Modeling Cadmium Induced Metabolic Shifts in Human Pulmonary Cells

Valentin Vigeant[1][a], Jean-Paul Comet[1][b], Gilles Bernot[1][c] and Jean-Yves Trosset[2][d]
[1]*Laboratoire d'Informatique, Signaux et Systèmes de Sophia Antipolis, Université Côte d'Azur, Euclide B, 2000 Rte des Lucioles, 06900 Sophia Antipolis, France*
[2]*Bioinformation Research Laboratory, Sup'Biotech, Villejuif, France*

Keywords: Discrete Modeling, Cadmium, Lung Cells, Cancerization, Formal Methods.

Abstract: Non-genotoxic carcinogens (NGTxCs) induce cancer without directly altering the genetic material, making their mechanisms of action particularly challenging to predict. Cadmium, a well-known NGTxC classified as carcinogenic to humans, has been associated with several cancers, including lung cancer through chronic exposure. In this work, we investigate the carcinogenic effects of cadmium through metabolic reprogramming. We adapt a previously formally validated model of the metabolism regulation (Gibart et al., 2021c) in order to study cadmium exposure. It integrates major cadmium-induced perturbations such as oxidative stress, mitochondrial dysfunction, and glycolytic shift. Using the formal verification tool TotemBioNet, we systematically explore and filter all dynamics consistent with the global behaviors of healthy, cancerous, and apoptotic cells. They are investigated and classified into three subgroups. Our approach successfully reproduces characteristic features of carcinogenic metabolism, including fermentation activation and elevated ROS production, and demonstrates the relevance of discrete qualitative modeling for a systemic analysis of NGTxC-induced metabolic dysregulations.

1 INTRODUCTION

Among human carcinogens, the most well-known are mutagenic agents, whose effects can be directly observed at the genetic code level. However, some carcinogens are classified as non-genotoxic (NGTxC) and are capable of inducing cancer without directly causing genetic mutations. Unlike most drugs, cadmium has no specific molecular or receptor target; its mechanism of action is inherently systemic, impacting multiple cellular pathways and processes. Moreover, the diverse modes of action of non-genotoxic carcinogens, together with their tissue and species-specific effects, make predicting their carcinogenic potential highly challenging (Hernández et al., 2009). Understanding the causal chains between exposure to a non-genotoxic carcinogen and cancer onset is thus an important challenge.

Cadmium is a NGTxC (Waalkes, 2000), classified as carcinogenic to humans in 1993, chronic exposure through smoking is associated with lung cancer in humans. Additionally, this heavy metal can have significant health consequences if ingested, leading to prostate, pancreas or even breast cancer (Satarug et al., 2010).

It has been observed that cancer cells exhibit distinct metabolic adaptations compared to healthy cells (Qu and Zheng, 2024; Nagao et al., 2019; Gao et al., 2014). In this study, we propose a model to investigate how cadmium systemically induces cancer-associated metabolic reprogramming. These metabolic shifts are multiple (Qu and Zheng, 2024; Chen et al., 2015). To unravel how cadmium exposure reshapes cellular metabolism, we developed a computational model capturing the key regulatory mechanisms of metabolic control. As cadmium exposure is closely associated with tobacco smoking and lung carcinogenesis, this study specifically investigates its effects on lung cells. We develop a computational model that captures cadmium-induced metabolic reprogramming, providing mechanistic insights and identifying potential markers of its carcinogenic action.

[a] https://orcid.org/0009-0000-6880-8848
[b] https://orcid.org/0000-0002-6681-3501
[c] https://orcid.org/0000-0002-7149-0559
[d] https://orcid.org/0000-0001-9805-6237

Vigeant, V., Comet, J.-P., Bernot, G. and Trosset, J.-Y.
Modeling Cadmium Induced Metabolic Shifts in Human Pulmonary Cells.
DOI: 10.5220/0014424900004070

In *Proceedings of the 19th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2026) - Volume 2: BIODEVICES, BIOIMAGING, BIOINFORMATICS*, pages 539-550
ISBN: 978-989-758-802-0; ISSN: 2184-4305

Numerous approaches have been developed for modeling regulatory processes. Stochastic approaches highlight the random aspect of regulation (Chu et al., 2017). Differential systems assume a continuous shape of regulations (Wu et al., 2014). Discrete systems highlight the qualitative aspect of regulation (Mateus et al., 2007). Hybrid models that combine some of these frameworks try to offer the best of the different modeling frameworks (Behaegel et al., 2016). Because the observed metabolic shift corresponds to a qualitative change, we prefer a discrete modeling framework. The discrete framework of René Thomas is particularly well-suited from a qualitative perspective (Thomas, 1991). Here, we use an extension of Thomas' modeling framework through multiplexes, logical entities that allow the modeling of combinatorial regulation (Khalis et al., 2009).

Here we extend the metabolic regulation model initially proposed in (Khoodeeram, 2021), and subsequently consolidated in (Gibart et al., 2021c). This model focuses on the regulatory signals of the main catabolic and anabolic pathways of eukaryotic cells, together with cofactors and the most important nutrients.

To represent cadmium-induced perturbations of metabolism, we incorporated the major regulatory pathways affected, such as oxidative stress (Wang et al., 2022), variations in glycolysis via HIF-1α activity (Hart et al., 2001), and disruption of the mitochondrial respiratory chain (Jing et al., 2012). The regulatory rules that govern their functioning have to be precisely detailed in the model. Contemporary biochemical and biological literature gives generally sufficient information in order to infer the rules (cooperation/concurrences) and parameters (regulation strength) of the model.

The model is first constructed based on 'local' biological knowledge of regulatory interactions among metabolic components. Its global dynamics are then explored using formal verification methods, enabling a systematic, computer-assisted analysis of all possible state transitions (Bernot et al., 2004; Bernot et al., 2019).

TotemBioNet (Boyenval et al., 2020; Gibart et al., 2021b) which includes GreenBioNet (Gibart et al., 2021a), provides all the necessary tools to uncover the underlying kinetic parameters of a model, and verify its consistency with biological observations (expressed in formal languages). Model checking of temporal properties applied to experimental measurements is intensively used. In our context, we design a model that aligns with the global temporal properties observed in healthy, apoptotic and cancerous cells.

This paper is organized as follow. In Section 2, we briefly outline the theoretical foundations, the associated methodology, and the tools enabling efficient algorithmic enumeration and verification. Relevant information about cadmium and its effects on metabolic regulation, as well as their graphical representation, are detailed in Section 3. Section 4 describes how we adapted the model of metabolism regulation for our case study. This adaptation led to a simplified, more effective model. The validation of our model is presented in Section 5, revealing all parameterized models compatible with the properties identified for either of 3 different cell fate classes: healthy, cancerous, apoptotic. Lastly, Section 6 concludes and provides some perspectives.

2 METHODOLOGIES & TOOLS

2.1 Interaction Graph

In first place, a model of biological regulatory network rely on a directed interaction graph between variables whose nodes represent the main elements of the regulation network and arcs the regulations. As shown in Figure 1, metabolites, metabolic pathways, and biomass are generally represented by variables. In discrete modeling, quantitative measures such as activity levels, rates, or concentrations are abstracted by qualitative levels (positive integers). The levels of a particular variable are defined by carefully chosen thresholds, associated with the regulations of the considered variable on its targets. Each qualitative level is associated with a distinct set of regulated targets. For instance, level 1 of variable ***LBP*** corresponds to a scenario in which ***LBP*** regulates ***ATP*** and **Box**. Assigning a level to each variable defines a possible state of the system.

Figure 1 also shows the regulations as rectangles called multiplexes. They formalize the rules of cooperation/competition between several regulators. For instance, the **AnO** multiplex regulates the Krebs cycle when glycolysis is greater than or equal to 1 and ***O2*** is present. For context, **AnO** represents the action of acetyl-CoA as main precursor of the Krebs cycle. More generally, each multiplex contains a formula that specifies the conditions under which regulation occurs. Small rectangles containing symbols ! represents inhibition at level 1, and with +1 (resp. -1) represents activation (resp. inhibition) at level 1.

2.2 Dynamic Parameters

In R. Thomas' modeling framework, the evolution of variables is guided by a target level determined solely

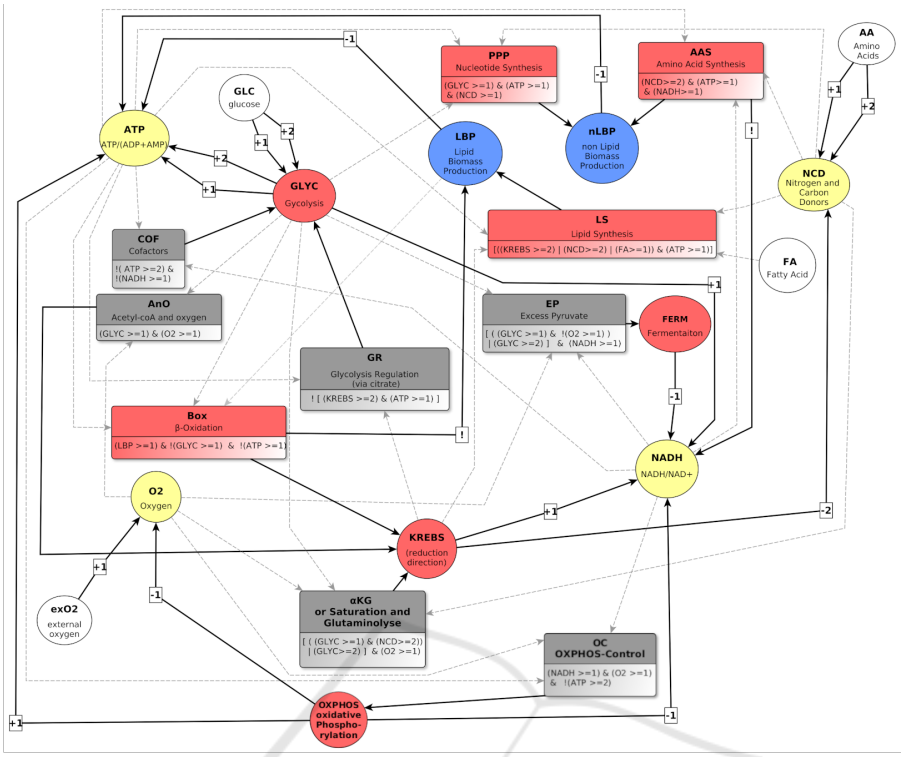

Figure 1: Initial interaction graph of cellular metabolism from (Gibart et al., 2021c). Circles denote variables and rectangles represent multiplexes; blue, yellow, and red nodes correspond respectively to biomass, metabolites, and pathways. Solid arcs indicate targets, dashed arcs sources (deducible from formulas). Logical symbols: ! = negation, & = conjunction, | = disjunction.

by the active regulations they receive. A dynamic parameter refers to the value a variable tends toward based on a specific set of active regulations (multiplexes), where a multiplex is considered active if and only if its logical formula evaluates to true in the current state. Each set ω of active multiplexes corresponds to a dynamic parameter denoted $K_{v,\omega}$. This parameter represents a local behavior, i.e., the value that would hold if ω remained stable indefinitely. However, since the set of active regulations changes over time, the corresponding attraction values also evolve. Nevertheless, these attraction values give the direction towards which the system is attracted (Snoussi, 1989). When two variables tends to evolve, we do not know which one will cross its threshold first, leading to several possible local evolutions. R. Thomas' modeling framework proposes to represent the global behavior of the system as a so-called states graph where nodes correspond to states, and arcs represent the possible local evolution of the system (in the direction of the attraction values). Since any potential regulation or multiplex may be either active or inactive, the number of required parameters grows rapidly: if a variable has n predecessors x_1 to x_n in the graph, then it has 2^n possible combinations of its active regulations (i.e., all subsets ω of $\{x_1,\ldots,x_n\}$). As a result, 2^n parameters are required to describe its dynamics. This exponential growth highlights the necessity of simplifying the model as much as possible.

2.3 Toward Automated Validation

A biological observation generally leads to dynamic knowledge which can be taken into consideration for restricting the values of kinetic parameters. For example, the statement that a high dose of cadmium leads to the activation of fermentation is crucial information for the modeling process. Such properties can be formally expressed in *Computation Tree Logic* (CTL) (Clarke and Emerson, 1982), a temporal logic designed to describe the possible trajectories of the studied system. Its tree-like representation of time naturally accommodates the indeterminism of possible behaviors: a given state of the regulatory network may lead to different traces with distinct sequences of observations.

- Time operators of CTL are: X "neXt", G "Globally", F "in the Future", U "Until".
- They are accompanying path quantifiers: A "for all path", and E "there exists a path".

The previously mentioned behavior can be formalized as: $[(Cd > 1) \Rightarrow AF\,(FERM = 1)]$. This formula means that, starting from a state where Cd is at a level strictly greater to 1, for All possible trajectories in the Future, the system will reach a state where the fermentation is active.

CTL formulas express a broad set of dynamic system properties, including temporal reachability (*td*), which specifies whether a variable's value converges toward or fails to converge to (*!td*) a given target. It also captures *oscillatory* behaviors (*osc*) when the variable keep passing one or more thresholds. This makes CTL a powerful tool to rigorously specify, verify, and interpret complex behaviors in regulatory network models. Many known asymptotic behaviors can be captured using CTL formulas, and a parameterized model need only be considered if its behavior aligns with this knowledge. The process verifying whether a model satisfies a particular temporal formula is known as model checking.

2.4 Model Checking

The complexity of intertwined regulations and the large number of possible system states make any manual verification of temporal properties infeasible, thus requiring automated procedures. The main software used in this work is *TotemBioNet* (Boyenval et al., 2020). It enables the systematic exploration of all admissible parameterizations of a Thomas' regulatory model. Its inputs consist of a formal description of the interaction graph between variables, the admissible ranges for kinetic parameters, and a set of biologically known global behaviors expressed in the form of temporal logic formulae (CTL). For each possible parameterization, the software automatically constructs the corresponding state graph and applies CTL model checking. As a result, TotemBioNet produces the exhaustive list of parameterized models that are consistent with the known biological phenotypes.

This approach provides a rigorous and reproducible way to link interaction graph to observed cellular behaviors expressed in the form of CTL formulae, thereby facilitating both model validation and biological interpretation.

3 EFFECTS OF CADMIUM

3.1 Cadmium, a NGTxC Mimicking Cell Hypoxia

Cadmium raises significant biological concerns due to its multifaceted toxicity. First, it is classified as a human carcinogen (non-genotoxic, often referred to as a class 1 carcinogen). Second, its effects differ substantially depending on the duration of exposure: while acute exposure to a given dose may produce limited effects, chronic exposure (defined as repeated exposure to the same dose) leads to severe pathologies, including tissue damage and cancer development. In addition, similarities can be observed between the metabolic changes induced by cadmium exposure and those associated with the Warburg effect, a metabolic shift in which cells prioritize glycolysis and fermentation over respiration, even in the presence of oxygen (Larson-Casey et al., 2020; Nagao et al., 2019). Notably, the disruptive actions of cadmium on mitochondrial metabolism mimics the changes happening when a cell lacks oxygen (Jing et al., 2012), leading to the activation of fermentation to produce enough ATP, similarly to the Warburg effect.

3.2 Initial Model

Cadmium has been identified as a highly potent regulator of metabolism, exerting its effects on multiple critical components of the system. The identification, comprehension, and accurate integration of these actions into the model in the most abstract form possible constitutes a pivotal element of the modeling process. The similarity with the Warburg effect led us to design a model illustrating the impact of cadmium on metabolism. It is based on the model developed in (Gibart et al., 2021c) that was initially validated for representing the Warburg effect. This model offers a general framework for central carbon metabolism regulation and can be applied to study systemic metabolic responses under various biological perturbations. The variables and multiplexes correspond to well-known, easily observable, and interpretable elements of cell carbon metabolism as seen in Figure 1. External supplies for cellular metabolism include $exO2$ (external oxygen), FA (fatty acids), GLC (glucose), and AA (amino acids). LBP and $nLBP$ represent lipidic and non-lipidic biomass production. ATP and $NADH$ denote the ratios ATP/(ADP+AMP) and NADH/NAD$^+$. NCD refers to nitrogen and carbon donors. $OXPHOS$, $O2$, $KREBS$, and $GLYC$ correspond to oxidative phosphorylation, oxygen, the Krebs cycle, and glycolysis. Multiplexes include: **EP** (excess pyruvate), **LS** (lipid synthesis), **OC** (OXPHOS control), **GR** (glycolysis regulation via citrate from the Krebs cycle), **BOX** (β-oxidation), **AnO** (Acetyl-CoA formation in presence of oxygen), **COF** (glycolytic cofactors), **PPP** (nucleotide synthesis through the Pentose Phosphate Pathway), **AAS** (amino acid synthesis), and **αKG** (accumulation of α-ketoglutarate from

glutaminolysis or glycolytic saturation).

The Warburg effect being well documented, it has been possible to determine the majority of the parameters of the initial model. Out of more than a hundred parameters, only seven remain unconstrained. Since most parameters have already been identified, incorporating the effects of cadmium will be easier. The initial model's limited behavioral flexibility actually facilitates this process: the range of behaviors it encompasses directly defines the constraints for the enriched model. Ultimately, a more constrained enriched model offers a deeper understanding of the underlying mechanisms.

3.3 Cadmium Mechanistic Effects

Our approach brings five important modifications to the initial model. All these successive enhancements to the model result in the interaction graph shown in Figure 2.

$\circ$ $Cd \in 0,1,2$: A cell exposed to cadmium in its environment (external cadmium is represented by the variable *exCd*, see subsection 4.1) leads to the passage and ultimately the presence of cadmium in the cell (expressed by *Cd*). While $Cd = 0$ corresponds to a low, non-effective concentration, $Cd = 1$ indicates a level sufficient to trigger a cellular response. $Cd = 2$ represents an overload that the cell cannot efficiently handle. Cadmium accumulation disrupts mitochondrial metabolism, primarily inhibiting complexes III and V of the respiratory chain (Wang et al., 2022; Branca et al., 2020b). This leads to electron leakage and increased production of Reactive Oxygen Species (*ROS*, new variable), captured in the model via effects on *NADH* through the deregulation of alpha-ketoglutarate (αKG) (Cho et al., 2018). The inhibition of *NADH* regeneration permits elevated *ROS* levels through the **noRos** multiplex. Cadmium also directly impairs oxidative phosphorylation (Genchi et al., 2020), represented by a direct arc from *Cd* to the *OXPHOS* variable. In the interaction graph, the impacted elements can be found in the gray frame.

$\circ$ $GSH \in 0,1$: The implications of cadmium chelators in regulating ROS levels and the availability of free cadmium, requires their consideration in models. They are abstracted in the form of the variable *GSH* that stands for the principal chelator of cadmium: level 0 stands for a basal or depleted quantity of GSH, while level 1 represents an elevated level of GSH. Under low/moderate cadmium exposure ($Cd = 1$), cells activate an adaptive mechanism that increases glutathione production (leading to an increase of *GSH*). This is reflected by the arc labelled +1 from *Cd* to *GSH*. Consequently, the chelation of cadmium is enhanced, helping to limit ROS accumulation. Conversely, at elevated cadmium concentrations, a decline is evident, attributable to the saturation induced by cadmium and ROS (Branca et al., 2020a). Both arcs labelled -2 coming from theses variables transcribes it. If glutathione becomes insufficient, free cadmium can accumulate, leading to mitochondrial dysfunction, enhanced ROS generation and other forms of cellular damage. In the interaction graph, the impacted elements can be found in the blue frame.

$\circ$ $ABC \in 0,1$: Cadmium efflux is primarily mediated by ATP-binding cassette (ABC) transporters, represented by variable *ABC*. Level 0 corresponds to inactive or saturated transporters, while level 1 indicates efficient cadmium excretion. This excretory activity is modeled by an inhibitory arc labelled -1 from *ABC* to *Cd*. ABC-mediated transport requires ATP (Thévenod and Wing-Kee, 2024). This is captured in the interaction graph as a negative loop between *ABC* and *ATP*. Since only chelated cadmium can be exported, the model includes a multiplex to represent the condition that both *Cd* and *GSH* must be present to activate *ABC*. Formally, this is expressed as: $\square$ **Gluthathion-Cd/chelation** $(Cd \geq 1) \wedge (GSH \geq 1)$. In the interaction graph, the impacted elements can be found in the green frame.

$\circ$ $ROS \in 0,1,2$: As mentioned, another effect of cadmium in the cell is the rise in ROS. Level 0 of the *ROS* variable indicates a healthy level of ROS in the cell. Level 1 corresponds to a high level of ROS that can be managed by the cell, and level 2, corresponds to a concentration that, if sustained long enough, triggers apoptosis through a self-defense mechanism. While mitochondrial respiratory chain disruption contributes to ROS accumulation, it does not fully explain the observed increase. Additional factors play major roles. The inhibition of enzymes such as Superoxide dismutase (SOD), catalase, and glutathione peroxidase, which are essential for managing ROS has been shown to cause such an increase (Hernández-Cruz et al., 2022). In the interaction graph, the abolition of the system responsible for the management of ROS is represented by the **noROS** multiplex $\square$ **noROS** $\neg\big((AA \geq 2) \wedge (NADH \geq 1) \wedge ((ATP \geq 1) \vee (ROS \geq 1))\big) \wedge \neg(GSH \geq 1)$. The depletion of available enzymes responsible for managing ROS is exacerbated by the competition between ROS and cadmium for certain enzymes, such as glutathione, which functions as a chelator for cadmium (Gaubin et al., 2000). There is then a reduction of the detoxification of the ROS as well as an increase in production due to electron leakage. The production of ROS through oxidative phosphorylation in addition to cadmium could lead to a threshold cross-

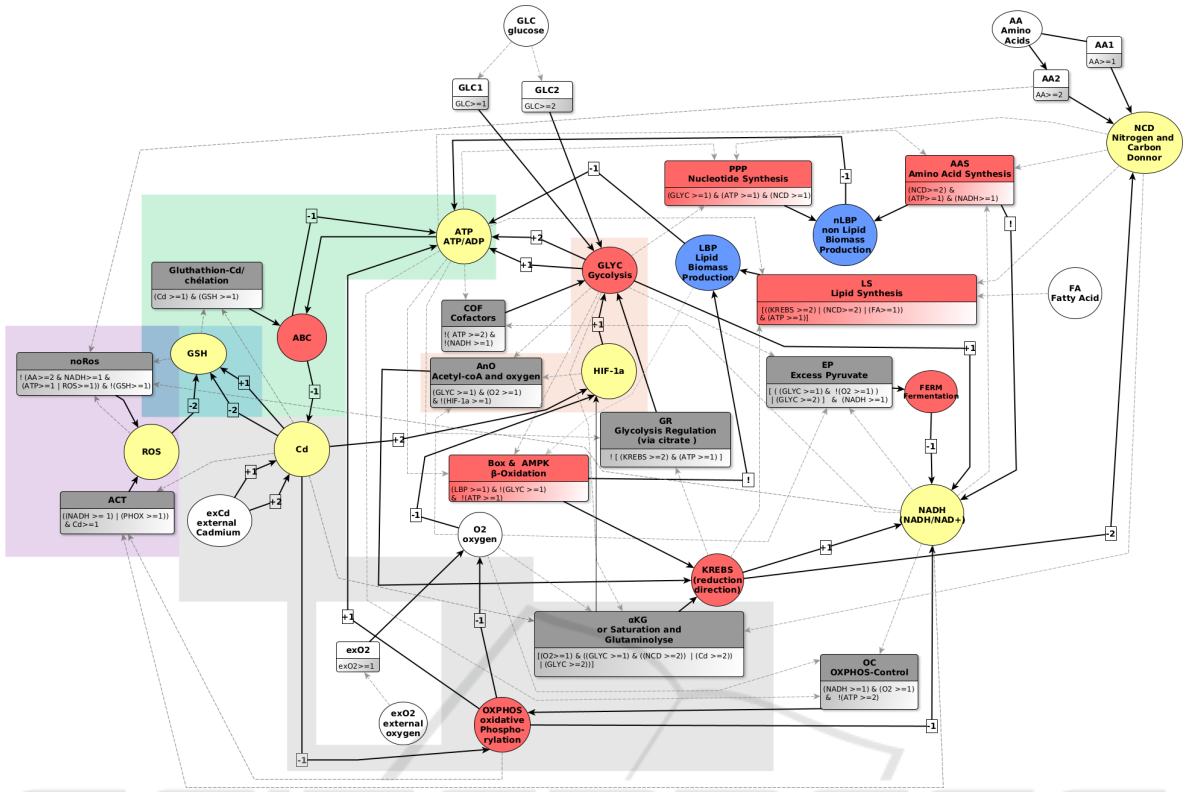

Figure 2: Interaction graph of cellular metabolism taking into account the main cadmium effects. .

ing. This action is represented in the **ACT** (activators) multiplex, who gathers activators of ***ROS*** $\square$ **ACT** $((\textit{NADH} \geq 1) \vee (\textit{OXPHOS} \geq 1)) \wedge (\textit{Cd} \geq 1)$. The formula translates that **ACT** remains active whenever ***Cd*** is present and at least one of ***NADH*** or ***OXPHOS*** is active. In the interaction graph, the variable ***ROS*** and the impacted elements can be found in the purple frame.

$\circ$ $\textit{HIF-1a} \in 0, 1$: Lastly, the disruption of the respiratory chain can induce a pseudo-hypoxia state within the cell (Jing et al., 2012; Kim et al., 2006). Despite the presence of enough oxygen, an increase in Hypoxia-Inducible Factor-1α (HIF-1α) levels occurs, leading to a substantial migration of glucose transporters to the membrane (Jing et al., 2012). This increased level is represented by the level 1 of the variable ***HIF-1a***, while the level 0 stands for a basal level. Under cadmium exposure, ***HIF-1a*** receives conflicting signals: cadmium promotes its stabilization (+2 arc from ***Cd***) by inhibiting prolyl hydroxylases (PHD) which uses O2 as a cofactor, through mitochondrial disruption and oxidative stress. In this case, an increase in certain intermediate metabolites (e.g., succinate, fumarate) can inhibit PHD, preventing the degradation of HIF-1α, even when oxygen is present. Meanwhile, the presence of oxygen typically triggers its degradation (-1 arc from ***O2***). These opposing influences create a regulatory "dilemma" that ultimately determines the level of HIF-1α in the cell. Glucose uptake by cells can thus be augmented without an escalation in the extracellular glucose concentration. The +1 arc from ***HIF-1a*** to ***GLYC*** represents this effect. This in turn enables the acceleration of glycolysis, which, in certain instances, results in the induction of fermentation within the cell. This phenomenon is analogous to the Warburg effect (Kim et al., 2006). Another consequence of HIF-1α stabilization is its inhibition of pyruvate conversion to acetyl-CoA via upregulation of PDK1[1], which phosphorylates and inactivates pyruvate dehydrogenase (PDH), the enzyme responsible for pyruvate's oxidative decarboxylation (Kim et al., 2006; Semenza, 2012). As a result, ***HIF-1a*** is included in the multiplex acetyl-CoA and oxygen (**AnO**). The inactivation of PDH implies a lack of efficiency of the regulatory mechanisms on ***HIF-1a*** such as those mediated by α**KG** or Saturation and Glutaminolyse multiplexe. The impacted elements are in the orange frame.

[1]NCBI Gene ID 5163: PDK1 pyruvate dehydrogenase kinase 1 [Homo sapiens], https://www.ncbi.nlm.nih.gov/gene/5163, Accessed: 2025-10-07.

4 MODEL ADAPTATION

4.1 Cell Environment Specification

The environmental variables representing what is available in the environment of the cell are: *exO2* (external oxygen), *FA* (fatty acids), *GLC* (external glucose), *AA* (amino acids), which represent essential supplies, and finally *exCd* (external Cadmium). To prevent confounding effects from restrictive conditions such as hypoxia or nutrient deprivation, the four environmental variables (excluding *exCd*) are held at their level 1 (Gibart et al., 2021c), to ensure normal cellular function. Under these conditions, only three distinct environments remain, defined solely by the level of cadmium exposure. The first environment, where $exCd = 0$, stands for a lack or an insignificant amount of cadmium exposure. When $exCd = 1$, the level of cadmium exposure is enough to induce a response and adaptations in the cell metabolism, without killing the cell. At $exCd = 2$, the level of cadmium exposure is unmanageable for the cell, and can cause cell death.

4.2 Consequences on Variables

After defining the cellular environments under consideration, some internal variables can be simplified without loss of essential information, based on different principles. First, if the value of a variable is solely determined by a predecessor that has been fixed, the variable itself can also be fixed, by assuming it has already reached its target value at the time of study (as the environment is constant, these internal variables will eventually reach this value). This is the case, for instance, with the variable *O2*. In earlier work on the metabolic graph (Khoodeeram, 2021), physiological fluctuations of internal oxygen were represented through a three-valued variable. Under physiological conditions, *O2* levels would oscillate between values 1 and 2, without ever reaching hypoxia (value 0). Although such fluctuations do occur, their timescale ($\sim$40 minutes) is irrelevant to our study, which focuses on the effects of chronic cadmium exposure. Therefore, in the simplified graph, the *O2* variable is reduced to two levels. Here, the value 0 still corresponds to intracellular hypoxia, whereas the value 1 now encompasses the former levels 1 and 2, both representing oxygen levels sufficient for normal cellular utilization. Since *exO2* is fixed, the parameters indicate that *O2* will always converge toward 1. Concerning *HIF-1a*, as said in subsection 3.3, cadmium stabilises HIF-1α through the inhibition of PDH, implying that regulation mechanisms of HIF-1α are not effective. This is why, in the simplified version, there is no need for them to appear as predecessors of *HIF-1a*.

In addition, the formulas for some multiplexes can be simplified. Fixed external variables, and local variables under exclusive control by a fixed variable do not appear anymore, and certain conditions become always true or always false. First, the simplifications of *O2* and *NCD* induces changes in the following multiplexes:

- **AnO** (Acetyl-coA and O2) controlling *KREBS* becomes **AcoA** : $[(GLYC \geq 1) \& \neg(HIF\text{-}1a \geq 1)]$
- **PPP** (nucleotide synthesis) controlling *nLBP*: $[(GLYC \geq 1) \& (ATP \geq 1)]$
- **OC** (Oxphos Control) controlling *OXPHOS*: $[(NADH \geq 1) \& \neg(ATP \geq 2)]$
- **αKG** (or Saturation and Glutaminolyse) on *KREBS*: $[((GLYC \geq 1) \& \neg(Cd \geq 2)) \vee (GLYC \geq 2)]$
- **EP** (Excess Pyruvate) controlling *FERM*: $[(GLYC \geq 2) \& (NADH \geq 1)]$
- **LS** (lipid synthesis) on *LBP*: $[(ATP \geq 1)]$

Here, since the variable *FA* is fixed at 1, the first condition is always validated, allowing to simplify the formula. Only one condition is then left on *ATP*, the multiplex can be substituted by an arc labelled +1 from *ATP* to *LBP*. Furthermore, the multiplex **AAS** (Amino Acid Synthesis) controlling *nLBP*: $[(NCD \geq 2) \& (ATP \geq 1) \& (NADH \geq 1)]$ is never verified since *NCD* is fixed at 1. The formula cannot be validated, so the multiplex can be suppressed. This implies that the inhibition on *NADH* is not reached. Consequently, the parameters $K_{NADH,\texttt{FERM GLYC1 OXPHOS}}$ and $K_{NADH,\texttt{FERM GLYC1 KREBS}}$ which were free in the initial model, disappear.

4.3 Abstraction of Cd's Chelators

The variable *GSH* represents the levels of free glutathione in the cell. Free glutathione is required both for cadmium chelation and for the regulation of ROS (Nzengue et al., 2008). From a modeling point of view, the role and actions of the variable *GSH* can be covered by other elements of the graph. As said in subsection 3.3, there are two levels of *GSH*, and this is sufficient since its effects are observed when *GSH* is at level 1. Indeed, when $GSH = 0$, effective cadmium chelation does not occur, either by absence of internal cadmium ($Cd = 0$), or because *GSH* is saturated ($Cd = 2$ implies saturation of *GSH*). Similarly, one of the conditions for no excessive *ROS* accumulation through the **noRos** multiplex is that the *GSH* level remains high. However, at higher cadmium doses, where ROS production becomes substantial, GSH availability becomes limited (Gaubin et al., 2000). This depletion of free GSH leads to oxidative stress and cellular damage.

In Figure 2, the role of *GSH* in *ROS* regulation is encoded within the formula of the multiplex **noROS**, whose activation requires that *GSH* remains strictly below 1. Indeed, achieving level 1 of *GSH* necessitates cadmium as an activator of *GSH* (i.e. $Cd = 1$). Consequently, as illustrated in Figure 3, by bypassing *GSH* and replacing its constraint within the **noROS** multiplex with the requirement that $Cd \geq 2$, we still represent the effect of elevated free cadmium, and the regulatory dynamics on *ROS* are preserved (note that when $Cd = 0$, *ROS* is not considered activated). We can then deduce that $K_{GSH,\texttt{Cd1 Cd2 ROS}} = 1$ and $K_{GSH,\texttt{Cd1 Cd2}} \in 0,1$: these parameters of *GSH* are the only ones for which the activation (denoted **Cd1**) and the absence of inhibition (denoted **Cd2**) are resources, and consequently they are the only ones allowed to take the value 1.

Furthermore, *GSH* is required for the activation of the multiplex Glutathione–Cd/chelation. Indeed, cadmium must first be chelated for ABC transporters to excrete it. The representation of this mechanism can be simplified by modeling a direct action of *Cd* on *ABC*. In practice, an inhibition at level 2 for *Cd* is sufficient, since it parametrically encodes the constraint that the Glutathione–Cd/chelation multiplex cannot be activated when *GSH* is below level 1, i.e., insufficient to counterbalance the effects of *ROS* and *Cd*. In this case, *Cd* is not a resource to the variable *GSH* that would tend to the value 0. At the same time, the absence of inhibition when $Cd \leq 2$ plays the role of an active predecessor (just as $Cd = 1$ does for *ABC* via *GSH*).

After performing all simplifications, the resulting interaction graph can be seen in Figure 3.

4.4 Parameter Values

We are left with 109 parameters after simplifications [2]. Here, we discuss the values of all parameters; among them, all but 13 can be fixed.

The fate of the variable *Cd* is mostly dictated by the entry of cadmium from the environment. This is enough to fix to 0 the parameters whose resources do not contain external cadmium ($K_{Cd} = 0$ and $K_{Cd,\texttt{ABC}} = 0$). When the cell is exposed to cadmium, we need to take into account the excretory activity. While ABC transporters can function properly (i.e. when $ABC = 1$), the accumulation of cadmium cannot reach its peak, so $K_{Cd,\texttt{exCd1 exCd2}}$ is fixed to 1, while $K_{Cd,\texttt{exCd1 exCd2 ABC}}$ is fixed to 2, as ABC transporters saturation is reached. For $K_{Cd,\texttt{exCd1 ABC}}$, the external exposition is not high enough to provoke a sufficient entry to get to the highest threshold of *Cd*, so $K_{Cd,\texttt{exCd1 ABC}}$ is fixed to 1.

[2]Values of the parameters are available on request.

Some of the parameters of *OXPHOS* have their values changed because *Cd* is able to inhibit *OXPHOS*. When *Cd* is a resource of *OXPHOS* (*Cd*=0), we keep the parameter values established before the introduction of *Cd*: $K_{OXPHOS,\texttt{Cd}} = 0$ and $K_{OXPHOS,\texttt{OC Cd}} = 1$. When *Cd* is not a resource of *OXPHOS* ($Cd > 0$), $K_{OXPHOS,\texttt{OC}} = 0$ since *Cd* inhibits *OXPHOS*.

To take into account the necessary chelation on cadmium by the gluthatione, and the need of ATP for ABC transporters to function, the only parameter of *ABC* that makes possible the crossing of threshold 1 is the one where all predecessors are resources. Thus, all parameters are equal to 0 except $K_{ABC,\texttt{ROS ATP Cd}}$ which equals 1.

Regarding the variable *ROS*, we can only fix the parameters accounting for the situations where none of its predecessors are resources, implying no elevation in ROS levels ($K_{ROS} = 0$) and where all its predecessors are resources, corresponding to maximal ROS production ($K_{ROS,\texttt{ACT noROS}} = 2$).

For the variable *HIF-1a*, the sole presence of cadmium results in an elevation of HIF-1α, thus $K_{HIF,\texttt{Cd}} = 1$ and $K_{HIF} = 0$.

Furthermore, as *HIF-1a* is able to lead to a rise in glucose uptake from the external environment, the influence of *HIF-1a* on *GLYC* leads to a behavior of *GLYC* similar to the one with high level of nutrients (glucose) used by the cell. Indeed the parameters on *GLYC* with a set of resources containing *HIF-1a* get the same fixed values as those which have **GLC2** as a resource in (Gibart et al., 2021c): $K_{GLYC,\texttt{HIF}} = 0$, $K_{GLYC,\texttt{GR HIF}} = 0$, $K_{GLYC,\texttt{COF HIF}} = 1$, $K_{GLYC,\texttt{COF GR HIF}} = 2$.

Among the 7 free parameters of the initial model (see Table 1 of (Gibart et al., 2021c)), 2 have been fixed (both parameters concerning *NADH*, see end of section 4.2), and 5 remain free parameters of the current model (those controlling *ATP*). Moreover, some new parameters arise:

- $K_{ATP,\texttt{LBP OXPHOS ABC}}$: 1..2
- $K_{ATP,\texttt{LBP GLYC1 ABC}}$: 0..2
- $K_{ATP,\texttt{LBP GLYC1 GLYC2 ABC}}$: 1..2
- $K_{ATP,\texttt{LBP OXPHOS GLYC1 ABC}}$: 1..2
- $K_{ATP,\texttt{LBP OXPHOS GLYC1 GLYC2 ABC}}$: 1..2
- $K_{ROS,\texttt{noROS}}$: 0..2
- $K_{ROS,\texttt{ACT}}$: 0..2
- $K_{Cd,\texttt{exCd1}}$: 0..1

Among these 8 new parameters, the 5 controlling *ATP* sprout from the addition of the resource *ABC* over the pre-existing free parameters of *ATP*. ATP consumption by ABC transporters occurs in parallel with multiple other energy-demanding processes already captured in the reference model. In the absence of information supporting a dominant or threshold-crossing effect of ABC activity on ATP levels, the corresponding parameters were assigned values identical to those of their counterparts in the initial model

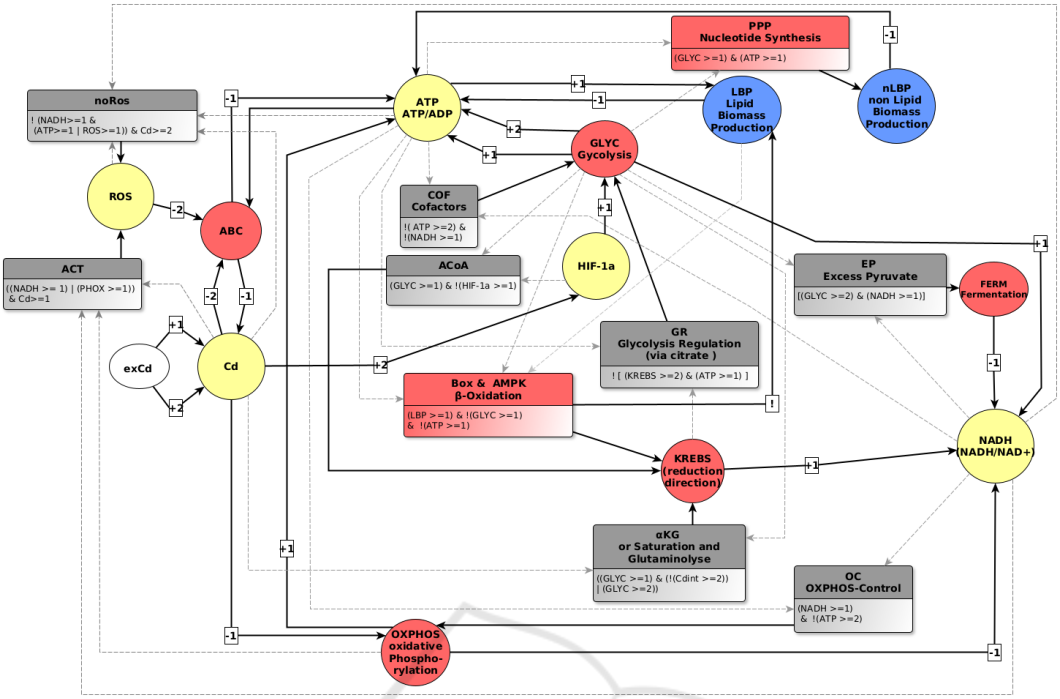

Figure 3: Simplified interaction graph of cellular metabolism taking into account the main cadmium effects.

(the initial model was designed for a context where ABC was supposed to a be resource). This conservative choice avoids introducing artificial ATP depletion while remaining consistent with the validated metabolic framework. The 5 parameters without ABC in resources remain free.

The remaining free ROS parameters correspond to intermediate regulatory configurations, in which either production or detoxification mechanisms dominate. The biological literature does not provide sufficient evidence to discriminate between moderate and strong ROS accumulation in these situations. Consequently, these parameters were deliberately left unconstrained in order to preserve the ability of the model to reproduce both adaptive and deleterious outcomes.

Finally, the free parameter on cd can take two values to account for the variability of the influx of cadmium in the cell for a given exposition.

The remaining parameters unchanged in the new model retain their values from the initial model (Gibart et al., 2021c).

5 MODEL VALIDATION

5.1 Global Biological Knowledge

Model validation is a key step to assess the biological relevance of the regulatory network. It verifies whether the model can reproduce known cellular phenotypes under normal and cadmium-exposed conditions, ensuring that the modeling framework captures the essential metabolic mechanisms of cadmium toxicity. From the perspective of studying the overall behavior of the system, one should describe the available knowledge on its global phenotypes in terms of its variables. For that matter we use Computation Tree Logic (CTL). Extensions such as Fair Path CTL, restrict verification to paths where no possible transition is ignored indefinitely.

The significant amount of knowledge available about the studied system is translated into CTL formulas that are then arranged in a validation matrix to facilitate their inventory. The rows of the matrix represent the different possible environments controlling the system (the rows of Table 1 list the possible levels of Cadmium present in the environment) whereas the columns list the relevant elements of the interaction graph. Each cell contains a CTL formula that indicates how the element associated with the column behaves over time in the environment represented by

Table 1: Validation matrix. Each row accounts for an environment, each column for a variable and each box for a formula. The boxes representing the formula associated with osc(1) (resp. osc(2)) means that the variable can oscillate anywhere between 0 and 1 (resp. 2). td(0) (resp. td(1), td(2)) means that the variable tends to 0 (resp. 1, 2). !td(0) (resp. !td(1), !td(2)) means that the variable does not tend towards 0 (resp. 1, 2). Theses formulas are interpreted following the semantic of Fair path CTL.

Context	Phenotype									
exCd	***ATP***	***GLYC***	***nLBP***	***LBP***	***KREBS***	***FERM***	***OXPHOS***	***NADH***	***HIF-1a***	***ROS***
0	osc(2)	osc(2)	!td(0)	!td(0)	osc(2)	td(0)	osc(1)	osc(1)	td(0)	!td(2)
1	osc(2)	osc(2)	!td(0)	!td(0)		td(0)			td(0)	!td(2)
2	osc(2)	osc(2)		!td(0)		!td(0)	!td(1)	!td(1)	!td(0)	!td(0)

the row. A cell may be left blank if no information is available. The environment where $exCd$=0 regroups the normal homeostatic behaviors of cellular metabolism. The two other environments account for different cadmium exposure settings.

5.2 Computer Aided Validation

Model checkers are essential tools for automatically ensuring consistency between the dynamics of the model and the expected biological behaviors. In this context, we used **TotemBioNet** (Boyenval et al., 2020) and its matrix validation module **GreenBioNet** (Gibart et al., 2021c). Special attention was paid to the free parameters left undefined during the initial modeling phase (see subsection 4.4). Their systematic exploration allowed us to identify which regulatory interactions are indispensable for the model to reproduce the observed biological phenotype, highlighting the mechanistic foundations of cadmium-induced metabolic reprogramming.

As 9 boolean and 4 trivaluated parameters are free, 41472 possible parameterizations remain ($2^9 \times 3^4$). Those 41472 models ought to be tested for their consistency with the validation matrix. Using fair-path CTL, 1800 parameterized models are compliant with the desired behaviors, including fermentation activation (see subsection 3.1). A clustering of these 1800 models was performed to separate models expressing different cellular fates. A key marker for these differentiated phenotypes is the level of ROS over time. The behavior of the ROS variable was therefore thoroughly investigated:

• Two hundreds parameterized models were found to tend to $ROS = 2$. As the level of ROS stays too elevated for survival, this is interpreted as cell death.

• On the contrary, eight hundreds parameterized models can be interpreted as cell survival through adaptation. Among them, six hundreds exhibit an oscillation on their level of ROS between the levels 0 and 1, never reaching the level 2, and another two hundreds tend to stay at level 1.

• Interestingly, four hundreds shows perpetual oscillations of the ROS level between 1 and 2, allowing for the mechanisms leading to a cancerous state to operate without the cell dying. This is interpreted as a potential oncogenic fate.

• Finally, four hundreds parameterized models oscillates between all possible levels of ROS, making it ambiguous as we have no indication of how much time is spent at each level, due to the nature of the modeling framework.

These results answer the question of cadmium-induced carcinogenesis in lung cells through metabolic regulation, with 400 models supporting this hypothesis.

6 CONCLUSION

In this work, we propose an abstract regulatory model based on current biological knowledge and a well-established reference model of metabolic regulation. Powerful formal methods and modeling tools such as model checking were used to determine the ties between different cell states and cell fates, enabled by the flexibility provided by the free parameters. The obtained models reproduce key dynamical behaviors of the system, including the activation of fermentation as well as non-genotoxic carcinogenesis induced by cadmium, which is the central result of our analysis. This work shows that abstract regulatory modeling enables reasoning at the system level, beyond individual regulatory interactions. By identifying classes of models corresponding to coherent metabolic behaviors, this approach provides an intermediate layer of interpretation that adds value before quantitative comparison with experimental data.

Several perspectives for improvement remain open. First, the reduction of indeterminate components in the validation matrix and the fixing of additional parameters would increase the accuracy of the model according to the available experimental observations. Second, extending the scope of the model to other cellular types or to additional NGTxC molecules and mechanisms would provide a broader validation and highlight the generality of our ap-

proach. Third, all 1,800 models validated can be classified into three categories, depending on the cellular outcome shown. This modeling approach can facilitate designing experiments to determine the class to which the *in vivo* model belongs. This would require a back-and-forth between experimentation and modeling, where experimental results would guide us in refining the already validated models.

ACKNOWLEDGEMENTS

This research partially supported by the NewgenTOXiv project, founded by the programme d'investissements d'avenir (PIA). We acknowledge Dr Georges Vassaux and his team at the IPMC in Valbonne for fruitful discussions and valuable input during the development of this study.

REFERENCES

Behaegel, J. et al. (2016). A hybrid model of cell cycle in mammals. *J*BCB, 14(01):1640001.

Bernot, G., Comet, J.-P., et al. (2019). A genetically modified Hoare logic. *T*CS, 765:145–157.

Bernot, G. et al. (2004). Application of formal methods to biological regulatory networks: extending Thomas' asynchronous logical approach with temporal logic. *J*TB, 229(3):339–347.

Boyenval, D. et al. (2020). What is a cell cycle checkpoint? the TotemBioNet answer. In *P*roceedings of the 18th Intl Conf. CMSB, volume 12314 of *L*NCS, pages 362–372.

Branca, J. J. V., Fiorillo, C., et al. (2020a). Cadmium-induced oxidative stress: Focus on the central nervous system. *A*ntioxidants, 9(6):492.

Branca, J. J. V., Pacini, A., et al. (2020b). Cadmium-induced cytotoxicity: Effects on mitochondrial electron transport chain. *F*rontiers in Cell and Developmental Biology, Volume 8 - 2020.

Chen, C.-Y., Zhang, S.-L., et al. (2015). Cadmium toxicity induces ER stress and apoptosis via impairing energy homoeostasis in cardiomyocytes. *B*iosci. Rep., 35(3).

Cho, H. J., Cho, H. Y., et al. (2018). Nadp+-dependent cytosolic isocitrate dehydrogenase provides nadph in the presence of cadmium due to the moderate chelating effect of glutathione. *J*BIC, 23(6):849–860.

Chu, B. K., Tse, M. J., et al. (2017). Markov state models of gene regulatory networks. *B*MC Systems Biology, 11(1):14.

Clarke, E. M. and Emerson, E. A. (1982). Design and synthesis of synchronization skeletons using branching time temporal logic. In Kozen, D., editor, *L*ogics of Programs, pages 52–71, Berlin, Heidelberg. Springer Berlin Heidelberg.

Gao, Y., Lu, Y., Huang, S., et al. (2014). Identifying early urinary metabolic changes with long-term environmental exposure to cadmium by mass-spectrometry-based metabolomics. *E*nvironmental Science & Technology, 48(11):6409–6418.

Gaubin, Y., Vaissade, F., et al. (2000). Implication of free radicals and glutathione in the mechanism of cadmium-induced expression of stress proteins in the a549 human lung cell-line. *B*iochimica et Biophysica Acta (BBA) - Molecular Cell Research, 1495(1):4–13.

Genchi, G., Sinicropi, M. S., et al. (2020). The effects of cadmium toxicity. *I*nternational Journal of Environmental Research and Public Health, 17(11).

Gibart, L., Collavizza, H., and Comet, J.-P. (2021a). Greening R. Thomas' framework with environment variables: a divide and conquer approach. In *P*roceedings of the 19th Intl Conf. on CMSB, vol. 12881 of *L*NBI, pages 36–56.

Gibart, L. et al. (2021b). TotemBioNet enrichment methodology: Application to the qualitative regulatory network of the cell metabolism. In *P*roceedings of the 14th Intl Joint Conf. BIOINFORMATICS, volume 3, pages 85–92.

Gibart, L., Khoodeeram, R., et al. (2021c). Regulation of eukaryote metabolism: An abstract model explaining the warburg/crabtree effect. *P*rocesses, 9:1496.

Hart, B. A., Potts, R. J., and Watkin, R. D. (2001). Cadmium adaptation in the lung - a double-edged sword? *T*oxicology, 160(1-3):65–70.

Hernández, L. G., van Steeg, H., et al. (2009). Mechanisms of non-genotoxic carcinogens and importance of a weight of evidence approach. *M*utation Research/Reviews in Mutation Research, 682(2):94–109.

Hernández-Cruz, E. Y., Arancibia-Hernández, Y. L., et al. (2022). Oxidative stress and its role in cd-induced epigenetic modifications: Use of antioxidants as a possible preventive strategy. *O*xygen, 2(2):177–210.

Jing, Y., Liu, L.-Z., et al. (2012). Cadmium increases HIF-1 and VEGF expression through ROS, ERK, and AKT signaling pathways and induces malignant transformation of human bronchial epithelial cells. *T*oxicol. Sci., 125(1):10–19.

Khalis, Z. et al. (2009). The SMBioNet method for discovering models of gene regulatory networks. *G*enes, Genomes and Genomics, 3(special issue 1):15–22.

Khoodeeram, R. (2021). Discrete modelling of the energy metabolism regulation of eukaryotic cells and formal validation of its dynamics. modeling and simulation. *H*AL Theses.

Kim, J.-w., Tchernyshyov, I., et al. (2006). Hif-1-mediated expression of pyruvate dehydrogenase kinase: A metabolic switch required for cellular adaptation to hypoxia. *C*ell Metabolism, 3(3):177–185.

Larson-Casey, J. L., Gu, L., et al. (2020). Cadmium-mediated lung injury is exacerbated by the persistence of classically activated macrophages. *J*. Biol. Chem., 295(46):15754–15766.

Mateus, D., Gallois, J.-P., et al. (2007). Symbolic modeling of genetic regulatory networks. *J*BCB, 5(2B):627–640.

Nagao, A., Kobayashi, M., et al. (2019). Hif-1-dependent reprogramming of glucose metabolic pathway of cancer cells and its therapeutic significance. *I*nt. J. Mol. Sci., 20(2).

Nzengue, Y., Steiman, R., et al. (2008). Oxidative stress and DNA damage induced by cadmium in the human keratinocyte HaCaT cell line: Role of glutathione in the resistance to cadmium. *T*oxicology, 243(1):193–206.

Qu, F. and Zheng, W. (2024). Cadmium exposure: Mechanisms and pathways of toxicity and implications for human health. *T*oxics, 12(6):388.

Satarug, S., Garrett, S. H., et al. (2010). Cadmium, environmental exposure, and health outcomes. *E*nv. Health Perspectives, 118(2):182–190.

Semenza, G. L. (2012). Hypoxia-inducible factors in physiology and medicine. *C*ell, 148(3):399–408.

Snoussi, E. H. (1989). Qualitative dynamics of piecewise-linear differential equations: a discrete mapping approach. *D*ynamics and Stability of Systems, 4(3-4):565–583.

Thomas, R. (1991). Regulatory networks seen as asynchronous automata: A logical description. *J*ournal of Theoretical Biology, 153(1):1–23.

Thévenod, F. and Wing-Kee, L. (2024). Cadmium transport by mammalian atp-binding cassette transporters. *B*iometals, 37(3):697–719.

Waalkes, M. P. (2000). Cadmium carcinogenesis in review. *J*ournal of Inorganic Biochemistry, 79(1):241–244.

Wang, Y., Chi, H., et al. (2022). Cadmium chloride-induced apoptosis of hk-2 cells via interfering with mitochondrial respiratory chain. *E*cotoxicology and Environmental Safety, 236:113494.

Wu, H., Lu, T., et al. (2014). Sparse additive ordinary differential equations for dynamic gene regulatory network modeling. *J*ournal of the American Statistical Association, 109(506):700–716.