33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Modeling signaling pathways in articular cartilage

Ioannis N. Melas, Aikaterini D. Chairakaki, Alexander Mitsos, Zoe Dailiana, Christopher G. Provatidis and Leonidas G. Alexopoulos

Abstract—**Chondrocytes, the only cell type in articular cartilage, are responsible for maintaining the composition of cartilage extracellular matrix (ECM) through a complex interplay of anabolic and catabolic stimuli. Although understanding the way chondrocytes respond to stimuli is of utmost importance for shedding light into the etiology of joint diseases, an integrative approach to studying their signaling transduction mechanisms is yet to be introduced. Herein, we propose an approach that combines high throughput proteomic measurements and state of the art optimization algorithms to construct a predictive model of chondrocyte signaling network, downstream of 78 receptors of interest.**

I. Introduction

Articular cartilage is an avascular, aneural connective tissue that covers the ends of bones and transfers loads with minimal wear and friction. Chondrocytes, the only cell type found in cartilage, are responsible for maintaining the extracellular matrix (ECM) composition. Their function is orchestrated by a complex signaling network and a variety of environmental perturbations. In pathological cases (e.g. osteoarthritis) an imbalance of anabolic and catabolic processes occurs, resulting in degradation of the ECM and eventually, total destruction of the joint [1]. Modeling of chondrocytes' signaling mechanisms is a major endeavor for academia and pharmaceutical industries as they attempt to either block pro-inflammatory pathways leading to cartilage degeneration, or stimulate pro-growth pathways compensating for the loss of ECM's structural integrity.

The study of signaling in chondrocytes is traditionally performed in a reductionist way, by modeling the effects of a few well known stimuli (i.e. IL1a/b - Interleukin 1 alpha/beta, TNFa-Tumor Necrosis Factor alpha, TGFa - Transforming Growth Factor alpha, etc.) on few downstream proteins (i.e. NFkB - Nuclear Factor kappa beta, ikb, etc.) and subsequently on cells' phenotype and tissue development or degradation [2-9]. However, discovery of therapeutic interventions requires a deeper understanding of cell's machinery that cannot be accomplished by studying each cascade independently [10-12]. A systems-level approach needs to be implemented to address key issues such as robustness, redundancy and modularity that play key role in cellular function and determine cartilage homeostasis. Starting point to this endeavor is the construction of a chondrocyte-specific signaling network, that will capture the signaling reactions taking place in human chondrocytes and form the basis for developing cartilage-specific therapeutic interventions..

Herein, we aim at constructing a chondrocytes specific signaling network, based on high throughput proteomic data and a priori knowledge of protein connectivity. To that effect, chondrocytes isolated from a single donor are placed in a 96 well plate and stimulated with single treatments of 78 stimuli. The activation state of 17 key phosphoprotein signals is measured using xMAP technology [14] resulting in a dataset of more than 1300 data points. When this dataset is compared against canonical pathways obtained from the literature, discrepancies are observed attributed to the generic nature of the canonical pathways. Therefore, an optimization scheme is implemented, built on a Boolean framework, to prune the canonical pathway so it best fits the dataset at hand [15,16]. An objective function is introduced, consisting of two terms; the first one quantifies the mismatch between experimental measurements and pathway predictions; and the second one corresponds to the map's size. An Integer Linear Programming (ILP) formulation is used to remove all reactions from the canonical pathway that appear not to be functional in the interrogated cell type [16], thus minimizing the objective function:

$$\sum a_j^k \mid x_j^k - x_j^{k,m} \mid + \sum b_i y_i \qquad (1)$$

where,

- $j=\{1, \ldots, n_s\}$ are the species (nodes) of the pathway,
- $i=\{1, \ldots, n_r\}$ is the set of reactions,
- $k=\{1, \ldots, n_e\}$ is the set of experimental conditions
- x_j^k is the simulated value of species j in experiment k
- $x_j^{k,m}$ is the measured value of species j, experiment k
- a_j^k are user-defined constants (for species j in experiment k)
- y_i is the presence (or absence) of reaction i (assumes only Boolean values)
- b_i are user-defined weights (for reaction i)

Complete formulation can be found in [16].

The optimized map includes only the reactions in accordance to the experimental dataset and emulates the functional characteristics of chondrocytes' signaling network. The research presented in this paper is amongst the first attempts to construct an integrative model of chondrocytes signaling mechanisms and study the cross-talks developing downstream of 78 receptors of interest.

Manuscript received April 14, 2011.

Ioannis N. Melas, Aikaterini D. Chairakaki, Christopher G. Provatidis and Leonidas G. Alexopoulos are with the National Technical University of Athens, Mechanical Engineering Department, Heroon Polytechneiou 9, Zografou 15773, Athens, Greece.

Alexander Mitsos is with the Massachusetts Institute of Technology, Mechanical Engineering Department, 77 Mass Ave, MIT3-158, Cambridge, MA, 02139 USA.

Zoe Dailiana is with the University Hospital of Larissa, Department of Orthopedic Surgery, Viopolis 41110, Larissa, Greece

Correspondence should be addressed to Leonidas G. Alexopoulos Tel +30 210 7721666. email : leo@mail.ntua.gr. Web: http://users.ntua.gr/leo

II. Results

A. Phosphoproteomic dataset

A library of 78 stimuli was screened (see Figure 1) and the phopsphorylation levels of the following 17 signals were measured: AKT, JUN, CREB, ERK, GSK3, HISTH3 (Histone H3), HSP27 (Heat Shock Protein 27), IKB, IRS1S, JNK, MAP2K1, MAPK14, TP53.

Chondrocytes do not respond to all stimuli in the library, however 18 cytokines (DEFB1, BTC (Betacellulin), BMP4 (Bone Morphogenetic Protein 4), EGF (Epidermal Growth Factor), FGF2 (Fibroblast Growth Factor 2), HBEGF (Heparin-Binding EGF-like Growth Factor), IGF1 (Insulin-like Growth Factor 1), IL19, IL1A, IL1B, IL6, INS (Insulin), NOG (Noggin), ODN2006, FLAGELLIN, TGFA, TGFB1 (Tumor Growth Factor B1), TNF) are identified as having a clear effect on some of the 17 signals. As positive control we observe that TGFA and EGF activate MEK, INS and IGF1 activate AKT; IL1B and TNF activate IKB, and IL6 activates STAT3, observations already reported in literature. Our screening procedure reveals some other major players: BMP4 activates HSP27, BTC and EGFR ligand activate MAP2K1 and CREB, FGF2 and HBEGF activate IGF1, IL19 activates HSP27, MAPK14 and CREB. In adittion, some less known players have been identified: FLAGELLIN, a Toll Like Receptor 5 (TLR5) ligand activates MAPK14, MAP2K1, IKB, HSP27, CREB and JUN; DEFB1, a member of the defensin family implicated in the resistance of epithelial surfaces to microbial colonization activates CREB and HSP27; NOG, a protein that inhibits TGF-B signal transduction by binding to TGF-B family of ligands activates CREB and HSP27; ODN2006, an oligodeoxynucleotide activates JUN.

B. Pathway optimization

Out of the original 1064 reactions present in the canonical pathway, just 67 of them are conserved by the ILP algorithm. The optimized map is plotted in Figure 2. As positive control observations, BTC, EGF and TGFA signal through the EGFR and activate AKT, MAP2K1, CREB and GSK3 in accordance to the dataset of Figure 1. TNF and IL1B signal though the MAP3Ks and TRAF6 and activate IKB, MAPK14, JUN, HSP27, CREB and MAP2K1; IL6 activates STAT3 via JAK2, again in general accordance to the dataset in Figure 1 and the literature. Moreover, INS and IGF1 signal via pathways similar to TGFA, BTC, EGF and HBEGF. FGF2 signals via the FGFR and activates MAP2K1, CREB. Concerning the novel players identified previously (FLAGELLIN, DEFB1, NOG and ODN2006): FLAGELLIN signals via TRAF6 in similar fashion to IL1B and activates IKB, MAPK14, JUN, HSP27, CREB, MAP2K1; DEFB1 signals through TLR4 and activates HSP27 and CREB; while NOG and ODN2006 are not present in the optimized map, since there was limited information on these ligands in the pathway databases and we could not correlate them with any of the activated signals

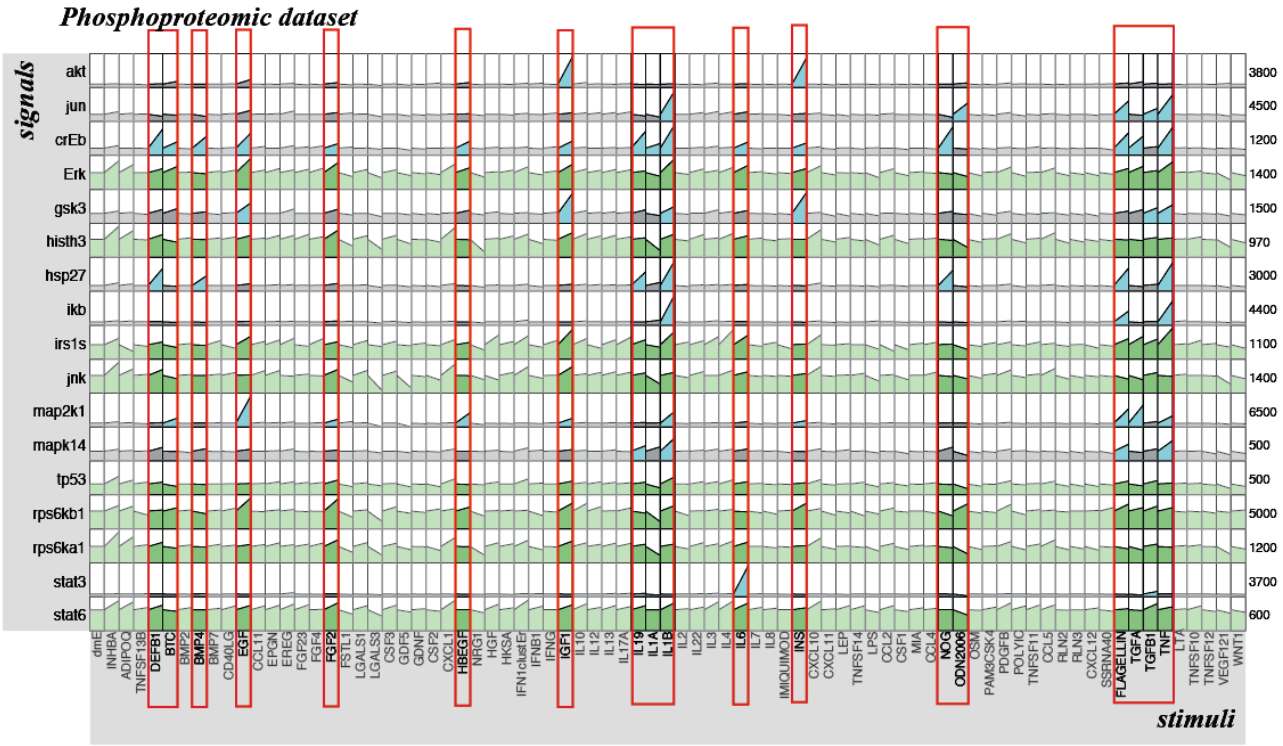

Figure1 – High throughput phosphoproteomic dataset. Columns correspond to imposed stimuli, rows correspond to measured signals. Each subplot represents the time course of each measurement, only 2 time points are incorporated in this dataset (stimulated, unstimulated). The color of each subplot represents the activation state of each measurement, grey for deactivated, blue for activated, green when the signal/noise ratio is too low to distinguish between activated and deactivated state. The values on the right stand for the maximum phosphorylation values (in fluorescent units) for each signal.

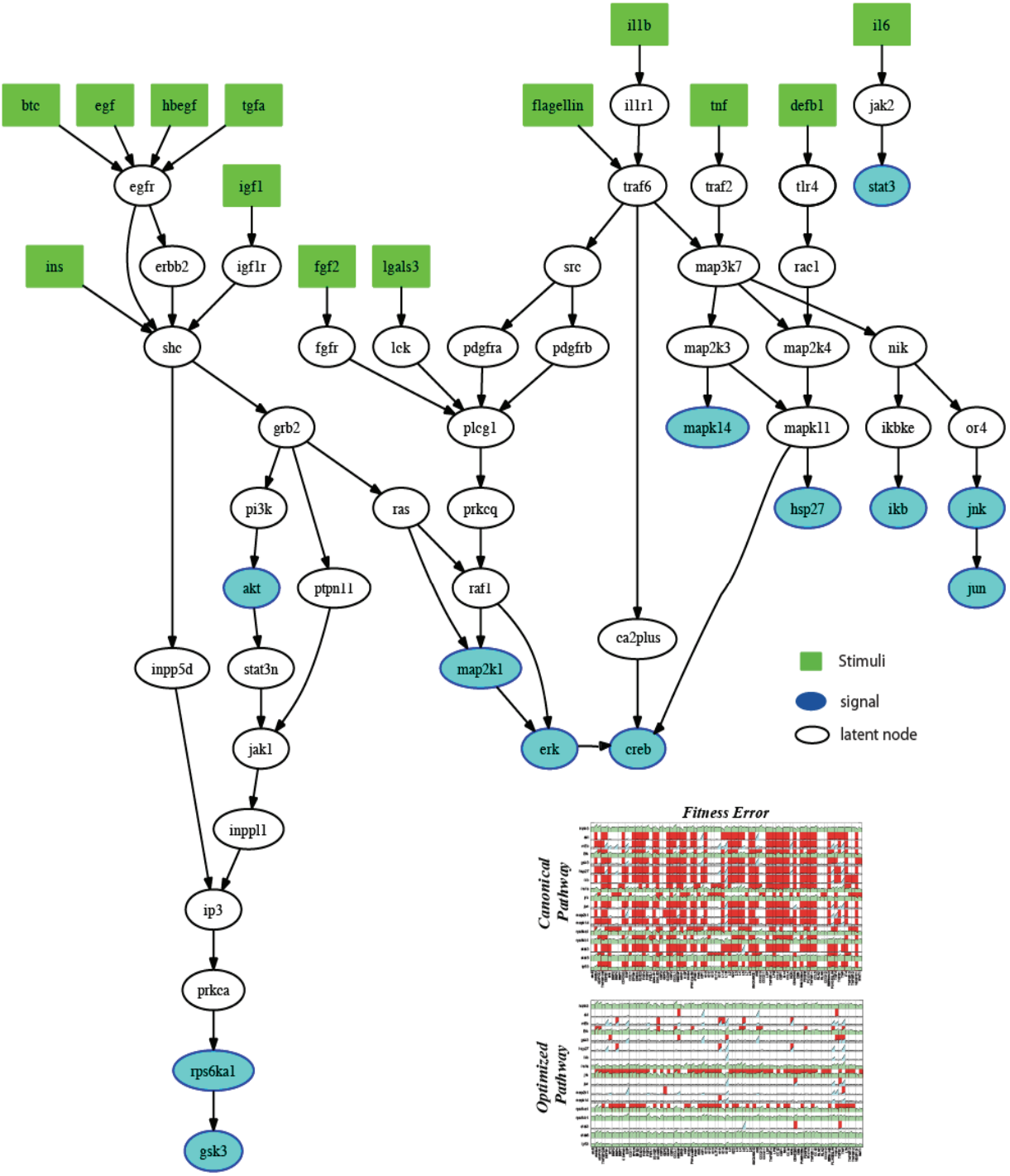

Figure 2 –Pathway after optimization with phosphoproteomic data (Fig.1). Insert: Missmatch error (red background) between experimental data before (top panel) and after (bottom panel) ILP optimization.

MATERIALS & METHODS

C. Chondrocytes' isolation – Proteomic assay

Cartilage tissue was obtained from the femoral head of human donor according to standardized protocols [17]. Briefly, human cartilage tissue obtained under approved protocols was digested with collagenase and pronase overnight and the isolated chondrocytes were placed in collagen coated 96 well plates at 40,000 cells per well.

Chondrocytes were stimulated by single treatments of 78 cytokines and for each sample we collected cell lysates at 20 minutes. The phosphorylation level of 17 key proteins (AKT, JUN, CREB, ERK, GSK3, HISTH3, HSP27, IKB, IRS1S, JNK, MAP2K1, MAPK14, TP53, RPS6KB1, RPS6KA1, STAT3, STAT6) was measured using the bead-based ELISA type assays (xMAP technology, Luminex Inc). The signals were chosen based on the availability of the reagents and quality controls performed at the early phases of the experimental setup.

D. Data normalization

Phosphorylation activity, measured via xMAP technology, is measured in arbitrary fluorescent units that is dependent on the antibody pair used for detection. For instance, MAP2K1 ranges from 280 units to 6500 units (under EGF), while GSK3 ranges from 500 units to 1500 units. Variations such as these may be due to protein abundance or assay calibration issues and do not necessarily reflect that MAP2K1 is more activated than GSK3. Consequently, two challenges emerge from this variability. First identifying whether a signal is activated or not, and second, normalizing the data in a manner that the ILP is not biased in favor of the higher values.

Herein we introduce the following procedure: (i) A bimodal distribution is assumed for the phosphorylation values of each measured signal consisting of an activated mode and a deactivated mode. The bimodal distribution is calculated using Matlab statistical toolbox; (ii). For each datapoint the frequencies respective to the two modes are evaluated. (iii) Their ratio is normalized to the range [0,1] based on a hill function filter and subsequently the normalized value is imported to ILP (i.e. if a raw datapoint falls solely in the activated distribution curve, then it assumes a normalized value of one).

E. Construction of the canonical pathway

The canonical pathway is constructed downstream of 78 receptors of interest and in the neighborhood of 17 measured proteins. KEGG and Ingenuity were used for obtaining the reactions of the generic topology. The HUGO Gene Nomenclature Committee database (http://www.genenames.org) was used to ensure standardized protein names. The pathway was built and visualized using Graphviz (http://www.graphviz.org/). The resulting topology consists of 533 species and 1064 reactions. A Boolean framework was used to simulate signal transduction upon stimulation. Proteins are connected via logic gates (AND/OR/NOT) and assume only Boolean values (0/1) [16]. Starting at the receptor level, the signal is propagated based on the network topology and the activation state of downstream proteins is evaluated.

III. Discussion

In this paper we have successfully combined state of the art optimization algorithms (ILP), solved by commercial solvers (CPLEX- http://www.ilog.com/products/cplex/) and high throughput proteomic measurements to construct a functional, predictive model of chondrocytes signaling network. Proteomic measurements were performed using the xMAP technology. A library of 78 stimuli was used and the activation state of 17 key phosphoproteins was measured. Cells from a single donor were used. Data with phosphoproteomic activity was then used by the ILP optimization algorithm to prune a canonical pathway of 533 nodes and 1064 reactions so it best fits the phosphoproteomic responce of chondrocytes. As a result we have validated previously reported signaling motifs such as TNF and IL1B induced IKB activation; IL6 induced STAT3 activation; TGFA, EGF, INS and IGF1 induced AKT and MAP2K1 activation. Unknown players were also identified such as FLAGELLIN, a TLR5 ligand and DEFB1, a member of the defensin family. FLAGELLIN was found to signal in very similar way to IL1B, activating IKB, MAPK14, JUN, HSP27, CREB and MAP2K1. DEFB1 activates HSP27 and CREB. Both of these findings indicate that chondrocytes assume an active role in immuno-surveillance. In contrast to reductionist approaches mostly used today, the approach presented herein views the signaling network in a holistic manner, and draw conclusions about chondrocytes function on a systems level.

IV. References

[1] Guilak, F., et al., *The role of biomechanics and inflammation in cartilage injury and repair.* Clinical Orthopaedics and Related Research, 2004(423): p. 17-26.

[2] Visse, R. and H. Nagase, *Matrix metalloproteinases and tissue inhibitors of metalloproteinases - Structure, function, and biochemistry.* Circulation Research, 2003. **92**(8): p. 827-839.

[3] Goldring, M.B. and S.R. Goldring, *Osteoarthritis.* Journal of Cellular Physiology, 2007. **213**(3): p. 626-634.

[4] Palmer, A.W., et al., *Composition-function relationships during IL-1-induced cartilage degradation and recovery.* Osteoarthritis and Cartilage, 2009. **17**(8): p. 1029-1039.

[5] DeLise, A.M., L. Fischer, and R.S. Tuan, *Cellular interactions and signaling in cartilage development.* Osteoarthritis and Cartilage, 2000. **8**(5): p. 309-334.

[6] Hartmann, C. and C.J. Tabin, *Dual roles of Wnt signaling during chondrogenesis in the chicken limb.* Development, 2000. **127**(14): p. 3141-3159.

[7] Goldring, M.B., K. Tsuchimochi, and K. Ijiri, *The control of chondrogenesis.* Journal of Cellular Biochemistry, 2006. **97**(1): p. 33-44.

[8] Roman-Blas, J.A. and S.A. Jimenez, *NF-kappa B as a potential therapeutic target in osteoarthritis and rheumatoid arthritis.* Osteoarthritis and Cartilage, 2006. **14**(9): p. 839-848.

[9] Liacini, A., et al., *Inhibition of interleukin-1-stimulated MAP kinases, activating protein-1 (AP-1) and nuclear factor kappa B (NF-kappa B) transcription factors down-regulates matrix metalloproteinase gene expression in articular chondrocytes.* Matrix Biology, 2002. **21**(3): p. 251-262.

[10] Kitano, H., *Systems biology: A brief overview.* Science, 2002. **295**(5560): p. 1662-4.

[11] Boccaletti, S., et al., *Complex networks: Structure and dynamics.* Physics Reports-Review Section of Physics Letters, 2006. **424**(4-5): p. 175-308.

[12] Aggarwal, K. and K.H. Lee, *Functional genomics and proteomics as a foundation for systems biology.* Brief Funct Genomic Proteomic, 2003. **2**(3): p. 175-84.

[13] Fisher, J. and T.A. Henzinger, *Executable cell biology.* Nature Biotechnology, 2007. **25**(11): p. 1239-1249.

[14] Alexopoulos, L.G., et al., *Networks Inferred from Biochemical Data Reveal Profound Differences in Toll-like Receptor and Inflammatory Signaling between Normal and Transformed Hepatocytes.* Molecular & Cellular Proteomics, 2010. **9**(9): p. 1849-1865.

[15] Saez-Rodriguez, J., et al., *Discrete logic modelling as a means to link protein signalling networks with functional analysis of mammalian signal transduction.* Molecular Systems Biology, 2009. **5**:331.

[16] Mitsos, A., et al., *Identifying Drug Effects via Pathway Alterations using an Integer Linear Programming Optimization Formulation on Phosphoproteomic Data.* Plos Computational Biology, 2009. **5**(12).

[17] Guilak, F., et al., *The effects of matrix compression on proteoglycan metabolism in articular cartilage explants.* Osteoarthritis and Cartilage, 1994. **2**: p. 91-101.