 through the same interface close to its N-terminal end. Thus these interactions cannot take place at the same time. In addition, TRAF6 binds to Ubc13 (ubiquitin-conjugating enzyme E2 N) through an interface, which is next to its C-terminal end. Since the 3D structure of full-length TRAF6 has not been identified yet, it is not known whether its N-terminal and C-terminal ends are close to each other. Therefore, it is still unclear whether the TRAF6 interactions with Ubc13 and CD40...
or RANK can take place simultaneously. This example stresses, once again, the necessity for full-length 3D structures of proteins and the importance of structural pathways.

Chemokines and cytokines attract tumor-promoting infiltrates (Tregs, TAMs, etc.) and facilitate angiogenesis and dissemination [24]. Some early transformed cells express chemokine receptors that contribute to their survival and invasion. For example, a chemokine receptor CXCR4 is often expressed on tumor cells and the quantity of this receptor is indicator of the metastatic potential [15,24].

Like TLR3, TLR4 also uses the TRIF-dependent pathway [43], but whether it can associate with Mal and TRAM simultaneously is unknown [44]. There are two TIR domains to be occupied, one on each TLR. TLRs dimerize when they bind their cognate ligands, again highlighting the merit of structural pathways which could model these interactions. Evidence suggests that the MyD88-dependent pathway takes place in the early stages of stimulation of the TLR4 signaling pathway. By contrast, TRIF-dependent pathway is activated at later stages [38]. TRAM can function in the absence of MyD88 [45]. Kagan et al. suggested that MyD88- and TRIF-dependent pathways are activated sequentially. They proposed a mechanism in which TLR4 on the plasma membrane first associates with Mal and MyD88, leading to NF-κB activation. Then, TLR4 is endocytosed and gets associated with TRAM and TRIF, leading to transcription of IFN-α and IFN-β [46].

TLRs are highly expressed in tumor cells and they confer chemoresistance, escape from immunity, and promote tumor growth [47,48]. For instance, MyD88 and NF-κB signaling through TLR4 on ovarian cancer cells enable these cells to escape from immunity and contribute to cancer progression [20,49]. On the other hand, there is also evidence to the contrary: TLR activation of cancer cells helps to inhibit the proliferation of cancer cells [20]. These conflicting results demonstrate the complexity of TLR action on cancer cells. A complete structural TLR pathway can help to forecast what will be the outcome: inhibition or promotion of tumor.

As can also be seen in Fig. 1, MyD88 and NF-κB are key orchestrators of TLR signaling which is important in cancer related inflammation [42]. In advanced cancers, TAMs have defective NF-κB activation [24]. Inhibition of IKK-β in myeloid lineage reduced cancer related inflammation in intestine and colitis-associated cancer [24].

Similar to NF-κB, STAT3, another transcription factor, is constitutively active in cancer and immune cells and participates in driving inflammation to cancer. It promotes survival, proliferation, and metastasis of cancer cells [21,24]. STAT3 down-regulates...
tumor suppressor protein p53, and thereby inhibits apoptosis. It also stimulates VEGF and metalloprotease expressions, increasing angiogenesis and tissue rearrangement [21].

Below, we describe how we go about the structural modeling of these pathways and multimolecular associations related to them, such as that of the key scaffolding protein assembly of MyD88 (Fig. 3).

3. An overview of the rationale and method for constructing structural pathways

Constructing pathways [1,2,50,51], in particular structural pathways, is challenging. First, experimental structural data are limited, which necessitates exploiting homology models. Homology models lead to increased noise in the predictions [52]. Second, the accuracy of high-throughput data relating to interactions is unclear. Third, ideally, conformational changes in the protein that take place following allosteric events should be taken into account. For example, binding of ligands to the LRR facing the endosomal lumen of Toll-like receptor TLR9 dimer induce allosteric changes in the cytoplasmic TIR domains of the dimer [53]. These conformational changes are responsible for activating TLR9 signaling rather than receptor dimerization [53]. The presence of post-translational modifications (PTMs) such as phosphorylation, methylation, and acetylation may also lead to conformational change. For instance, IRAK1 goes through a conformational change following phosphorylation by IRAK4 [54]. In the unphosphorylated state, IRAK1 is in a closed-inactive state and cannot associate with MyD88 through its death domain and TRAF6 through its C-terminal domain. When it is phosphorylated by IRAK4, it undergoes a conformational transition to an active open form. In the active state, IRAK1 can associate with both MyD88 and TRAF6 [54]. In another example, C/EBP protein, which plays an important role in inflammatory cytokines transcription [55], undergoes a conformational change upon phosphorylation and gets activated [56]. Conformational changes, regardless of their origin – allosteric binding or posttranslational modifications – may alter the interactions with partner proteins. Fourth, a structural pathway does not provide the protein availability at any given time; for example, some proteins are expressed only during certain phases of the cell cycle. Similarly, protein localization can also come into play. If a protein is restricted to the nucleus, it cannot interact with another protein present in the mitochondria.

Computational strategies to predict interactions between proteins, and thus to construct structural pathways, include docking and template-based techniques (reviewed elsewhere [2]). In most protein–protein docking methods, proteins are treated as rigid bodies in their native form. Consequently, prediction of interactions via docking approaches is often poor [52]. A major hurdle in the prediction of protein–protein interactions is that a large number of solutions are typically obtained, and in the absence of additional biochemical data about the location of the binding site, it is very difficult to distinguish between the native and other interactions. Modeling of pathways encounters the additional difficulty of the combinatorial complexity, since a large number of proteins would have to be sampled. ‘Traditional’ docking approaches may not work for three reasons: first, they will produce too many false positive solutions, since they will always find complementary surfaces; second, their time requirements prohibit their application on a large scale; and third, rigid body docking considers only minor flexibility.

Template-based PPI prediction techniques make use of previous interface knowledge. In nature, there are thousands of different proteins with diverse 3D structures, but the binding surfaces or interfaces are more conserved compared to the global structures.
of proteins [57–61]. That is, different proteins use the similar interfaces to interact with their partners (as displayed in Fig. 6); thus, previously identified interfaces from structures of protein complexes can be used to predict novel interactions. If the surfaces of two proteins of interest structurally match with the template interface, theoretically these two proteins can interact. Template-based prediction methods take much less CPU time than docking and they can be used for pathway-, or proteome-scale predictions. The success of template-based prediction depends on whether the interface is present in the template set.

To address the modeling of the inflammation and cancer network, we propose to use PRISM (PRotein Interactions by Structural Matching) [16,17], a motif-, knowledge-based algorithm. PRISM exploits our observation that protein–protein interfaces are conserved, and consist of recurring architectural motifs, much like the motifs observed in single chain proteins. In particular, interfaces can be conserved even when the global structures and the functions of the proteins are entirely different. Some examples for shared and distinct interfaces are given in Fig. 7.

The difficulty in the prediction is compounded by the fact that a particular protein can interact with other proteins via the same or different interfaces (Fig. 7). If a protein interacts with its two partners through the same interface, these two interactions are mutually exclusive. Likewise, if the interfaces overlap, those interactions are also mutually exclusive. Conversely, if two partners bind to a particular protein through different interfaces, with no overlap, these interactions can co-exist [62–64].

As we noted above, hub proteins interact with tens or hundreds of partners. Due to the limited surface area of a protein, a single protein cannot interact with all of its partners simultaneously. Structural pathways provide the knowledge of which interactions can take place simultaneously and thus the respective pathways co-activated/inhibited, and which are mutually exclusive [62]. Even though the structures of many of these proteins are available in the PDB, and of some in multiple forms (e.g. IRF3 and MAPK), their interactions provide the cellular network, which ultimately is how the cell is regulated. Fig. 7 depicts some of these key proteins and their interactions.

4. Conclusions

Structural pathways are important. They are essential to the understanding of how oncogenic mutations work and to figuring out alternative parallel pathways in drug resistant mutants. Structural pathways also help to understand the inter-relationship among linked phenomena, as in the case of inflammation and cancer. Cell biology provides a global overview of the behavior of the cell, tissue and the organism under different sets of conditions; the structures of single proteins and their coherent interactions provide insight into the dynamic changes in the proteins, such as those taking place through post-translational modifications, binding events and mutations, and into their interactions. While not addressed in this review, it behoves us to recall that beyond the challenging construction of structural pathways, there is also a need to obtain a mechanistic insight into single proteins, their modifications, interactions and broadly, their changing landscapes. Why is insight into the dynamic landscape of single proteins important? Perceiving protein’s behavior can help to forecast allosteric transitions, and regulation; it can help relate oncogenic mutations to their constitutive consequences. Oncogenic mutations bypass the autoinhibited state by shifting the population from the physiologically favored inactive state to an active state. In principle, this can take place via three mechanisms: first, by stabilizing the active conformation, as in the case of the T790M mutation in EGFR which promotes the assembly of an enzymatically active kinase conformation by
stabilizing the hydrophobic R spine; second, by disrupting critical interactions in the inactive conformation, thus destabilizing it with respect to the active state, as in the case of the L858R mutation in the hydrophobic core of inactive EGFR, which leads to the shift of the population toward the active conformation; or third, by a combination of both mechanisms, as in the case of L858R which also stabilizes the c-on-Helix in inactive conformation from an intrinsically disordered structure through heterodimerization, and thus shifts the population even in the absence of ligand-induced receptor dimerization [3]. Such insight into the kinase landscape can help to classify oncogenic mutations, and thus identify their origin of their oncogenic mechanisms.

Structures, their associations, and their conformational ensembles are critical to fully understand the hallmarks of cancer, pathways to cancer, the loss of control through constitutive mutations and the bevy of changing interactions; they are key to reveal and understand in detail the steps that drive a normal cell to become a cancer cell.

Here, we focused on the link between inflammation and cancer, and how we go about constructing the structural pathways as a paradigm for mapping cancer signaling pathways in the cell, which is a major on-going project in our lab.

Conflict of interest statement

The authors declare that there are no conflicts of interest.

Acknowledgements

This project has been funded in whole or in part with federal funds from the National Cancer Institute, National Institutes of Health, under contract number HHSN261200800001E. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government. This research was supported (in part) by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research. O.K. acknowledges the support of Science Academy (of Turkey).

References
