



INSTITUT DE FRANCE
Académie des sciences

Comptes Rendus

Biologies

Pauline Gloaguen, Yves Vandenbrouck, Jacques Joyard and Gilles Curien

**ChloroKB, a cell metabolism reconstruction of the model plant
*Arabidopsis thaliana***

Volume 344, issue 2 (2021), p. 157-163

Published online: 2 July 2021

<https://doi.org/10.5802/crbio.49>



This article is licensed under the
CREATIVE COMMONS ATTRIBUTION 4.0 INTERNATIONAL LICENSE.
<http://creativecommons.org/licenses/by/4.0/>



Les Comptes Rendus. Biologies sont membres du
Centre Mersenne pour l'édition scientifique ouverte
www.centre-mersenne.org
e-ISSN : 1768-3238



Articles / Reviews / *Articles / Revues*

ChloroKB, a cell metabolism reconstruction of the model plant *Arabidopsis thaliana*

ChloroKB, une reconstruction du métabolisme cellulaire de la plante modèle Arabidopsis thaliana

Pauline Gloaguen^a, Yves Vandenbrouck^b, Jacques Joyard^{*, a} and Gilles Curien^a

^a Université Grenoble Alpes, CNRS, CEA, INRAe, IRIG-LPCV, UMR5168, 38000 Grenoble, France

^b Université Grenoble Alpes, Inserm, CEA, IRIG-BGE, U1038, 38000, Grenoble, France

E-mails: pauline.gloagen@cea.fr (P. Gloaguen), yves.vandenbrouck@cea.fr (Y. Vandenbrouck), jacques.joyard@univ-grenoble-alpes.fr (J. Joyard), gilles.curien@cea.fr (G. Curien)

Jacques Joyard was the award winner of the 2020 prize Georges Morel from the Académie des sciences / Jacques Joyard a été lauréat du prix 2020 Georges Morel de l'Académie des sciences

Abstract. Can we understand how plant cell metabolism really works? An integrated large-scale modelling of plant metabolism predictive model would make possible to analyse the impact of disturbances in environmental conditions on cellular functioning and diversity of plant-made molecules of interest. ChloroKB, a Web application initially developed for exploration of *Arabidopsis* chloroplast metabolic network now covers *Arabidopsis* mesophyll cell metabolism. Interconnected metabolic maps show subcellular compartments, metabolites, proteins, complexes, reactions, and transport. Data in ChloroKB have been structured to allow for mathematical modelling and will be used as a reference for modelling work dedicated to a particular issue.

Résumé. Peut-on comprendre comment fonctionne réellement le métabolisme des cellules végétales? Un modèle prédictif intégré à grande échelle du métabolisme des plantes permettrait d'analyser l'impact des perturbations des conditions environnementales sur le fonctionnement cellulaire et la diversité des molécules d'intérêt fabriquées par les plantes. ChloroKB, une application Web initialement développée pour l'exploration du réseau métabolique du chloroplaste d'*Arabidopsis*, couvre désormais le métabolisme des cellules du mésophylle d'*Arabidopsis*. Des cartes métaboliques interconnectées décrivent les compartiments subcellulaires, les métabolites, les protéines, les complexes, les réactions et le transport. Les données de ChloroKB ont été structurées pour permettre la modélisation mathématique et seront utilisées comme référence pour les travaux de modélisation consacrés à une question particulière.

* Corresponding author.

Keywords. Knowledge base, Metabolism, Plant cell, Chloroplast, Modelling.

Mots-clés. Base de connaissances, Métabolisme, Cellule végétale, Chloroplaste, Modélisation.

Manuscript received and accepted 26th April 2021.

Will we one day be able to model plant metabolism as we can model the flight of an aircraft? Beyond the obvious advantage of large-scale modelling of plant metabolism for the optimised production of molecules of interest, for example, having an integrated predictive model at our disposal would allow us to totally change the way we can validate our knowledge. Such a validated model would mean that we really understood how plant cell metabolism works. It would then be possible to analyse the impact of disturbances in environmental conditions on cellular functioning, an essential preliminary step for a detailed understanding of how plants live in their environment. The task seems immense because plant cell metabolism is certainly the most complex eukaryotic metabolism to understand given the nature of the energy source used (light) and the existence of an additional compartment that uses and controls this energy (the chloroplast). Endocytosis of a photosynthetic microorganism 1.5 billion years ago by a heterotrophic eukaryotic host not only added a subcellular compartment but also completely reorganised the metabolic flows within the host cell and the transport reactions between this compartment and the other cell compartments [1]. It is therefore not possible to use bacterial or yeast models for plant cell metabolism modelling. The situation is made even more complex by the fact that higher plants are made of about ten different cell types with quite different metabolisms in leaves, roots, meristems, phloem, pollen, fruits, seeds, etc. Upstream of predictive modelling, it is therefore necessary to collect experimental data, reference them, compare them and transform this knowledge into an “entity” that can be analysed by both humans and computers.

This work has been carried out for about ten years for the model plant *Arabidopsis* and is available online (<http://chlorokb.fr>). Initially ChloroKB (for Chloroplast Knowledge Base) focused on chloroplast metabolism [2]: it is thus possible to explore the biosynthesis pathways for amino acids (e.g., see *L_branched_Chain_Aminoacids*), lipids (e.g., see *Galactolipid synthesis*; [3]), or pigments (chlorophylls, carotenoids [4], etc.; e.g., see Chloro-

phyll_synthesis), their subcellular location (within the thylakoids, stroma and envelope membranes) and their interactions with other cell compartments (Figure 1) [5].

With recent updates, the reconstructed network now covers a very large part of *Arabidopsis* mesophyll cell metabolism with nearly 1600 proteins and more than 1000 metabolites. This network is presented in the form of interconnected metabolic maps (currently 189) showing subcellular compartments, metabolites, proteins, complexes, reactions, and transport. Each “object” (protein, transporter, complex) is clickable, allowing the user to access the data available online but scattered throughout the literature. Currently this reconstructed network is a tool for exploring metabolism which very quickly provides access to integrated knowledge of a particular metabolism (e.g., see *Ascorbate glutathion pathway*) (Figure 2).

ChloroKB differs from other metabolic databases (BioCYC, Uniprot, KEGG) by (i) its high level of data integration, (ii) the consideration of subcellular compartmentalisation and thus metabolite transport, (iii) the representation of “fine” controls such as complex formation, activity modification by allosteric control or phosphorylation (e.g., see *PEP carboxylase regulation*). Sub-cellular localisation has been improved compared to the version initially published with a coding which now considers the different cellular membranes and inter-membrane spaces (plasma membrane, endoplasmic reticulum, inner and outer membranes of mitochondria and chloroplast). It will thus eventually be possible to model membrane lipid metabolism and intermembrane traffic. The use of CellDesigner [6] as a standard for the representation of biochemical networks facilitates the reading of the network and communication, especially with young researchers, who must in a short period of time acquire knowledge accumulated over more than fifty years. In addition to this visual aspect, which allows ChloroKB to be used as a memory aid, as a source of expert information or as an interactive exploration tool, ChloroKB’s data have been structured in such a way as to allow for mathematical modelling. Flux balance analysis modelling

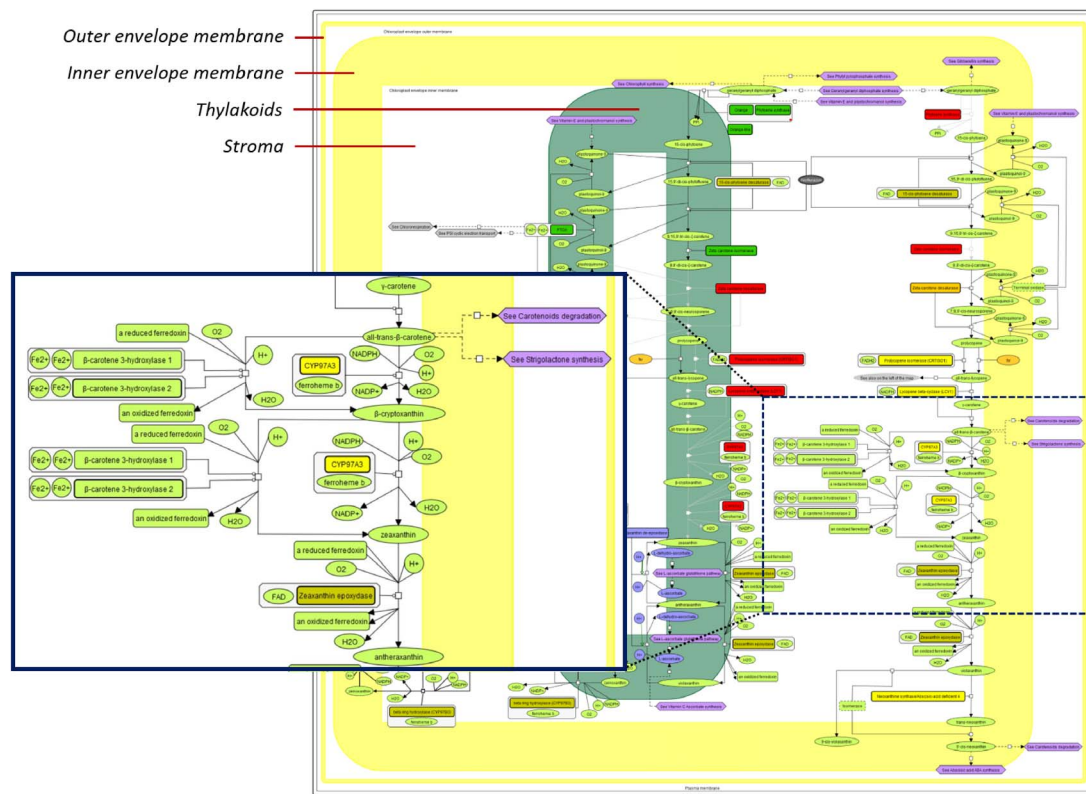


Figure 1. Exploration of carotenoid metabolism in chloroplasts. The different colours correspond to experimental evidence for protein localisation in the envelope (yellow), or the thylakoid (green) or absence of evidence (red). Demonstrated reactions are coded in black and hypothetical reactions in grey. The violin icons allow navigation from map to map. Zooming option in the online version facilitates the understanding.

for example allows prediction of steady-state fluxes distribution at the system level under some assumptions (for example optimization of biomass production or of a specific flux) [7]. This kind of modelling requires a stoichiometric matrix of the network. A stoichiometric matrix captures mathematically the structure of a network and consists in a table with reactions in columns and metabolites in rows containing +1 or -1 or 0 or other values depending on

French version

Pourrons-nous un jour modéliser le métabolisme d'une plante comme nous sommes désormais capables de modéliser le vol d'un avion? Au-delà des avantages évidents d'une modélisation à grande échelle du métabolisme de la plante pour, par

the amount of each metabolites used in each reaction or transport steps of the network. A complete stoichiometric matrix of the network represented in ChloroKB will soon be available online, which can be used as a reference for modelling work dedicated to a particular issue [8–10]. Finally, a large amount of quantitative data (metabolite concentrations) is available which will be useful to constrain metabolic models of the Flux balance analysis type.

exemple, la production optimisée de molécules d'intérêt, disposer d'un modèle prédictif intégré permettrait de changer totalement notre manière de valider nos connaissances. Un tel modèle signifierait que nous avons réellement compris comment ce

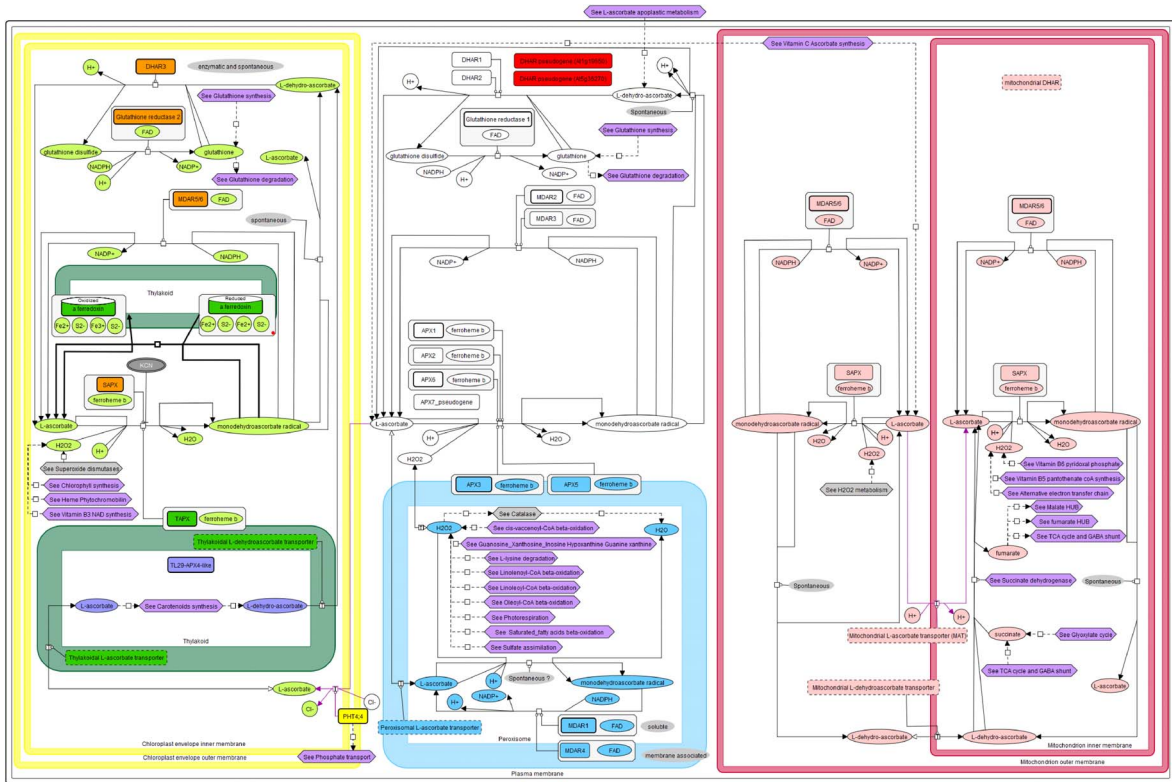


Figure 2. Exploration of ascorbate-glutathione pathway in photosynthetic *Arabidopsis* cells. A complete overview of this metabolism is provided across 6 subcellular compartments, namely thylakoid (green) plastid (yellow), peroxisome (blue), mitochondrion intermembrane space and matrix (red) and cytosol (white). Apoplactic L-ascorbate metabolism is represented in another map (top violin icon).

métabolisme fonctionne. Il serait alors envisageable d’analyser la réponse cellulaire aux modifications des conditions environnementales, étape préliminaire essentielle à la compréhension fine de la vie de la plante dans son environnement. La tâche semble immense, car le métabolisme de la cellule végétale est certainement le métabolisme eucaryote le plus complexe à appréhender, étant donné la nature de la source d’énergie utilisée (la lumière) et l’existence d’un compartiment supplémentaire, qui utilise et contrôle cette énergie (le chloroplaste). L’endocytose d’un microorganisme photosynthétique il y a un 1,5 milliard d’années par un hôte eucaryote hétérotrophe a non seulement ajouté un compartiment subcellulaire, mais a totalement réorganisé les flux métaboliques au sein de la cellule hôte et les réactions de transports entre ce compartiment et les autres compartiments [1]. Il n’est donc pas possible d’utiliser des modèles de bactéries ou de levures

pour modéliser le métabolisme de la cellule végétale. La situation est rendue encore plus complexe par le fait que les plantes supérieures sont constituées d’une dizaine de types cellulaires aux métabolismes très différents (cellules chlorophylliennes, cellules racinaires, cellule de phloème, pollen, cellules méristématiques, etc.). En amont de la modélisation prédictive, il est donc nécessaire de collationner les données expérimentales, et de transformer cette connaissance en une « entité » lisible à la fois par l’être humain et par l’ordinateur. C’est ce travail qui a été effectué depuis une dizaine d’années pour la plante modèle *Arabidopsis thaliana* et qui est disponible en ligne (<http://chlorokb.fr>). Initialement ChloroKB (pour *Chloroplast Knowledge Base*) était centré sur le métabolisme du chloroplaste [2] : il est ainsi possible (i) d’explorer les fonctions chloroplastiques et en particulier les voies de biosynthèse des acides

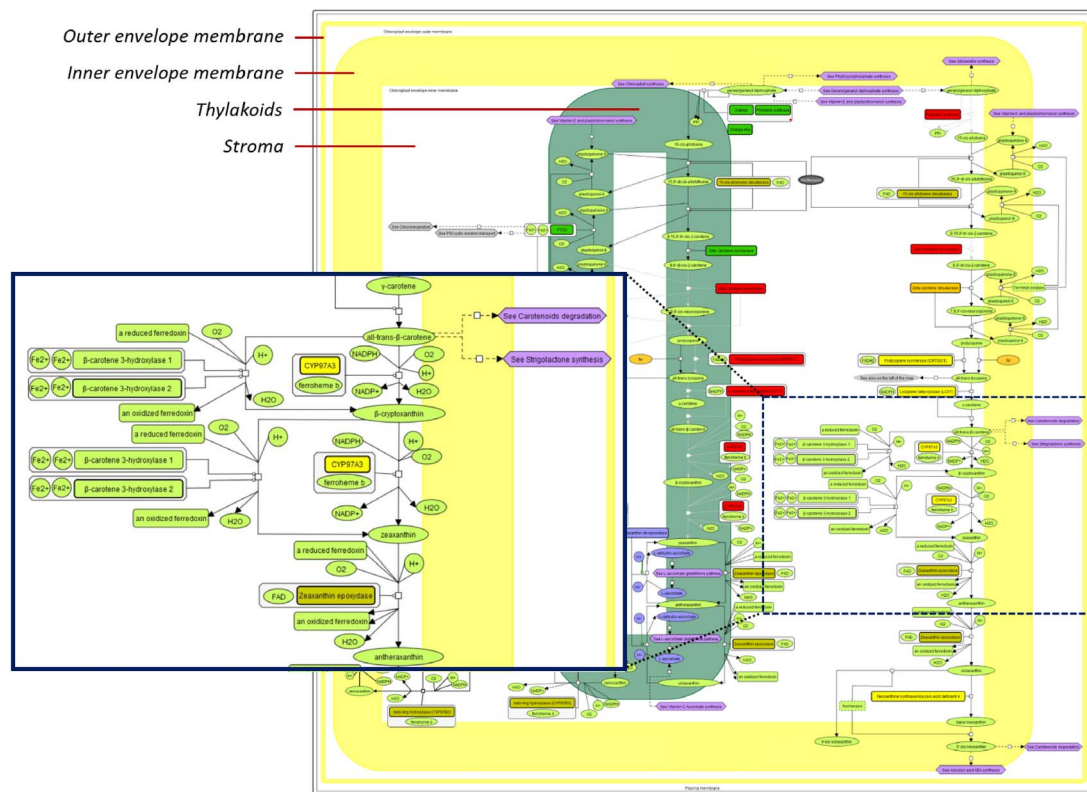


FIGURE 1. Exploration du métabolisme des caroténoïdes dans les chloroplastes. Les différentes couleurs correspondent aux preuves expérimentales de la localisation des protéines dans l’enveloppe (jaune), ou le thylakoïde (vert) ou à l’absence de preuves (rouge). Les réactions démontrées sont en noir et les réactions hypothétiques en gris. Les icônes violettes permettent de naviguer de carte en carte. L’option de zoom dans la version en ligne facilite la compréhension.

aminés (voir L_branched_Chain_Aminoacids), des lipides (voir Galactolipid synthesis; [3]), ou des pigments (chlorophylles, caroténoïdes... ; e.g. voir Chlorophyll_synthesis, [4]) ainsi que (ii) leur localisation subcellulaire (au sein des thylakoïdes, du stroma et de l’enveloppe) et (iii) leurs interactions avec l’ensemble des compartiments cellulaires (Figure 1) [5].

Avec ses mises à jour récentes, le réseau reconstruit de la cellule chlorophyllienne de *Arabidopsis thaliana* couvre désormais une très grande partie du métabolisme cellulaire. Il se présente sous la forme de cartes métaboliques interconnectées (actuellement au nombre de 189) sur lesquelles figurent les compartiments subcellulaires, les métabolites, les protéines, les complexes, les réactions et les transports. Chaque « objet » (protéine, transporteur, com-

plexe) est cliquable, permettant à l’utilisateur d’accéder aux données certes disponibles en ligne mais dispersées dans la littérature. Actuellement ce réseau reconstruit est un outil d’exploration du métabolisme qui permet très rapidement d’avoir accès à une connaissance intégrée d’un métabolisme particulier (e.g. voir Ascorbate glutathion pathway) (Figure 2).

ChloroKB se distingue des autres bases de données métaboliques (BioCYC, Uniprot, KEGG) par (i) son niveau élevé d’intégration des données, (ii) la prise en compte de la compartimentation subcellulaire et donc du transport de métabolites, (iii) la représentation des contrôles « fins » comme la formation de complexes, la modification d’activité par contrôle allostérique ou par phosphorylation (e.g. voir PEP carboxylase regulation). Le niveau de pré-

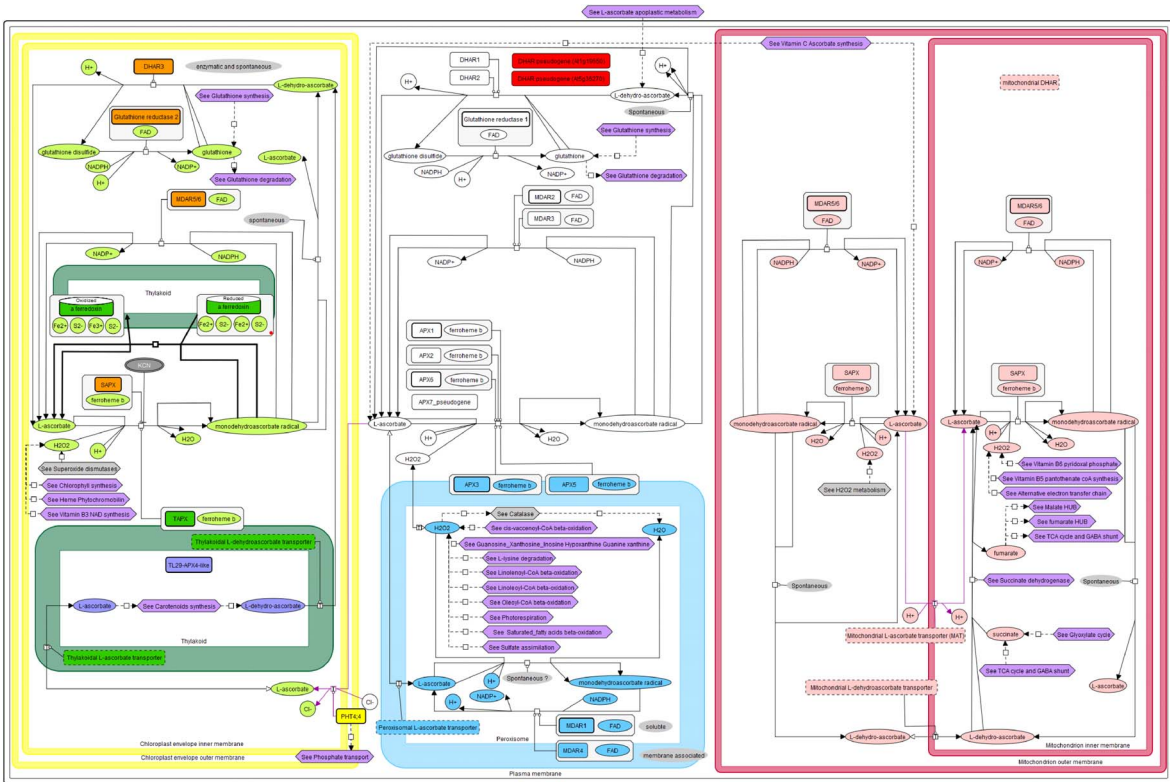


FIGURE 2. Exploration de la voie glutathion-ascorbate dans les cellules photosynthétiques d'*Arabidopsis*. Un aperçu complet de ce métabolisme est fourni à travers 6 compartiments subcellulaires, à savoir les thylakoïdes (vert), le plaste (jaune), le peroxyssome (bleu), l'espace intermembranaire et la matrice de la mitochondrie (rouge) et le cytosol (blanc). Le métabolisme apoplastique du L-ascorbate est représenté sur une autre carte (icône violette en haut).

cision des localisations a été amélioré par rapport à la version initialement publiée avec un codage qui prend désormais en compte les différentes membranes et les espaces intermembranaires (membrane plasmique, réticulum endoplasmique, membranes internes et externes de la mitochondrie et du chloroplaste). Il sera ainsi à terme possible de modéliser le métabolisme lipidique membranaire et le trafic intermembranaire. L'utilisation de CellDesigner [6] comme standard de représentation des réseaux biochimiques facilite la lecture du réseau ainsi que la communication, en particulier avec les jeunes chercheurs, qui doivent en peu de temps acquérir une connaissance accumulée pendant plus de cinquante ans. Au-delà de cet aspect visuel qui permet d'utiliser ChloroKB comme aide-mémoire, comme source

d'informations expertisées ou comme outil d'exploration interactive, les données de ChloroKB ont été structurées de manière à permettre une modélisation mathématique. Une méthode de modélisation appelée « Flux balance analysis » permet par des méthodes de programmation linéaire de calculer la distribution des flux à l'état stationnaire dans le système métabolique sous certaines hypothèses d'optimisation (de la biomasse par exemple ou d'un flux de métabolite particulier) [7]. Ce type de modélisation requiert une abstraction du réseau qui se présente sous la forme d'une matrice de stœchiométrie (un tableau contenant les réactions en colonnes et les métabolites en ligne). L'occurrence et la stœchiométrie des réactions et des transports sont représentées par des 0, +1 or -1, +2 etc. Une matrice de

stœchiométrie complète du réseau reconstruit sera d'ici peu disponible en ligne sur le site de ChloroKB et pourra servir de référence pour des travaux de modélisation dédiés à une question particulière [8–10]. Enfin de très nombreuses données expérimentales quantitatives (concentrations de métabolites) sont accessibles et seront utiles pour contraindre les modèles métaboliques de type « Flux balance analysis ».

References

- [1] M.-A. Selosse, J. Joyard, “Symbiose et évolution : à l'origine de la cellule eucaryote”, 2021, Encyclopédie de l'Environnement, [en ligne ISSN 2555-0950]: <https://www.encyclopedie-environnement.org/vivant/symbiose-evolution-lorigine-de-cellule-eucaryote/>.
- [2] P. Gloaguen, S. Bournais, C. Alban, S. Ravel, D. Seigneurin-Berny, M. Matringe, M. Tardif, M. Kuntz *et al.*, “ChloroKB: A Web application for the integration of knowledge related to chloroplast metabolic network”, *Plant Physiol.* **174** (2017), p. 922-934.
- [3] J. Joyard, M. Ferro, C. Masselon, D. Seigneurin-Berny, D. Salvi, J. Garin, N. Rolland, “Chloroplast proteomics and the compartmentation of plastidial isoprenoid biosynthetic pathways”, *Mol. Plant.* **2** (2009), p. 1154-1180.
- [4] J. Joyard, M. Ferro, C. Masselon, D. Seigneurin-Berny, D. Salvi, J. Garin, N. Rolland, “Chloroplast proteomics highlights the subcellular compartmentation of lipid metabolism”, *Prog. Lipid Res.* **49** (2010), p. 128-1385.
- [5] N. Rolland, G. Curien, G. Finazzi, M. Kuntz, E. Maréchal, M. Matringe, S. Ravel, D. Seigneurin-Berny, “The biosynthetic capacities of the plastids and integration between cytoplasmic and chloroplast processes”, *Annu. Rev. Genet.* **46** (2012), p. 233-264.
- [6] A. Funahashi, Y. Matsuoka, A. Jouraku, M. Morohashi, N. Kikuchi, H. Kitano, “CellDesigner 3.5: A versatile modeling tool for biochemical networks”, *Proc. IEEE* **96** (2008), no. 8, p. 1254-1265.
- [7] L. J. Sweetlove, R. G. Ratcliffe, “Flux-balance modeling of plant metabolism”, *Front. Plant Sci.* **2** (2011), article no. 38.
- [8] A. Arnold, Z. Nikoloski, “Bottom-up Metabolic Reconstruction of Arabidopsis and Its application to determining the metabolic costs of enzyme production”, *Plant Physiol.* **165** (2014), p. 1380-1391.
- [9] S. Shameer, R. G. Ratcliffe, L. J. Sweetlove, “Leaf energy balance requires mitochondrial respiration and export of chloroplast NADPH in the light”, *Plant Physiol.* **180** (2019), p. 1947-1961.
- [10] M. A. Blatke, A. Brautigam, “Evolution of C4 photosynthesis predicted by constraint-based modelling”, *Elife* **8** (2019), article no. e49305.