

Proceeding Paper

# Information Processing in Living Cells: Mapping a Logic Gate to a Gene Regulation Event <sup>†</sup>

Tilmann Wurtz 

Independent Researcher, 14653 Tullinge, Sweden; tilmann.wurtz@yahoo.com

<sup>†</sup> Presented at the Conference on Theoretical and Foundational Problems in Information Studies, IS4SI Summit 2021, online, 12–19 September 2021.

**Abstract:** In living cells, signal chains process the information that flows between the signal reception and the gene usage. This paper investigates the conjecture that logic gates play a role. In the mammalian preimplantation embryo, cells either become part of the placenta, or part of the embryo proper; i.e., they enter either the trophectoderm (TE) line, or the inner cell mass (ICM) line. TE cells express a caudal-related homeogene (CDX2), whose induction is regulated by two signal chains. A logic AND gate was mapped to the regulative DNA region of the CDX2 gene, which integrated the signaling from both chains. The gate accounts for all of the available experimental findings, which suggests that it is instrumental for CDX2 induction.

**Keywords:** intracellular signal chains; information processing; preimplantation embryo; logic gates; interactive networks

## 1. Introduction

Life is characterized by the usage and generation of information, and by cells and organisms that embody information. Information flow has continued from the origin of life, when RNA molecules degraded and synthesized each other [1], up till today, with all of the diversity of the biosphere. Organisms harbor different systems of data processing that are associated with cells, tissues, individuals, and social interactions [2]. Cells, the building blocks of all tissues, function as computing machines: they receive signals, process them, and respond by making biologically active molecules. Signal input occurs via specific receptors that recognize cell contacts, secondary messengers, hormones, and growth factors. Receptor molecules encode signals into intracellular information carriers: the signal chains [3]. The signal chain components address the genome. Genes serve as memory, from which information can be retrieved for the protein synthesis that is required in a given context. The decoding of information occurs when new gene products are integrated into the cell body, or when they are secreted for construction, metabolism, or for signaling to other cells.

There are many receptors that are specific for various cues, as well as many signal chains. Since signal chains interact (crosstalk), the combined signal chains represent a network that is able to process data [3]. The signals are transferred between the chain components by sequential allosteric interactions and/or chemical modifications. The signal chain components in cells may be active or inactive, and an active component of one chain may stimulate or interrupt the signal transmissions of other chains. Sometimes, a chain component needs several events to be activated, but, in the end, the signal transmission is binary: the components are in an ON or OFF state. Most signal chain components are not freely diffusible; they are arranged in assemblies that are called, “signalosomes” [4].

The purpose of this work is to study the cellular information flow as a computing system, with technical information processing as a reference.



**Citation:** Wurtz, T. Information Processing in Living Cells: Mapping a Logic Gate to a Gene Regulation Event. *Proceedings* **2022**, *81*, 157. <https://doi.org/10.3390/proceedings2022081157>

Academic Editor: Mark Burgin

Published: 29 August 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 2. Hypothesis

The circuits for electronic data processing are composed of diodes and transistors, which are grouped in logic gates. These gates combine and filter input data according to Boole's algebra. The calculations are thus automatized in a binary fashion. Since cells also propagate and process signals in a binary way, the hypothesis is forwarded that logic-gate-like mechanisms can operate in cells too. A test for this idea is to map logic gates to biological events. Some of the enzymes and elements of signal chains have already been associated with logic gates; any factor that is activated or silenced, or allosterically modified with two or more input signals and one output signal, would allow that. For example, different mitogen activated protein kinases (MAPK) are associated with multiple signaling chains ([en.wikipedia.org](https://en.wikipedia.org)). They are activated by MAPK-activating factors (MAPKK), which are themselves activated by the dual phosphorylation by other factors (MAPKKK). MAPKK may thus constitute AND or OR gates in the signaling network [3]. However, the mere possibility of constructing a logic gate does not map a logic gate to a biological event. The following criteria are set as the attributes for a signal chain component that is to be mapped to a logic gate with a biological function:

- It integrates in a logic-gate-like structure;
- It is part of a network;
- It is involved in defined biological decision making;
- The gate helps to interpret the experimental findings.

With regard to decision making, biological functions are usually analyzed by the mutation of the factor in question, and by the observation of the effect on an associated process. Many components of signal chains have been identified in cancer cells that are genetically mutated or misexpressed. They are part of networks and they may contribute to logic-gate-like structures. However, mapping these components to logic gates in cancer formation is hampered because of the multitude of possible reactions with unclear significances. For example, if an oncogene is mutated to become a constitutively active transforming growth factor (TGF-beta), or a RAS-GTPase (which is named after the rat sarcoma virus), there are several candidate signaling pathways ([www.sinobiological.com](http://www.sinobiological.com), accessed on 15 October 2021), and it is unclear a priori which one is relevant to the cancer formation. In addition, any mapping has to consider the networking of the gates. No such gates have been mapped so far, so that the impact of any newly mapped gate is difficult to assess.

## 3. The Model

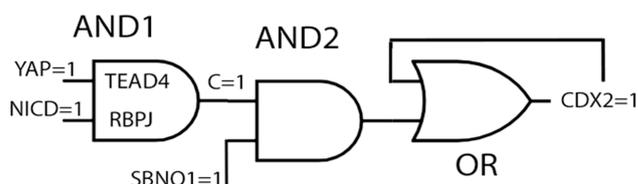
Two considerations may help to access an intracellular network for the identification and mapping of logic gates: First, the transcription regulation site of a specific gene seems more easily accessible than the signal reception or the intermediate signaling. There are only two copies of each gene in a diploid cell, and the analysis concerns the final steps of the signal chain. There, the biological effect can be observed without considering the downstream gates. The mapping of the gates that are situated in the middle of a signal chain is much more difficult. Second, we may presume that signal processing systems develop as a part of an organism: in early embryos, little cell interaction occurs, and little processing is needed, whereas complex tissues require elaborate systems.

The model that is described here is based on a gene that is induced in the preimplantation embryos of mammals during the first cell differentiation [5,6]. The fertilized egg has divided three times, which produces eight equivalent cells inside the zona pellucida. At this stage, the cells stick closer together (contraction stage), whereupon a cell polarity is established. The outside of the cells points to the zona pellucida, and the inside is in contact with other cells. Further cell divisions proceed according to two alternatives: symmetric divisions, which occur longitudinally along the outside–inside axis and which maintain the polarities of both daughter cells; and transversal asymmetric divisions, which render one daughter cell nonpolar because this cell is surrounded by other cells, with no zona pellucida contact. Further along in their development, polar cells become trophoblast

(TE) cells, and nonpolar cells become inner cell mass (ICM) cells. The TE cells end up in the placenta, and the ICM cells end up in the embryo, including in the yolk sac. Several markers have been identified that are induced during this positioning and that consolidate the cell fate: octamer transcription factor 4 (OCT4) and sex determining region Y (SOX2) for the ICM cells; and CDX2 and DNA sequence GATA-binding protein (GATA3) for the TE cells. Here, the regulation of the gene, CDX2, is described. Initially, there were two models: The first was that it could be the outer cell position that triggers the CDX2 expression, and the second was that it could be the polarization. Moreover, redundant regulation was suggested, with both signals working independently [7].

The position and polarization are sensed by two signaling pathways: NOTCH (a transmembrane receptor, which refers to the notch mutation in *Drosophila*), and HIPPO (a protein kinase, which refers to the hippopotamus mutation in *Drosophila*) [5,6]. The cell surface protein NOTCH concentration is higher in peripheral cells, and the binding of the ligands induces the release of the NOTCH intracellular domain (NICD) as the signal carrier. The NICD enters the nucleus and binds to a recombination signal-binding protein (RBPJ), which is part of the TE enhancer of CDX2. The cell polarization is sensed by the HIPPO pathway. One of its mediating factors is protein angiominin (AMOT), which is concentrated in the polar cells that are close to the zona pellucida. There, abundant actin fibers sequester AMOT, and the HIPPO pathway is interrupted. In nonpolar cells, AMOT is not sequestered and the HIPPO signaling is active. Active HIPPO signaling leads to the phosphorylation of the transcription factor, “yes-associated protein” (YAP), which, in this form, cannot enter the nucleus. Thus, in polar cells where AMOT is sequestered, the YAP enters the cell nucleus and binds to its target at the trophectoderm enhancer of the CDX2 gene, the TEA domain protein (TEAD4). Nonpolar cells do not allow the YAP to enter the nucleus, and CDX2 is not expressed. Taken together, the TE enhancer of the CDX2 gene contains the factors, TEAD4 and RBPJ, which are the targets for the transcription factors, YAP and NICD. CDX2 is induced if both YAP and NICD are present. It has been proposed that the TE enhancer of the CDX2 gene constitutes a logic AND gate [2]. The AND gate channels two input signals from the NOTCH (cell position) and the HIPPO (cell polarity) into one output, which is the expression of the CDX2 gene (“AND1” in Figure 1).

There is another factor that contributes to the regulation and expression of CDX2: strawberry notch homolog (SBNO1) [8]. This protein is a chromatin component that also binds to TEAD4 and RBPJ. The absence of SBNO1 does not interfere with the presence of TEAD4 and RBPJ, and, therefore, it is assumed that its action is downstream of these factors. If SBNO1 is absent, CDX2 is not expressed, and the embryos die before implantation [8]. It is here further assumed that it constitutes a variable for another AND gate, but there are much less data about its action compared with TEAD4/YAP and RBPJ/NICD (“AND2” in Figure 1). The CDX2 protein is itself an inducer of CDX2 gene transcription [9]. Since both YAP/NICD/SBNO1 and CDX2 can induce CDX2, an OR gate is suggested as the mode of action (Figure 1). However, it has not been proven whether the CDX2 protein can induce its own gene independently of the other factors.



**Figure 1.** Logic gates of the trophectoderm enhancer of the CDX2 gene. The TE enhancers in cells of the >8-cell embryos contains the nuclear targets, TEAD4 and RBPJ, as part of the AND1 gate. The signal chain factors, YAP (part of HIPPO) and NICT (part of NOTCH), are variable gate input factors. Positive input triggers the output, C = 1. The ensuing AND2 gate induces the transcription of CDX2, if the gate input is C = 1 and SBNO1 = 1. If YAP = 0, NICT = 0, or SBNO1 = 0, there will be no CDX2 induction. The CDX2 protein induces gene expression in a positive feedback loop [9].

#### 4. Evaluation of the Model

Notably, as a test for the involvement of NOTCH signaling in preimplantation development, the target, RBPJ, has been abrogated [10]. These null mutant embryos were not viable, but they died much later than implantation (namely, in mid-gestation). This suggests that the NOTCH signaling was dispensable for the induction of CDX2 during the TE differentiation. On the other hand, the NOTCH signaling was active during the TE differentiation, and the genetic constructs that lacked TEAD4 (which is necessary for CDX2 induction by YAP) could be rescued by NOTCH [11], which both suggest that NOTCH signaling does play a role. Furthermore, in 50% of all embryos, there are single cells with outside locations that are not polarized, and that are without the YAP in the cell nucleus (HIPPO active) [6]. These cells end up in the ICM, which suggests that an outer position is not sufficient for a TE fate, as absent polarization commits the cells to the ICM.

Taken together, different interpretations have previously been forwarded for the induction of CDX2:

1. The straight cooperation of the HIPPO and the NOTCH [6,11], where the missing phenotype of the RBPJ deficiency is not explained;
2. A redundant regulation, where either the HIPPO or the NOTCH is sufficient for the formation of an OR gate [7]. This model does not explain the cell fate of the nonpolarized surface cells;
3. Only HIPPO signaling is important, and not NOTCH [10]. Here, the rescue of the missing TEAD4 by the NOTCH is not taken into account [11].

All of the experimental findings are explained as if the HIPPO and NOTCH signals are taken as the input in a logic AND gate. The gate consists of the CDX2 trophoctoderm enhancer, plus the factors, TEAD4 and RBPJ, which are the targets of the YAP and the NICD. The YAP and the NICD are variables; if both are bound, the output = 1, and the gate is open, which allows transcription. In all other situations, the gate is closed. In the case of the absence of RBPJ, which is due to the KO mutation, the gate changes character, and it no longer has two inputs. Such a gate is called a “buffer gate”; i.e., the output has the same quality as the input, and the presence of YAP is sufficient to trigger an output = 1. In other situations, without the effector, both the TEAD4 and RBPJ may repress CDX2 expression and may prime the cells to the ICM. This gate model can account for the behavior of the RBPJ mutant. Accordingly, the missing phenotype of the RBPJ null mutant was not conclusive for the question as to whether NOTCH signaling contributed to the CDX2 induction in the wild type.

In terms of the fate of the nonpolarized surface cells, their CDX2 enhancers receive the NOTCH signal, but the HIPPO signal tells them, “no polarization”. Only one of the two gate inputs is positive. Since it is an AND gate, the output is  $C = 0$ , CDX2 is not induced, and the cells end up in the ICM.

In sum, the described AND1 gate is mapped to the process of the regulation of CDX2 expression by the following criteria:

1. The YAP and the NICD are integrated in a logic-gate-like structure around the CDX2 enhancer;
2. The components are part of a network;
3. The network regulates the transcription of CDX2, which is a key event in TE-ICM differentiation;
4. The AND gate model can explain the NICD and YAP cooperation, plus the RBPJ null phenotype.

The AND gate model can be tested, and it predicts that CDX2 induction is lost after the abrogation of either the NICD or the YAP. The alternative, which is an OR gate, would require the suppression of both factors.

## 5. The Outlook

By extrapolating from the TE enhancer of CDX2, it is concluded that logic-gate-like mechanisms (i.e., the combining and filtering of signals) are biologically relevant. Starting from a uniform cell population at the eight-cell stage, the regulative network is thought to develop to high complexity, together with the differentiation of multiple cell types, and the concomitant diversification of intracellular signal flow systems.

There are important differences between the AND gate that is described here and the gates that are in computers. In computers, the signals are distinct, and the gates serve as tools that are used to perform Boolean procedures without an influence on the signal quality. In cells, the sensing has to detect the concentration differences. The NOTCH concentration differences between the outer and inner cells are gradual, and it is apparent that the term, “HIPPO inactive”, which senses the polarization, refers to an important, but not absolute, sequestering of the AMOT protein by the actin filaments. Since two signals have to coincide for a positive gate output, the AND gate helps to switch the CDX2 expression ON or OFF as a clear-cut output signal. The ensuing unequivocal differentiation between the TE and ICM cells, and the formation of distinct cell layers, involves further regulatory steps. Thus, the CDX2 protein stimulates the expression of its own gene and suppresses ICM-specific genes [9]. Other genes interfere, and the interactions between the cells add to the regulation, so that, presumably, many more gates are associated.

Moreover, the processing modes in computers and cells are different. Conventional computers are Turing machines, where the processing steps occur in a linear way, one by one. Their processors contain transistors and logic gates that are set in place by design. The cells are different in both aspects. They receive many signals and activate many signaling chains at the same time. Many processing events occur simultaneously. The combined signal chain components are not locally fixed; however, they are not freely diffusible either. We do not know much about the topography and timing of the signaling, but we do know that the signal proteins are clustered into signalosomes around the cell membrane receptors, around the gene regulation site, and in between the assemblies that are held by scaffolding proteins [4]. The signal transfer inside signalosomes basically occurs in a linear way, and the interactions between the signalosomes take place only between certain components. The signal transfer between the chains at preferential hubs might make the whole network similar to communicating Turing machines. In addition, the cells undergo a continuous turnover of proteins and other materials, and the tagging of the signaling factors for decay is, by itself, an element of signal propagation [3]. Thus, the processing network is continuously replacing its signal carriers. There is a number of receptors of each sort, and all of the individual molecules feed their signals into the processing network. The network addresses the regulatory elements of the genes, and both induces and represses them. To put it simply, there is no motherboard with a defined place for every transistor, but, rather, there is a zoo of signal chain components, with many of them in the form of aggregates, which are engaged in fluctuating interactions. However, it is evident that the system is highly ordered because the development of the organisms proceeds reproducibly, and the response to the outer signals is predictable. Its robustness towards faulty components is limited because many mutations or misexpressions of the single signal-chain elements lead to cell death or cancer formation. The way the network works is, so far, elusive; however, it is of interest not only for the understanding of the information flows in biological systems, but also for the analysis of information processing mechanisms in general.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The author declares no conflict of interest.

## References

1. Joyce, G.F.; Orgel, L.E. Prospects for Understanding the Origins of the RNA World. In *The RNA World*; Gesteland, R.F., Atkins, J.F., Eds.; Cold Spring Harbor Laboratory Press: New York, NY, USA, 1993; pp. 1–25.
2. Wurtz, T. Nested information processing in the living world. *Ann. N. Y. Acad. Sci.* **2021**, *1500*, 5–16. [[CrossRef](#)] [[PubMed](#)]
3. Marks, F.; Klingmüller, U.; Müller-Decker, K. *Cellular Processing: An Introduction to the Molecular Mechanisms of Signal Transduction*, 2nd ed.; Taylor & Francis Group: New York, NY, USA; Abington, PA, USA, 2017.
4. Csizmok, V.; Follis, A.V.; Kriwacki, R.W.; Forman-Kay, J.D. Dynamic protein interaction networks and new structural paradigms in signaling. *Chem. Rev.* **2016**, *116*, 6424–6462. [[CrossRef](#)] [[PubMed](#)]
5. Menchero, S.; Rayon, T.; Andreu, M.J.; Manzanares, M. Signaling pathways in mammalian preimplantation development: Linking cellular phenotypes to lineage decisions. *Dev. Dyn.* **2017**, *246*, 245–261. [[CrossRef](#)] [[PubMed](#)]
6. Sasaki, H. Roles and regulation of HIPPO signaling during preimplantation mouse development. *Dev. Growth Differ.* **2017**, *59*, 12–20. [[CrossRef](#)] [[PubMed](#)]
7. Frum, T.; Ralston, A. Cell signaling and transcription factors regulating cell fate during formation of the mouse blastocyst. *Trends Genet.* **2015**, *31*, 402–410. [[CrossRef](#)] [[PubMed](#)]
8. Watanabe, Y.; Miyasaka, K.Y.; Kubo, A.; Kida, Y.S.; Nakagawa, O.; Hirate, Y.; Sasaki, H.; Ogura, T. Notch and Hippo signaling converge on Strawberry Notch 1 (Sbno1) to synergistically activate Cdx2 during specification of the trophectoderm. *Sci. Rep.* **2017**, *7*, 46135–46152. [[CrossRef](#)] [[PubMed](#)]
9. Niwa, H.; Toyooka, Y.; Shimosato, D.; Strumpf, D.; Takahashi, K.; Yagi, R.; Rossant, J. Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation. *Cell* **2005**, *123*, 917–929. [[CrossRef](#)] [[PubMed](#)]
10. Souilhols, C.; Cormier, S.; Tanigaki, K.; Babinet, C.; Cohen-Tannoudji, M. RBP-J kappa-dependent notch signaling is dispensable for mouse early embryonic development. *Mol. Cell. Biol.* **2006**, *26*, 4769–4774. [[CrossRef](#)] [[PubMed](#)]
11. Rayon, T.; Menchero, S.; Nieto, A.; Xenopoulos, P.; Crespo, M.; Cockburn, K.; Cañon, S.; Sasaki, H.; Hadjantonakis, A.K.; de la Pompa, J.L.; et al. Notch and Hippo converge on Cdx2 to specify the trophectoderm lineage in the mouse blastocyst. *Dev. Cell* **2014**, *30*, 410–422. [[CrossRef](#)] [[PubMed](#)]