Predictive modelling of acute Promyelocytic leukaemia resistance to retinoic acid therapy

José A Sánchez-Villanueva, Lia N'guyen, Mathilde Poplineau, Estelle Duprez, Élisabeth Remy, Denis Thieffry

OXFORD

Briefings in Bioinformatics, 2025, **26(1)**, bbaf002
https://doi.org/10.1093/bib/bbaf002
Case Study

Predictive modelling of acute Promyelocytic leukaemia resistance to retinoic acid therapy

José A. Sánchez-Villanueva[1], Lia N'Guyen[2], Mathilde Poplineau[2], Estelle Duprez[2,*], Élisabeth Remy[1,*], Denis Thieffry[3,4,*]

[1]Aix Marseille Université, CNRS, I2M, 163 avenue de Luminy, 13009 Marseille, France
[2]Integrative molecular biology in hematopoiesis and leukemia, Equipe Labellisée Ligue Contre le Cancer, CRCM, Inserm UMR1068, CNRS UMR7258, Institut Paoli-Calmettes, Aix Marseille Univ, 27 Bd Lei Roure, 13009 Marseille, France
[3]Department of Biology, École Normale Supérieure, 46 rue d'Ulm, 75005 Paris, France
[4]Institut Curie - INSERM U900 - Mines Paris, PSL Research University, 26 rue d'Ulm, 75005 Paris, France
*Corresponding authors. Estelle Duprez, E-mail: estelle.duprez@inserm.fr; Denis Thieffry, E-mail: denis.thieffry@ens.fr; Élisabeth Remy, E-mail: elisabeth.remy@univ-amu.fr

Abstract

Acute Promyelocytic Leukaemia (APL) arises from an aberrant chromosomal translocation involving the Retinoic Acid Receptor Alpha (RARA) gene, predominantly with the Promyelocytic Leukaemia (*PML*) or Promyelocytic Leukaemia Zinc Finger (*PLZF*) genes. The resulting oncoproteins block the haematopoietic differentiation program promoting aberrant proliferative promyelocytes. Retinoic Acid (RA) therapy is successful in most of the PML::RARA patients, while PLZF::RARA patients frequently become resistant and relapse. Recent studies pointed to various underlying molecular components, but their precise contributions remain to be deciphered. We developed a logical network model integrating signalling, transcriptional, and epigenetic regulatory mechanisms, which captures key features of the APL cell responses to RA depending on the genetic background. The explicit inclusion of the histone methyltransferase EZH2 allowed the assessment of its role in the resistance mechanism, distinguishing between its canonical and non-canonical activities. The model dynamics was thoroughly analysed using tools integrated in the public software suite maintained by the *CoLoMoTo* consortium (https://colomoto.github.io/). The model serves as a solid basis to assess the roles of novel regulatory mechanisms, as well as to explore novel therapeutical approaches *in silico*.

Keywords: acute promyelocytic leukaemia; epigenetic regulation; gene regulatory network; logical model; therapy resistance.

Introduction

Acute Promyelocytic Leukaemia (APL) is a subtype of Acute Myeloid Leukaemia predominantly originating from an aberrant chromosomal translocation involving the Retinoic Acid Receptor Alpha (RARA) gene with other genes. The Promyelocytic Leukaemia (*PML*) gene is the most frequent fusion partner found in patients, followed by the Promyelocytic Leukaemia Zinc Finger (*PLZF*) gene [1]. The two oncoproteins (PML::RARA and PLZF::RARA) codified by these gene fusions interfere with the normal haematopoietic differentiation process and favour the malignant transformation and proliferation of promyelocytes [2].

Current therapeutic approaches involve the induction of differentiation of the malignant cells with Retinoic Acid (RA), which in combination with arsenic trioxide successfully achieves complete remission in almost 100% of PML::RARA APL patients [3]. However, in the cases of PLZF::RARA APL patients, resistance to treatment and relapse occur, presumably involving various transcriptional factors, epigenetic remodelling enzymes, and signalling pathways [4]. But how precisely these different factors contribute to treatment resistance and relapse remains to be deciphered (Fig. 1).

In this study, we constructed a detailed and comprehensive logical model to decipher the signalling and regulatory mechanisms at the basis of RA resistance in APL cells. The model allows to represent the three cell populations (WT, PML::RARA, and PLZF::RARA patients), in the context of three qualitative levels of RA (absence, physiological level, and pharmacological level of RA). Our model encompasses the key regulatory components, including PML::RARA and PLZF::RARA fusion proteins, and their interactions documented in the literature or inferred from the analysis of single cell functional genomic data. The model also incorporates key epigenetic regulatory mechanisms, focusing in particular on the role of the EZH2 protein of the Polycomb Complex 2 (PRC2).

To analyse the dynamical properties of the resulting model, we took advantage of the integration of several software tools in a common computational framework developed by the Consortium for Logical Models and Tools (*CoLoMoTo*, http://www.colomoto.org). Currently integrating over twenty logical modelling and analysis tools developed by different groups, this framework is available as a *conda* package or as a *Docker* container. Furthermore, the use of *Jupyter* interactive notebooks allows the development and extensive documentation of complex analysis workflows combining complementary *CoLoMoTo* tools [5]. Figure 2 shows an overview of the model analysis workflow followed for this study, while more detailed descriptions of the tools used are provided in the Methods section.

In brief, the stable states of the model recapitulate the main phenotypes of differentiated and aberrant proliferative cells induced by RA treatment in the different APL genetic backgrounds. The dynamical analysis of the model further identifies the crucial points underlying cell fate decision between

Received: August 23, 2024. **Revised:** December 9, 2024

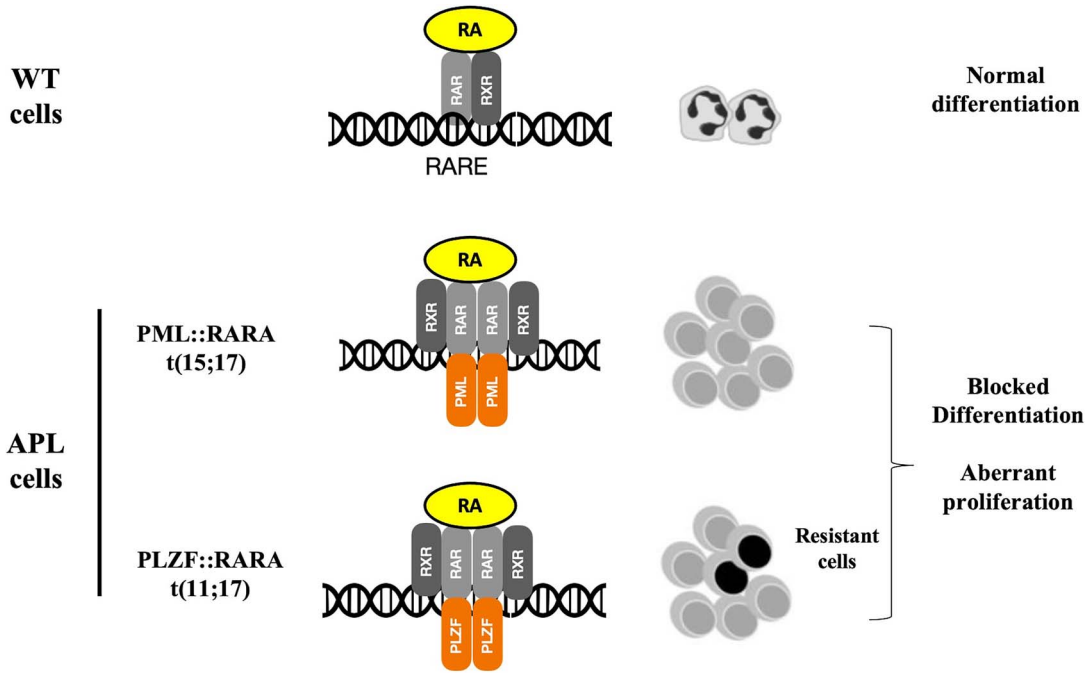

Figure 1. **Molecular mechanism underlying APL disease.** APL results from chromosomal translocation of RARA gene with the PML or the PLZF genes. Fusion proteins PML::RARA and PLZF::RARA drive the malignant transformation and proliferation of promyelocytes, blocking the differentiation and resulting in clonal growth of undifferentiated cells, which remain unresponsive to the normal physiological (nM) levels of RA. Current therapies are based on the induction of differentiation with pharmacological (μM) doses of RA. Resistance to this treatment and relapse involve interactions between the fusion proteins and several transcriptional factors, epigenetic remodelling enzymes and signalling components.

differentiation and aberrant proliferation. Furthermore, the simulations of EZH2 perturbations allow to distinguish the respective roles of canonical versus non-canonical activities of EZH2, and to highlight the key role of the non-canonical activity of EZH2 in the maintenance of the resistance to RA in APL PLZF::RARA cells.

Results

Construction of a logical model accounting for APL resistance to RA treatment

To build the logical model of the APL network, we gathered information from three main sources: (i) previously published computational models accounting for haematopoietic specification processes [6–10] (see Supplementary Table S1 available online at http://bib.oxfordjournals.org/ for a summary); (ii) a thorough literature review focusing on the distinct characteristics of the APL subtypes and their responses to RA therapy (for a review, see [4]); and (iii) the analysis of single cell omics data (RNAseq and ATACseq data) for RA-treated versus control cells from an APL mouse model expressing the PLZF::RARA translocation [11]. We have thus delineated the main regulatory and signalling mechanisms involved in the differentiation of myeloid cells in response to RA, as well as their blockade induced by APL fusion proteins.

Briefly, in the normal case (i.e. no fusion protein) and in the absence of RA, the Retinoic Acid Receptor (RARA) and the Retinoic X Receptor (RXR) protein heterodimers recruit epigenetic co-repressors (NCOR1, HDACs, PRC2, etc) to silence cell differentiation genes (CEBPA, CEBPB, CEBPE, SPI1, etc). The presence of the ligand RA induces a switch in the behaviour of the receptors, which results in the recruitment of epigenetic co-activators (NCOA1, PHF8, etc), which in turn enable the transactivation of genes provoking cell cycle arrest and triggering cell differentiation [12].

In the APL cases, the fusion proteins PML::RARA or PLZF::RARA predominantly recruit epigenetic co-repressors, even in the presence of RA physiological levels (nM range), which are insufficient to induce the recruitment of co-activators. Consequently, the differentiation program is blocked, causing cells to display an aberrant proliferative profile and to accumulate in the promyelocytic stage.

Pharmacological levels of RA (μM range) induce the proteasomal degradation of the PML::RARA fusion proteins and rescue the differentiation program in most of the patients with this variant [3]. In contrast, the stability of the PLZF::RARA fusion protein is maintained under high doses of RA, which causes the emergence of resistant phenotypes. Several epigenetic factors have been involved in this resistance mechanism, including HDAC enzymes, and the PRC1 and PRC2 complexes [2, 13–15].

Analysis of single cell omics data from the APL PLZF::RARA mouse model treated with RA revealed the presence of a cluster of cells with resistance features, characterised by the expression of genes involved in DNA replication and proliferation [11]. Interactions inference with SCENIC [16] revealed a transcriptional regulatory network controlled by E2F1 as a key element sustaining the resistance to RA [11]. The histone methyltransferase EZH2 was also identified as one of the most highly expressed genes in the resistant cluster. Further knocking down and pharmacological inhibition experiments pointed to the non-canonical activities of EZH2 as crucial for the E2F1 regulatory network and for the maintenance of the resistant phenotype [11].

Implemented with the *GINsim* software [17], the resulting regulatory graph encompasses 36 nodes and 131 arcs (Fig. 3). To each node is associated a Boolean variable representing the expression level of the component (see Methods), excepting the input node representing Retinoic Acid (RA), which is associated with a ternary variable (with three values 0, 1, and 2, corresponding to negligible, physiological, and pharmacological levels, respectively).

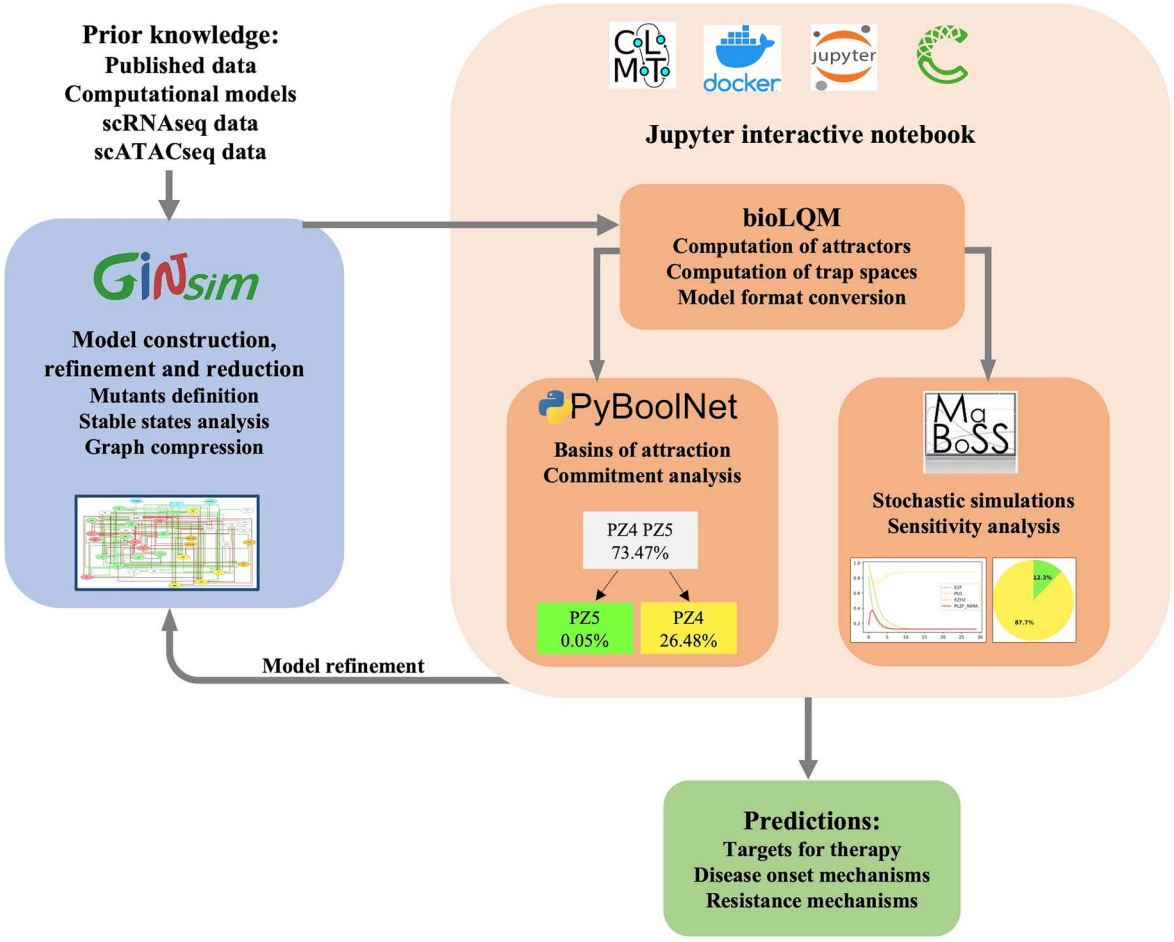

Figure 2. **Computational workflow used for the APL model construction, refinement and analysis.** A knowledge-based strategy was followed to build the APL model using the *GINsim* software (blue box). Tools of the *CoLoMoTo* package (*bioLQM*, *PyBoolNet*, *MaBoSS*) were used to analyse the dynamics of the model, following a workflow documented in a *Jupyter* interactive notebook (orange box). The resulting dynamical properties in turn served as a guide for model refinement and predictions (green box).

Furthermore, the activity level of each node is specified by a Boolean rule involving its regulators and logical operators (see Methods).

Besides the node RARA representing the wild-type RA receptor, and the nodes PML and PLZF representing respectively the wild-type PML and PLZF proteins, the input nodes PML_RARAg and PLZF_RARAg represent the two genetic fusions driving the APL disease. The consideration of different input combinations allows to evaluate separately the dynamical behaviour corresponding to a wild-type cell, to an APL cell expressing the PML::RARA translocation, or yet to an APL cell expressing the PLZF::RARA translocation, in response to different levels of RA. Noteworthy, in each of the two gene fusion situations considered, the model accounts for the simultaneous presence of wild-type RARA, PML, and PLZF encoded from the remaining wild-type alleles of these genes, as the fusion usually affects only one copy of the corresponding genes.

The APL model integrates signalling components (e.g. RARA, various kinases) and several key transcription factors (TP53, E2F1, SPI1, CEBPA, CEBPB, etc.) involved in the control of cell cycle of progenitor differentiation, together with key epigenetic regulatory factors and complexes, such as PRC2, other co-repressors (NCoR, HDACs) represented in the node CoRep_COMP, and the co-activatory factors (NCoA1, PHF8, etc) represented in the node CoAct_COMP. To the best of our knowledge, this is the first time that these three different levels of regulation have been seamlessly integrated in a full-fledge dynamical model.

To keep our logical model compact and dynamically coherent, we modelled receptor binding and regulatory complex formation in terms of directed arcs from the involved partners onto the regulatory targets, while co-occurrence requirements are encoded in logical rules. For example, while there is no arc between RA and RARA, the model includes arcs from both nodes onto each RARA target.

In our model, the stabilisation of the fusion protein (PLZF::RARA) presumably involves multiple positive regulatory circuits (also called positive feedback loops), which engage signalling components (CAK, USP37), transcription factors (CEBPA, CEBPB, MYC, E2F1, RB1, CDKN1A) and epigenetic regulators (EZH2, PRC2). More precisely, the stability of PLZF::RARA depends on the simultaneous presence of its three activators, PLZF::RARA coding gene, the CAK complex and the deubiquitinase USP37. The CAK complex is activated by MYC, whereas USP37 activation requires the presence of E2F1 and the absence of RB1.

In order to better visualise the organisation in our regulatory graph, we generated the strongly connected component graph (i.e. a graph gathering every subset of nodes mutually accessible in a single node) using an algorithm implemented in the software *GINsim*. The resulting graph comprises twelve components, including one central strongly connected component gathering 25

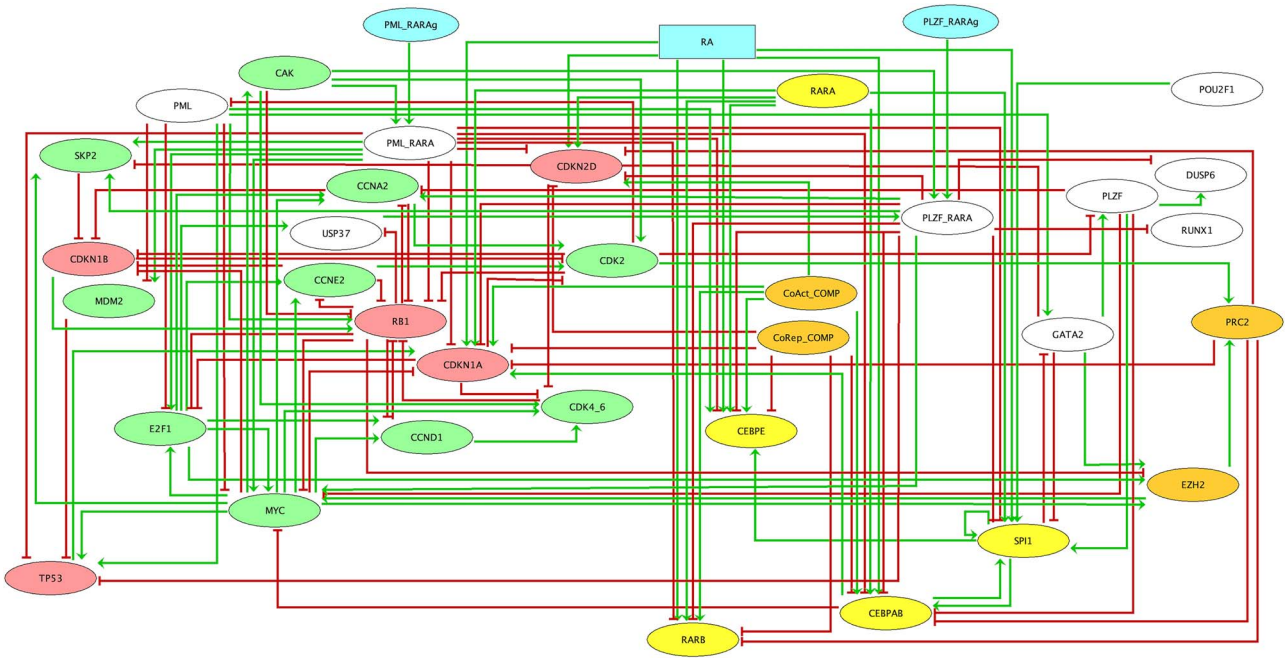

Figure 3. **Logical regulatory graph for the APL model.** Implemented with the *GINsim* software, this model integrates signalling, epigenetic and transcriptional regulatory factors to decipher the mechanisms underlying the resistance of APL cell in the course of RA treatment. This logical regulatory graph encompasses 36 nodes and 131 signed arcs. Internal nodes can be associated with four main functions: Cell cycle arrest (red); cell cycle progression (green); granulocytic differentiation (yellow); and epigenetic regulation (orange). Oval nodes denote Boolean components, whereas the rectangular one denotes a ternary component (RA). Green arrows and red blunt arcs denote positive and negative regulatory influences, respectively. Three input nodes (light blue) denote the presence of translocations (PLM_RARAg or PLZF_RARAg), as well as the RA level considered (No RA: 0, physiological level: 1, and pharmacological level: 2). The graph is complemented with logical rules determining the target level of each node as a function of the levels of its regulators. The model (including logical rules and model annotations) is available in the *GINsim* repository (zginml and SBML qual formats).

model nodes, which are involved in cell signalling, transcriptional regulation, and epigenetic remodelling (Supplementary Fig. S1 available online at http://bib.oxfordjournals.org/). The other components of this strongly connected component graph correspond to the seven input nodes and the four output nodes of the original graph. This organisation reflects the entanglement of the different regulatory layers underlying APL disease and resistance to RA.

The APL model is available in the *GINsim* repository, with the full list of logical rules and comprehensive annotations, in the native *GINsim* and *SBML* qual formats (http://ginsim.org/node/256). Furthermore, a supplementary file provides a comprehensive documentation of the model.

The stable states of the APL model recapitulate the behaviour of APL cells in response to RA treatment

The state space of interest splits into nine regions, one for each combination of input values considered, which correspond to the three genetic backgrounds (WT, APL PML::RARA, and APL PLZF::RARA) for the three different RA levels (absence, physiological or pharmacological levels). The input combinations accounting for the simultaneous presence of both fusion proteins (PML::RARA and PLZF::RARA) were excluded from this analysis, due to the lack of evidence of its existence in APL patients. As the APL model includes 29 nodes besides the seven input nodes, each of these regions thus encompasses $2^{29} = 536\,870\,912$ Boolean states.

Using an efficient algorithm included in the bioLQM Java library [18], we computed all the attractors in the APL model for each of the relevant input combinations (i.e. all possible combinations, excepting the simultaneous activation of PML_RARAg and PLZF_RARAg nodes). Of note, no cyclic attractors were detected with this analysis, a result that was further validated with the commitment analysis described in the following section. The resulting stable states are listed in the Fig. 4, focusing on the values of a subset of nodes enabling the association of each stable state with a specific cellular phenotype.

In the WT scenario, two stable states coexist in the absence of retinoic acid (RA = 0), corresponding to a proliferative phenotype (WT0) and a cell cycle arrest phenotype (WT1). For higher RA levels (level 1 or 2), a single stable state is obtained, which corresponds to a granulocytic differentiation phenotype (WT2 and WT3). Hence, our model recapitulates the strong differentiation effect of RA on normal cells observed in experimental settings [19].

Turning to the APL PML::RARA fusion scenario, in the absence of Retinoic Acid (RA = 0), we observe a bi-stability behaviour similar to the WT case (states PM0 and PM1). Under the physiological condition of RA (RA = 1), unlike the WT case, a proliferative stable state (PM3) coexists with a differentiated stable state (PM2). Interestingly, when RA is at pharmacological level (RA = 2), the proliferative stable state is lost, with only a differentiation stable state remaining (PM4). These results agree with published experimental and clinical data demonstrating that the PML::RARA fusion induces a form of APL responsive to high doses of RA, which trigger the differentiation of promyelocytes [20].

Finally, in the case of the APL PLZF::RARA fusion, in the absence or at physiological levels of RA, the model stable states are very similar to those obtained for the PML::RARA fusion. But interestingly, under pharmacological levels of RA, a proliferative stable state persists (PZ5), suggesting that cells do not (fully) respond to RA differentiation stimuli, even at pharmacological levels. This result is coherent with the observation of resistance to high

		RA	PML	PLZF	EZH2	PRC2	CDKN1A	CDKN1B	TP53	E2F1	MYC	SPI1	CEBPAB	RARB
WT	WT0	0	0	0	1	1	0	0	0	1	1	0	0	0
	WT1	0	1	1	1	0	1	1	1	0	0	0	0	0
	WT2	1	1	0	0	0	1	1	1	0	0	1	1	1
	WT3	2	1	0	0	0	1	1	1	0	0	1	1	1

		RA	PML	PLZF	EZH2	PRC2	CDKN1A	CDKN1B	TP53	E2F1	MYC	SPI1	CEBPAB	RARB
APL PML::RARA	PM0	0	0	0	1	1	0	0	0	1	1	0	0	0
	PM1	0	1	1	1	0	1	1	1	0	0	0	0	0
	PM2	1	1	0	0	0	1	1	1	0	0	1	1	1
	PM3	1	0	0	1	1	0	0	0	1	1	0	0	0
	PM4	2	1	0	0	0	1	1	1	0	0	1	1	1

		RA	PML	PLZF	EZH2	PRC2	CDKN1A	CDKN1B	TP53	E2F1	MYC	SPI1	CEBPAB	RARB
APL PLZF::RARA	PZ0	0	0	0	1	1	0	0	0	1	1	0	0	0
	PZ1	0	1	1	1	0	1	1	1	0	0	0	0	0
	PZ2	1	1	0	0	0	1	1	1	0	0	1	1	1
	PZ3	1	0	0	1	1	0	0	0	1	1	0	0	0
	PZ4	2	1	0	0	0	1	1	1	0	0	1	1	1
*	PZ5	2	0	0	1	1	0	0	0	1	1	0	0	0

Figure 4. **Stable states for the APL model.** The three tables list the stable states of the model for the wild-type (WT, top), the APL PML::RARA (PM, middle), and the APL PLZF::RARA (PZ, bottom) conditions, respectively. Each row represents a stable state characterised by the values of representative model components (cf. Figure 3). The first column associates a label with each stable state, with a colour denoting the corresponding phenotype: Proliferative (green), arrested (red) or differentiated (yellow). Each cell denotes the level of activity of the corresponding component (column) for the corresponding stable state (row): 0 (white background), 1 (light background) or 2 (dark background, for RA). The stable state labelled as PZ5 and highlighted with a star corresponds to the proliferative phenotype resistant to pharmacological doses of RA. The computation of the stable states was done with the *BioLQM* software included in the *CoLoMoTo* package and can be reproduced with the notebook provided together with the model in the GINsim model repository.

dose of RA in APL PLZF::RARA cells in experimental and clinical settings [20].

A cell fate commitment analysis rules out the presence of cyclic attractors and accounts for the decision between differentiation and aberrant proliferation

The *PyBoolNet* software was used to generate a broad overview of the full state space, revealing state commitment to each attractor [21, 22] (see also Methods). The comparison of the sizes of the basins of different attractors provides interesting hints regarding the likelihood of reaching each attractor, depending on genetic background and treatment.

The commitment analysis of the APL model outputs nine graphs, one per combination of inputs considered, corresponding to the WT, APL PML::RARA, and APL PLZF::RARA cases, each in the absence or under the presence of physiological or pharmacological levels of RA (Fig. 5). The nodes of the commitment graphs represent sets of states with common specific commitment potential, while the arcs connecting nodes denote the existence of pathways linking the states of the two connected state sets in the underlying state transition graph (STG, cf. Methods).

For the WT case with RA = 1 or 2, and for the APL PML::RARA case with RA = 2, the commitment graphs contain a single node gathering all the states, meaning that there is a unique attractor. Of note, these three situations lead to the same differentiation stable state. For the six remaining situations (WT with RA = 0, APL PML::RARA with RA = 0,1, and APL PLZF::RARA with RA = 0,1,2), the dynamics present two stable states. Each of the six resulting commitments graphs contain three nodes. Two nodes correspond to the stable states (representing proliferation versus differentiation cell fates) together with their basins of attraction, while the upper node corresponds to uncommitted cells, for which cell fate decision is still pending. In each of these graphs, the upper node gathers the majority (~ 74%) of the states (uncommitted progenitors). We further observe that a small proportion (less than 0.05%) of the states are fully committed to the stable state denoting cell proliferation, i.e. potential leukaemia initiating cells. Notably, in the PLZF::RARA scenario, the sizes of these basins are not affected by the increase in RA dose, which is coherent with the resistance and relapse features observed in this APL subtype. Finally, the third node of these graphs gathers the remaining states (~ 25%–26% of all states), which are uniquely committed to either cell differentiation or to cell cycle arrest.

Of note, the commitment analysis rules out the presence of cyclic attractors in the dynamics of the network. The analysis also reveals the different behaviours of the two APL subtypes in response to pharmacological RA in terms of stable states. In

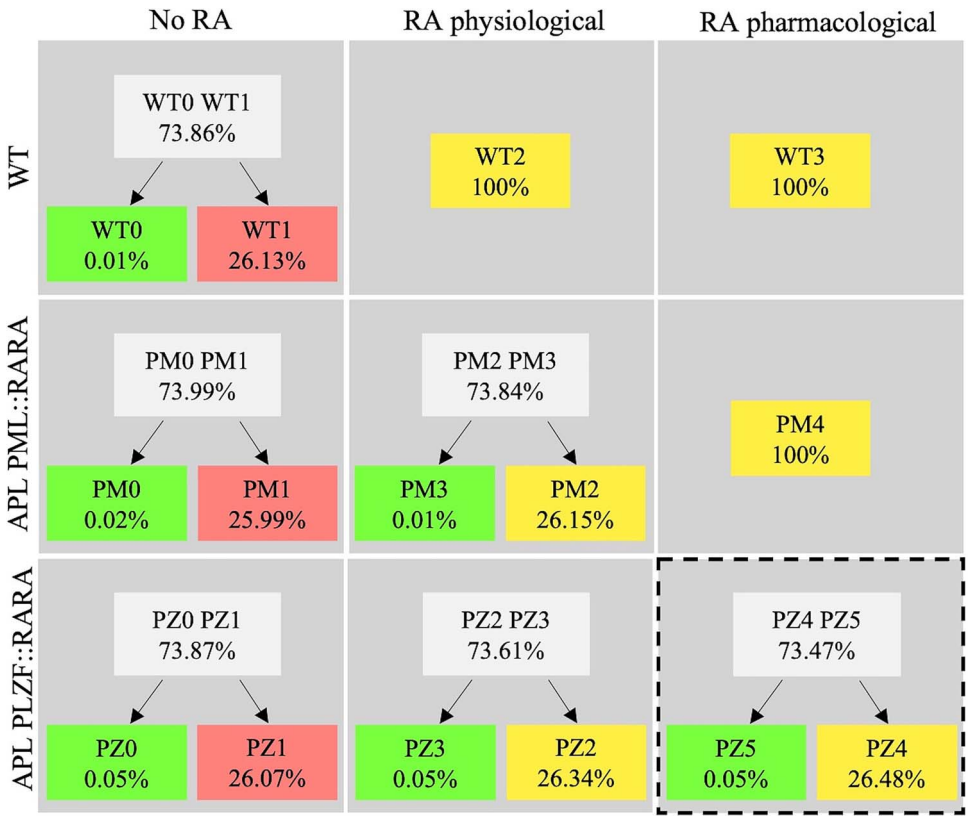

Figure 5. **Cell fate commitment analysis.** A) Cell fate commitment diagrams were computed for each relevant input combination (RA levels 0, 1, and 2) for each genetic background (WT, PLM::RARA, PLZF::RARA). Each node represents the subset of states committed to the corresponding labelled stable state(s), the percentage indicating the corresponding proportion of full state space. The colours of the nodes denote the corresponding phenotypes: Green for proliferative, red for arrested, yellow for differentiated. The arcs denote the existence of transitions between sets of states. These commitment diagrams were computed with the *PyBoolNet* software, included in the *CoLoMoTo* package. The case of PLZF::RARA cells under pharmacological doses of RA, which is the scenario for resistance and relapse, is surrounded by a dashed rectangle.

the PML::RARA APL case, there is a unique differentiated stable state, whereas an aberrant proliferative stable state coexists with a differentiated stable state in the PLZF::RARA APL case. This behaviour qualitatively matches the phenotypes observed in experimental and clinical observations [3, 4, 20].

Focusing on the impact of the presence of the PLZF::RARA fusion on cell fate decision, we aimed at further identifying the key transitions underlying stable state commitment. This required more accurate tracking of trajectories in the dynamics and monitoring of updated components along the trajectories. For this, we took advantage of an algorithm implemented in *GINsim* to reduce our model and thereby ease simulations (see Methods). We thus performed a simulation using a reduced model encompassing 24 nodes, starting from an initial state representing cells in the granulocyte-monocyte progenitor stage (GMP), expressing the PLZF::RARA fusion, and submitted to pharmacological dose of RA. As the resulting state transition graph (STG, cf. Methods) is too large to be displayed (it encompasses 12,296 states), we compressed it into a hierarchical transition graph (HTG, cf. Methods), which is displayed in Fig. 6. In this HTG, rectangular nodes represent stable states, while each oval node gathers a set of states sharing the same stable state reachability properties. *GINsim* further enables the labelling of each arc of the HTG with the corresponding transitions. Looking at the component values of the 191 states fully committed towards the aberrant proliferative attractor PZ5 (listed on the right of Fig. 6), we can visualise which nodes are frozen at specific values, including CEBPA, CDN1A, CDN1B, and RB1 all set to 0, and CCN2, E2F1, MYC, and POU2F1 all set to 1 in the HTG node. This analysis reveals the key roles of the inhibitions of CDKN1A and CEBPA/B in the commitment towards the proliferative fate, as well as of the activation of CDKN1A and of the inhibitions of E2F1 and MYC for the commitment towards the cell differentiation fate.

Simulation of EZH2 perturbations highlights the key role of the non-canonical activity of EZH2 in maintaining the resistant phenotype

The experimental results in the APL PLZF::RARA mouse model treated with RA pointed at the Polycomb complex subunit EZH2 as a key driver of the RA resistance. More specifically, the inhibition of EZH2 canonical enzymatic activity did not suppress the aberrant proliferative phenotype, while a complete degradation of EZH2 erased RA resistance [11].

Based on these observations, we designed two *in silico* experiments to assess the impact of EZH2 perturbations on the response of APL cells to RA treatments. The software bioLQM was used to define two variants of the APL model corresponding to (i) a complete knockdown of EZH2, or the use of a specific degrader such as MS1943 (model variant EZH2%0, also called EZH2 KO hereafter); (ii) an edgetic perturbation amounting to block the regulatory influence of EZH2 activity onto PCR2, mimicking the experimental inhibition of EZH2 methyltransferase activity with a chemical agent such as GSK126 (model variant EZH2:PRC2%0) [11]. With this approach, we implemented and tested the hypothesis of an additional non-canonical activity of EZH2 as a key mechanism in the maintenance of the resistant phenotype.

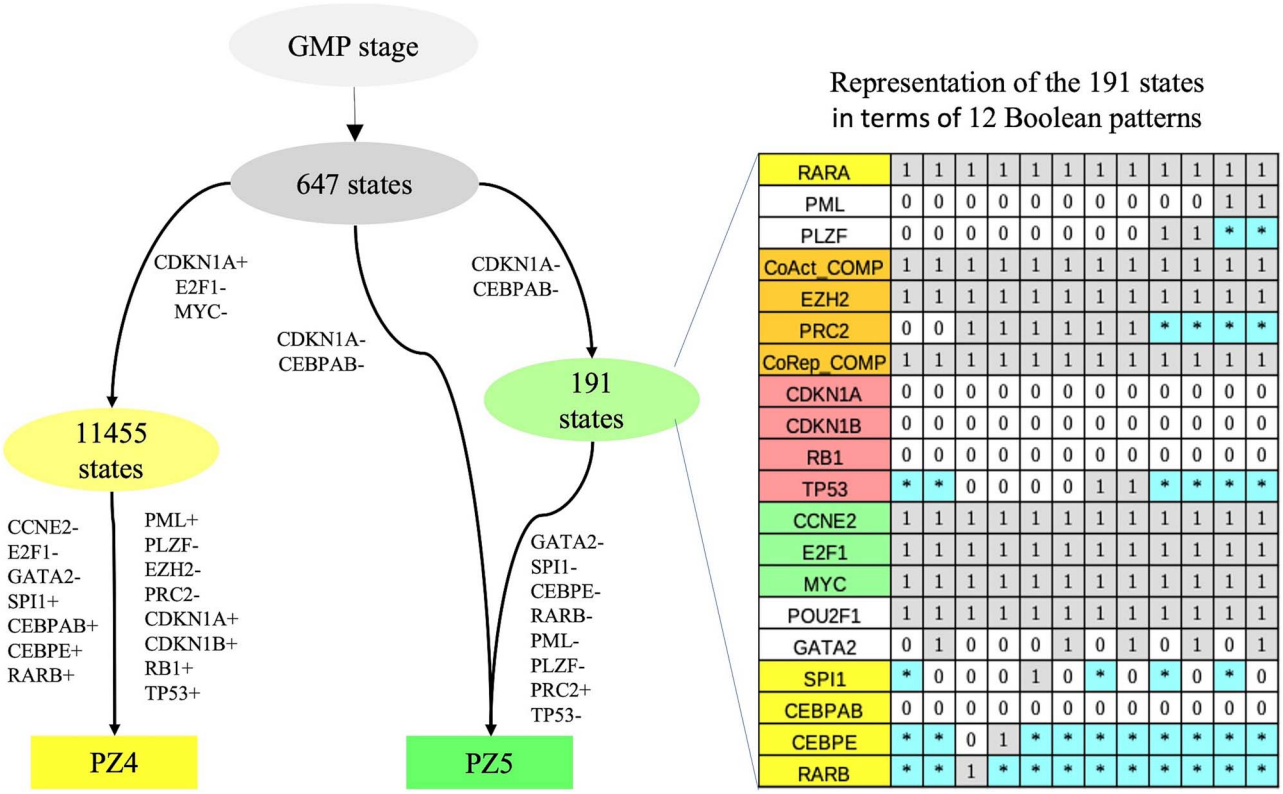

Figure 6. **Hierarchical transition graph.** Hierarchical transition graph (HTG) obtained for a reduced version of the model, using asynchronous updating. Each oval node represents a set of states associated with identical commitment potential, labelled with the corresponding number of states; each rectangular node denotes an attractor, i.e. the stable states PZ4 and PZ5 shown in Fig. 4. The arcs are labelled with the transitions underlying commitment decisions. The computation of the HTG was done with the *GINsim* software, selecting an initial state representative of the granulocytic/myeloid progenitor stage (RA set to 2, PLZF_RARAg, RARA, PML, PLZF, USP37, CoAct_COMP, EZH2, PRC2, CoRep_COMP, CDKN2D, RB1, SKP2, MDM2, CAK, CCNA2, CCND1, CCNE2, CDK4_6, CDK2, E2F1, MYC, DUSP6, POU2F1, RUNX1, SPI1, CEBPAB and CEBPE all set to 1, while the rest of the nodes were set to 0). The 191 states committed to the PZ5 attractor are listed in the table on the right. Each pattern (column) represents a set of states, while each row denote the level of activity of a model components (row): 0 (white background), 1 (light background) or free to take any of these two values (denoted by a star with a blue background).

Interestingly, the EZH2:PRC2%0 variant model displays the same stable states as the original model, while a complete loss of EZH2 leads to the loss of the proliferative stable state (cf. notebook provided with the model in the GINsim repository). To better characterise these differences, we took advantage of the software *MaBoSS* (included in the *CoLoMoTo* package) [23] to perform stochastic simulations with the two model variants, using the initial state representing cells in the GMP stage, expressing the PLZF::RARA fusion and submitted to pharmacological dose of RA (described in the caption of in Fig. 6), with default equiprobable transition rates (i.e. all transition rates were left to the default value 1).

10 000 *MaBoSS* simulations were computed for each of the three conditions and average results are displayed in Fig. 7.

In the absence of perturbation (Fig. 7, top), the differentiation marker SPI1 relatively quickly sets to a high probability (~0.88), while the proliferation marker E2F1 sets to a much lower probability (~0.12). This result is consistent with the experimental and clinical observations according to which RA exerts a potent differentiation effect, but relapses occur due to residual RA-resistant cells [4].

In the case of the edgetic perturbation of EZH2 activity (EZH2:PRC2%0), *MaBoSS* simulations result are very similar to the EZH2 WT case, as shown in the corresponding time-plots and pie charts. This result is also coherent with published experimental results, according to which the inhibition of EZH2 enzymatic activity with GSK126 barely affects the emergence RA-resistant clones [11].

Finally, in the case of a complete knock-down of EZH2, in coherence with the stable states analysis, *MaBoSS* simulations reproduce the loss of the proliferative stable state. The complete loss of the proliferative phenotype is in line with the published efficiency of EZH2 degrader MS1943 to eliminate RA-resistant clones [11].

More perturbation simulations are reported in the notebook provided with the model in the GINsim repository. These additional perturbations show the loss of the proliferative phenotype with the full knock-down of EZH2 even under physiological levels of RA (RA = 1).

Together, our model-based predictions points to a crucial role of EZH2 in RA resistance and leukemogenesis, which could be relevant for other AML subtypes [24]. This result further highlights the potential of EZH2 degraders for combinatorial therapeutic strategies.

A parameter sensitivity analysis delineates the main components impacting cell fate decision

Next, still focusing on the PLZF::RARA genetic background, we performed a parameter sensitivity analysis to evaluate the robustness of the model and to further assess the impact of each component of the network in cell fate decision. The software *MaBoSS* enables the configuration of two main types of perturbations for each node: full activations or inactivations, which reproduce loss- or gain-of-function experiments, or the assignment of specific transition rates, which enables simulations with different timescales [25].

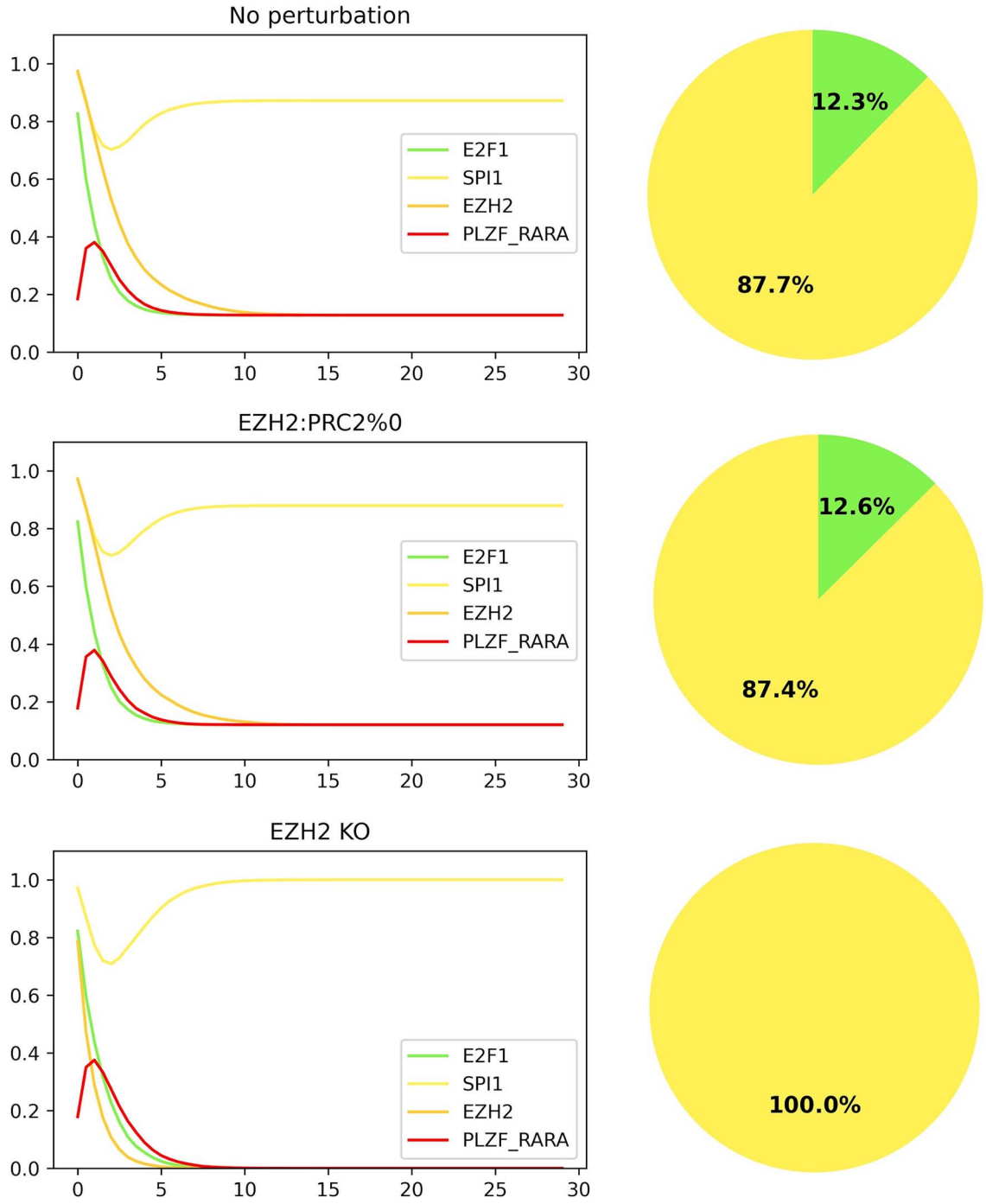

Figure 7. **Stochastic simulations assessing the effects of EZH2 perturbations.** The initial state defined in Fig. 6 was selected to perform stochastic simulations of PLZF::RARA cells under three different conditions: in the absence of perturbations (above), for an EZH2 perturbation blocking only its regulatory effect on PRC2 (middle), and for a complete EZH2 loss-of-function (below). The time plots (left) display the evolution of the probabilities (between 0 and 1) of activation of representative components over time (arbitrary units), which denote proliferative versus differentiated phenotypes. The piecharts (right) display the probabilities of the differentiated (yellow) versus proliferative (green) phenotypes at the end of these simulations. 10 000 simulations were performed to compute these probabilities with the *MaBoSS* software. These simulations can be reproduced with the notebook provided in with the model in the GINsim repository.

Starting from the same GMP representative state, simulations with different transition rates for the activation and the inactivation of each node pointed to the nodes whose perturbations partly or fully affect the reachability from the initial GMP state to the differentiated or aberrant proliferative state (Fig. 8). Noteworthy, a decrease of the activation transition rate of PLZF_RARA, or an increase of the inactivation transition rate of MYC, both greatly reduce the probability to reach the proliferative attractor (Fig. 8A). This result points to the role of PLZF-RARA and MYC in promoting RA-resistant clones. Perturbations affecting the transition rates of CDKN1A, CEBPAB, and E2F1 also strongly impact the probability of reaching the proliferative state (Fig. 8B). This further supports the crucial function of these nodes in the cell fate commitment decision, in coherence with the HTG of Fig. 6. Finally, this

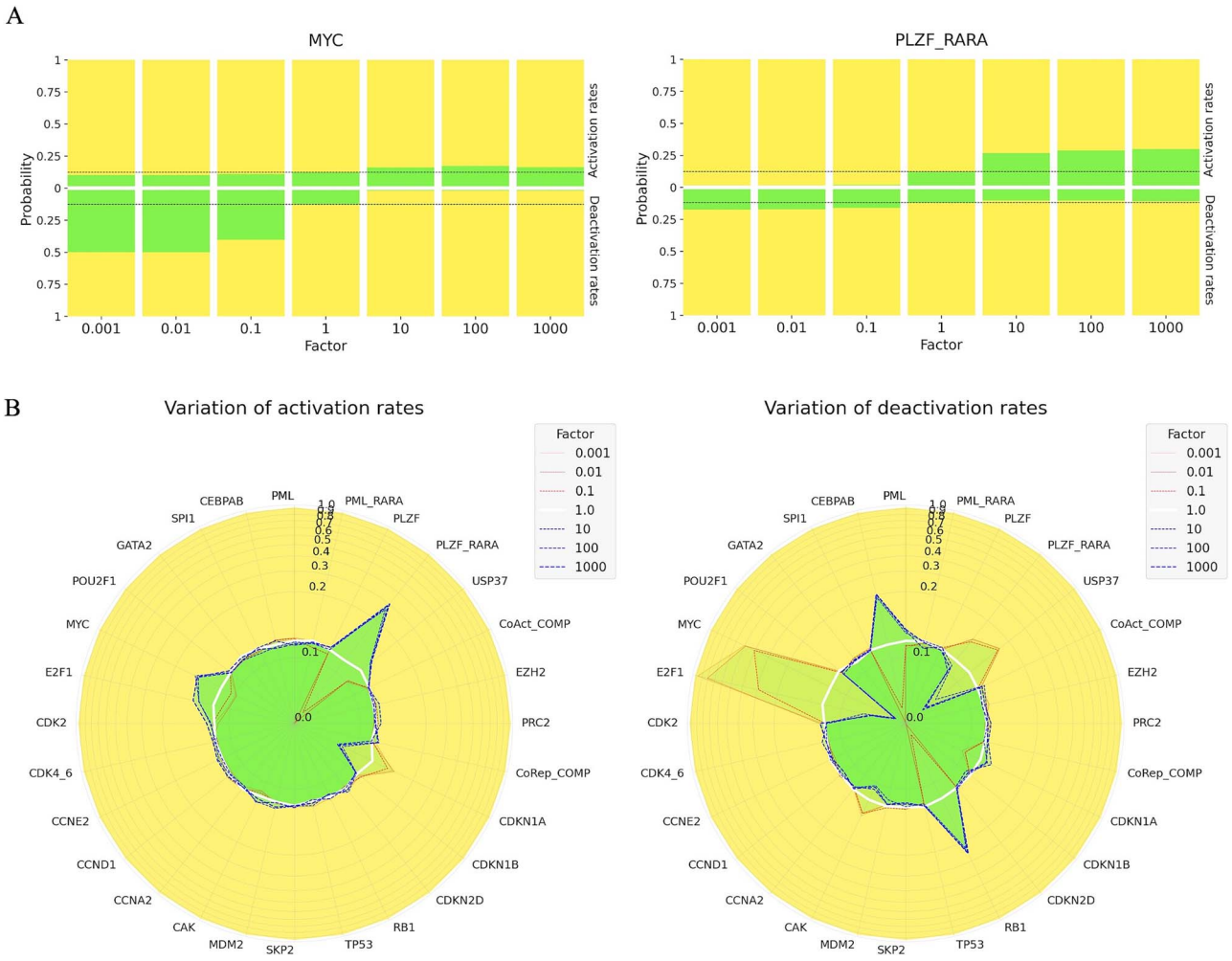

Figure 8. **Parameter sensitivity analysis of the APL model.** The initial state defined in Fig. 6 was selected to perform stochastic simulations of PLZF::RARA cells under variations of the activation or inactivation transition rates for each node (excepting input and output nodes), by multiplying them by one of the following factors: 0.001; 0.01; 0.1; 1; 10; 100; 1000. (A) Bi-directional bar plots displaying the sensitivity to the transition rates of the nodes MYC (left) and PLZF_RARA (right). In each vertical bar, the upper half corresponds to the activation rates, while the bottom half corresponds to the deactivation rates. The yellow part indicates the probability of differentiation, while the green part indicates the probability of proliferation. The black horizontal dashed lines denote the proportions of proliferation and differentiation for the default parameters set to 1. (B) Polar plot representation of the impact of the variations of the activation (left) and deactivation (right) rates for the full set of nodes. The probability (between 0 and 1) of the proliferative attractor under each condition is plotted as a vertex of a polygon, with the interior area filled in green. The default transition rate of 1 (white) is used as reference. Blue polygons represent transition rates higher than the default, whereas red polygons represent transition rates lower than the default. 10 000 simulations were performed to compute these probabilities with the *MaBoSS* software, with a time range set to 20 000 to enable the network to reach equilibrium. These simulations can be reproduced with the notebook provided with the model in the GINsim repository.

analysis identifies additional nodes affecting the probability of the proliferative state, including USP37 and CAK, which were reduced for the generation of the HTG shown in Fig. 6.

The results of our comprehensive parameter sensitivity analysis are reported in the notebook with the model in the GINsim repository.

Discussion

In this manuscript, we report the construction and the dynamical analysis of a logical model integrating the main signalling, transcriptional and epigenetic regulatory components driving the response of APL cells to RA therapy.

From a methodological point of view, we designed a robust computational workflow combining four different tools, developed by different groups, but seamlessly integrated in the *CoLoMoTo* software environment. The combination of these tools enabled us to refine our model and reveal its most salient dynamical properties. The integration of the tools in a common framework together with the use of *Jupyter* notebooks further offers solid guarantees for the reproducibility of our computational results, and simultaneously greatly eases further refinements or extensions of the current study.

From a biological point of view, our computational approach enabled us to capture and reproduce the general behaviour of the two most common APL variants in response to different levels of RA, as well as to simulate more refined details of the entangled regulatory mechanisms sustaining the resistance to RA. The explicit representation of the two aberrant translocations in addition to the wild type genes allows to implement and test mechanistic hypotheses about the stability of the proteins and its implication in the maintenance of the resistance, which is crucial to the identification of key pathways in response to treatment. Furthermore, as it explicitly integrates signalling, transcriptional and epigenetic regulatory components, our APL model opens novel prospects to decipher

the interplay between these different regulatory layers in cancer cells.

The representation of EZH2 as a separated node from the PRC2 complex enabled us to distinguish two types of EZH2 activities: (i) its canonical enzymatic activity, i.e. the catalysis of the repressive histone mark H3K27me3; and (ii) its so-called non-canonical activities, such as the methylation of non-histone proteins or the stabilisation of interaction partners [26, 27]. In this study, we focused on the impact of EZH2 non-canonical activity on the stability of the MYC oncogene, which presumably constitutes an important mechanism underlying RA resistance. The results from the EZH2 perturbation analysis and stochastic simulations recapitulate the experimental observations [11]. Hence, with our strategy, we successfully implemented the hypothesis of multiple concurrent activities of EZH2 into a predictive computational model, and assessed their respective impacts on a concrete scenario of leukemogenesis and therapy resistance.

Our model has proven to be useful to interrogate RA resistance in APLs and could be extended to better evaluate the EZH2 activity and its targets involved in RA resistance. A potentially interesting extension of this study would be the modelling of the contribution of concurrent genetic mutations to RA-resistance and relapse in APL, which has been frequently observed in the clinics and typically emerge in patients with the PML::RARA genetic background [28]. However, this would require the inclusion of additional signalling proteins and transcriptional factors, such as FLT3 and WT1, together with their regulatory interactions with current network components, which remain to be properly characterised. The impact of additional mutations could then be explored in the same way as we did for EZH2, or in a subtler manner, by tuning the rates used in *MaBoSS* simulations.

Methods

The logical modelling framework

Our modelling approach relies primarily on the delineation of a *regulatory graph*, i.e. a signed directed graph where each node represents a biological component, connected to others by means of signed arcs representing activations and inhibitions. A discrete variable is associated with each node, representing the expression level of the component. For the APL model, these nodes are each associated with a Boolean variable, which denotes negligible presence or inactivity (level 0), or significant presence or activity (level 1) of the component. An exception was made for the input node representing Retinoic Acid (RA), which is associated with a ternary variable. The three values of RA node respectively denote negligible concentrations (level 0), physiological differentiation doses (level 1), and higher pharmacological doses (level 2).

Logical rules are further associated with each node of the graph to model its behaviour as a function of the state of its direct regulators, expressed in terms of a combination of the corresponding Boolean variables with the logical operators AND (&), OR (|) and NOT (!). In some cases, we could rely on existing logical models in the literature and adopt the proposed rules. For novel nodes, we first used a default rule stating that the target node gets activated only in the presence of all its activators and none of its inhibitors. Iterative adjustments and refinements of these rules were then made to match experimental observations, in particular to obtain consistent sets of stable states.

The dynamical behaviours of Boolean networks are usually displayed in terms of *state transition graphs* (STG), i.e. directed graphs where each node denotes a state of the network (Boolean vector), and where the arcs denote transitions between pairs of states, which are enabled by the rules. We used an *asynchronous updating* scheme, where only one node can update its level at a time. The resulting dynamics is non deterministic and usually more complex than the dynamics obtained with the more simplistic synchronous updating scheme [29].

The characterisation of the dynamical properties of a logical model typically starts with the identification of the *attractors*, which represent the asymptotic behaviour of the model, and are usually associated with observed phenotypes in the experimental or clinical settings. The attractors can be of two types: *stable states*, with no update call (also called *fixpoints*), or *cyclic attractors*, made of circular sequences of two or more states. The sets of states (exclusively) leading to a given attractor defines its (exclusive) *basin of attraction* [22].

Computational tools

We used the software *GINsim* (v3.0.0b, http://ginsim.org) to build, edit and analyse the logical model of the APL network [17].

To analyse the dynamical properties of the APL model, we used a combination of software tools selected and assembled by the Consortium for Logical Model and Tools (*CoLoMoTo*, http://www.colomoto.org) [5]. Each of these tools provides specific functionalities in an interoperable framework, making the analysis workflow robust and reproducible (Fig. 2).

GINsim was used to build the regulatory graph (RG) defining our logical model, encode the logical rules associated with the different components of this graph, as well as to annotate each node. We further used an algorithm based on the Tarjan algorithm and implemented in *GINsim* to compress the RG into a *strongly connected component graph* (SCCG, see Supplementary Fig. S1 available online at http://bib.oxfordjournals.org/), which merge the RG nodes sharing mutual reaching properties [30].

GINsim was also used to generate state transition graphs (STG) with the asynchronous updating scheme. For this, a reduced model was defined using an algorithm implemented in *GINsim*, which takes a list of nodes to be reduced in input and iteratively recompute the logical rules of the targets of the reduced nodes. The reduced model preserves the stable states and other crucial dynamical properties of the original model [31]. Using a Tarjan-based algorithm implemented in *GINsim*, STG were further compressed into *Hierarchical State Transition graphs* (HTG, enabling a stronger compression compared to SCCG), where the sets of states sharing attractor reachability properties are merged into single nodes, which are then connected by arcs denoting changes in attractor reachability [30].

GINsim enables the encoding of different kinds of genetic or biochemical perturbations, including gain-of-function mutations (blocking of the values of the corresponding nodes to the level 1), loss-of-functions or full knock-downs (blocking the values of the corresponding nodes to the level 0), and the perturbation of specific interactions (aka edgetic perturbations) [17]. Similar functionalities are implemented in the bioLQM library (see below), also available in the *CoLoMoTo* package.

We used the *bioLQM* Java library W (http://www.colomoto.org/biolqm/index.html) for the computation of the attractors (amounting here to stable states), for the definition of variants of the model implementing specific perturbations (e.g. loss-of-function mutants), as well as for the conversion and export of models between tools of the *CoLoMoTo* package [5, 18].

The software *Pyboolnet* (https://github.com/hklarner/PyBoolNet) was used for the computation of the basins of attraction of the stable states of the model, as well as to perform an attractor commitment analysis. Given a Boolean model, *PyBoolNet* [21] relies

on the Computational Tree Logic and uses generic model checkers to compute the sets of states leading to the different attractors (i.e. the basins of attraction of these attractors), as well as to compute reachability properties [22].

Finally, we used the tool *MaBoSS* (https://maboss.curie.fr) to perform stochastic Boolean simulations. Relying on continuous time Markovian processes and on the Gillespie algorithm, *MaBoSS* enables the computation of the evolution of Boolean state probabilities over time [23].

Key Points

- Based on single cell multi-omics data and on an extensive literature revision, we built a logical model capturing and explaining key features of the responses of two main subtypes of Acute Promyelocytic Leukaemia (APL) cells to retinoic acid (RA) therapy.
- The stable states of the model recapitulate the phenotypes of differentiated and aberrant proliferative cells induced by RA treatment for the different APL genetic backgrounds.
- A commitment analysis identifies the crucial components underlying the decision between cell differentiation and aberrant proliferation.
- The simulations of different EZH2 perturbations allow to differentiate/distinguish the canonical and non-canonical activities of EZH2 and highlight the key role of the non-canonical activity of EZH2 in the maintenance of the resistance to RA in APL PLZF::RARA cells.
- A parameter sensitivity analysis further characterises the components of the network underlying cell fate decisions, pointing to potential targets for novel combinatorial therapy strategies.

Acknowledgements

The authors warmly thank Claudine Chaouiya for providing various useful suggestions on a previous version of this manuscript, Loic Paulevé for updating and maintaining the *CoLoMoTo* software environment, as well as Pedro Monteiro for managing the *GINsim* model repository.

Supplementary data

Supplementary data are available at *Briefings in Bioinformatics* online.

Conflict of interest: None declared.

Funding

This work has been supported by the French Plan Cancer-MIC ITMO, Project APL EpiNet (No. 21CM045-00).

Disclosure and competing interest statement

The authors state they have no competing interests or disclosures.

Data availability

The APL model is available in both native GINsim and SBML qual formats in the GINsim model repository, together with a python notebook enabling the reproduction of wild-type and mutant dynamical analysis results (http://ginsim.org/node/256).

References

1. dos Santos GA, Kats L, Pandolfi PP. *et al.* Synergy against PML-RARa: targeting transcription, proteolysis, differentiation, and self-renewal in acute promyelocytic leukemia. *J Exp Med* 2013;**210**:2793–802. https://doi.org/10.1084/jem.20131121.
2. Grignani F, De Matteis S, Nervi C. *et al.* Fusion proteins of the retinoic acid receptor-alpha recruit histone deacetylase in promyelocytic leukaemia. *Nature* 1998;**391**:815–8. https://doi.org/10.1038/35901.
3. Lo-Coco F, Avvisati G, Vignetti M. *et al.* Retinoic acid and arsenic trioxide for acute promyelocytic leukemia. *N Engl J Med* 2013;**369**:111–21. https://doi.org/10.1056/NEJMoa1300874.
4. Jimenez JJ, Chale RS, Abad AC. *et al.* Acute promyelocytic leukemia (APL): a review of the literature. *Oncotarget* 2020;**11**:992–1003. https://doi.org/10.18632/oncotarget.27513.
5. Naldi A, Hernandez C, Levy N. *et al.* The CoLoMoTo interactive notebook: accessible and reproducible computational analyses for qualitative biological networks. *Front Physiol* 2018;**9**:680. https://doi.org/10.3389/fphys.2018.00680.
6. Krumsiek J, Marr C, Schroeder T. *et al.* Hierarchical differentiation of myeloid progenitors is encoded in the transcription factor network. *PloS One* 2011;**6**:e22649. https://doi.org/10.1371/journal.pone.0022649.
7. Yuan R, Zhu X, Radich JP. *et al.* From molecular interaction to acute promyelocytic leukemia: calculating leukemogenesis and remission from endogenous molecular-cellular network. *Sci Rep* 2016;**6**:24307. https://doi.org/10.1038/srep24307.
8. Collombet S, Van Oevelen C, Ortega JLS. *et al.* Logical modeling of lymphoid and myeloid cell specification and transdifferentiation. *Proc Natl Acad Sci U S A* 2017;**114**:5792–9.
9. Wang M, Wang J, Zhang X. *et al.* The complex landscape of haematopoietic lineage commitments is encoded in the coarse-grained endogenous network. *R Soc Open Sci* 2021;**8**:211289. https://doi.org/10.1098/rsos.211289.
10. Hérault L, Poplineau M, Duprez E. *et al.* A novel Boolean network inference strategy to model early hematopoiesis aging. *Comput Struct Biotechnol J* 2022;**21**:21–33. https://doi.org/10.1016/j.csbj.2022.10.040.
11. Poplineau M, Platet N, Mazuel A. *et al.* Noncanonical EZH2 drives retinoic acid resistance of variant acute promyelocytic leukemias. *Blood* 2022;**140**:2358–70. https://doi.org/10.1182/blood.2022015668.
12. Kishimoto M, Fujiki R, Takezawa S. *et al.* Nuclear receptor mediated gene regulation through chromatin remodeling and histone modifications. *Endocr J* 2006;**53**:157–72. https://doi.org/10.1507/endocrj.53.157.
13. Boukarabila H, Saurin AJ, Batsché E. *et al.* The PRC1 Polycomb group complex interacts with PLZF/RARA to mediate leukemic transformation. *Genes Dev* 2009;**23**:1195–206. https://doi.org/10.1101/gad.512009.
14. He LZ, Guidez F, Tribioli C. *et al.* Distinct interactions of PML-RARalpha and PLZF-RARalpha with co-repressors determine differential responses to RA in APL. *Nat Genet* 1998;**18**:126–35. https://doi.org/10.1038/ng0298-126.
15. Ruthardt M, Testa U, Nervi C. *et al.* Opposite effects of the acute promyelocytic leukemia PML-retinoic acid receptor alpha (RAR alpha) and PLZF-RAR alpha fusion proteins on retinoic acid signalling. *Mol Cell Biol* 1997;**17**:4859–69. https://doi.org/10.1128/MCB.17.8.4859.
16. Aibar S, González-Blas CB, Moerman T. *et al.* SCENIC: single-cell regulatory network inference and clustering. *Nat Methods* 2017;**14**:1083–6. https://doi.org/10.1038/nmeth.4463.

17. Naldi A, Hernandez C, Abou-Jaoudé W. *et al.* Logical Modeling and analysis of cellular regulatory networks with GINsim 3.0. *Front Physiol* 2018;**9**:646. https://doi.org/10.3389/fphys.2018.00646.
18. Naldi A. BioLQM: a java toolkit for the manipulation and conversion of logical qualitative models of biological networks. *Front Physiol* 2018;**9**:1605. https://doi.org/10.3389/fphys.2018.01605.
19. Drumea K, Yang ZF, Rosmarin A. Retinoic acid signaling in myelopoiesis. *Curr Opin Hematol* 2008;**15**:37–41. https://doi.org/10.1097/MOH.0b013e3282f20a9c.
20. Ablain J, Rice K, Soilihi H. *et al.* Activation of a promyelocytic leukemia-tumor protein 53 axis underlies acute promyelocytic leukemia cure. *Nat Med* 2014;**20**:167–74. https://doi.org/10.1038/nm.3441.
21. Klarner H, Streck A, Siebert H. PyBoolNet: a python package for the generation, analysis and visualization of boolean networks. *Bioinformatics* 2017;**33**:770–2. https://doi.org/10.1093/bioinformatics/btw682.
22. Klarner H, Heinitz F, Nee S. *et al.* Basins of attraction, commitment sets, and phenotypes of Boolean networks. *IEEE/ACM Trans Comput Biol Bioinform* 2020;**17**:1115–24. https://doi.org/10.1109/TCBB.2018.2879097.
23. Stoll G, Caron B, Viara E. *et al.* MaBoSS 2.0: an environment for stochastic Boolean modeling. *Bioinformatics* 2017;**33**:2226–8. https://doi.org/10.1093/bioinformatics/btx123.
24. Tanaka S, Miyagi S, Sashida G. *et al.* Ezh2 augments leukemogenicity by reinforcing differentiation blockage in acute myeloid leukemia. *Blood* 2012;**120**:1107–17. https://doi.org/10.1182/blood-2011-11-394932.
25. Beal J, Montagud A, Traynard P. *et al.* Personalization of logical models with multi-omics data allows clinical stratification of patients. *Front Physiol* 2019;**9**:1965. https://doi.org/10.3389/fphys.2018.01965.
26. Lei A, Chen L, Zhang M. *et al.* EZH2 regulates protein stability via recruiting USP7 to mediate neuronal gene expression in cancer cells. *Front Genet* 2019;**10**:422. https://doi.org/10.3389/fgene.2019.00422.
27. Wang L, Chen C, Song Z. *et al.* EZH2 depletion potentiates MYC degradation inhibiting neuroblastoma and small cell carcinoma tumor formation. *Nat Commun* 2022;**13**:12. https://doi.org/10.1038/s41467-021-27609-6.
28. Fasan A, Haferlach C, Perglerovà K. *et al.* Molecular landscape of acute promyelocytic leukemia at diagnosis and relapse. *Haematologica* 2017;**102**:e222–4. https://doi.org/10.3324/haematol.2016.162206.
29. Thomas R, D'Ari R. Biological feedback. Boca Raton: CRC Press, 1990.
30. Bérenguier D, Chaouiya C, Monteiro PT. *et al.* Dynamical modeling and analysis of large cellular regulatory networks. *Chaos* 2013;**23**:025114. https://doi.org/10.1063/1.4809783.
31. Naldi A, Remy E, Thieffry D. *et al.* Dynamically consistent reduction of logical regulatory graphs. *Theor Comput Sci* 2011;**412**:2207–18. https://doi.org/10.1016/j.tcs.2010.10.021.

Briefings in Bioinformatics, 2025, **26(1)**, bbaf002
https://doi.org/10.1093/bib/bbaf002
Case Study